

Санкт-Петербургский государственный университет

Факультет прикладной математики —
процессов управления



**ПРОЦЕССЫ УПРАВЛЕНИЯ
И УСТОЙЧИВОСТЬ**

ТРУДЫ XII МЕЖДУНАРОДНОЙ НАУЧНОЙ
КОНФЕРЕНЦИИ АСПИРАНТОВ И СТУДЕНТОВ

Санкт-Петербург
4–7 апреля 2011 года



ИЗДАТЕЛЬСТВО
САНКТ-ПЕТЕРБУРГСКОГО УНИВЕРСИТЕТА

Санкт-Петербургский государственный университет

Факультет прикладной математики –
процессов управления



***ПРОЦЕССЫ УПРАВЛЕНИЯ
И УСТОЙЧИВОСТЬ***

ТРУДЫ XLII МЕЖДУНАРОДНОЙ НАУЧНОЙ
КОНФЕРЕНЦИИ АСПИРАНТОВ И СТУДЕНТОВ

Санкт-Петербург
4 – 7 апреля 2011 года



ИЗДАТЕЛЬСТВО
САНКТ-ПЕТЕРБУРГСКОГО УНИВЕРСИТЕТА
2011

УДК 517.51:517.9:519.71
ББК 22.1
П84

Рецензенты: д-р физ.-мат. наук, проф. *В. Л. Харитонов* (С.-Петербург. гос. ун-т),
д-р техн. наук, проф. *Ф. М. Кулаков* (С.-Петербургский институт
информатики и автоматизации РАН)

*Печатается по постановлению
Редакционно-издательского совета
факультета прикладной математики – процессов управления
Санкт-Петербургского государственного университета*

Процессы управления и устойчивость: Труды 42-й междуна-
П84 родной научной конференции аспирантов и студентов / Под ред.
А. С. Ерёмкина, Н. В. Смирнова. – СПб.: Издат. Дом С.-Петерб. гос.
ун-та, 2011. – 588 с.
ISBN 978-5-288-05197-5

В сборник включены работы студентов, аспирантов и сотрудников факультета прикладной математики – процессов управления СПбГУ и других высших учебных заведений, в том числе зарубежных, по математической теории процессов управления, математическим методам в механике и физике, математическому моделированию в медико-биологических системах, информационным и компьютерным технологиям, теории управления социально-экономическими системами.

Все работы докладывались на ежегодной 42-й международной научной конференции аспирантов и студентов «Процессы управления и устойчивость» (4–7 апреля 2011 года).

Сборник предназначен для студентов старших курсов физико-математических факультетов, аспирантов и научных работников.

ББК 22.1

ISBN 978-5-288-05197-5

© Авторы сборника, 2011
© Факультет ПМ – ПУ
С.-Петербургского государственного университета

Содержание

1. Математическая теория процессов управления.....	9
<i>Долгополж М.В.</i> Кодифференцируемые функции в нормированных пространствах	9
<i>Жабко А.П., Зараник У.П.</i> Область асимптотической устойчивости управляемых систем с запаздыванием	15
<i>Казанцева Т.Е.</i> О поведении одной динамической системы с двумя параметрами	18
<i>Кузнецова А.С.</i> Синхронизация в осцилляторах, соединенных через обратную связь	24
<i>Любич С.Я.</i> Математическая модель оптимального планирования расписания движения поездов	29
<i>Медведева И.В.</i> Модификация алгебраического метода исследования устойчивости дифференциально-разностных уравнений	35
<i>Морозова Д.Н.</i> Синтез линейных управлений с линейным запаздыванием	41
<i>Нгуен Динь Хуен.</i> Об асимптотической устойчивости некоторого класса нелинейных неавтономных разностных систем	47
<i>Пономарев А.А.</i> О выборе параметров метода «предиктор-корректор»	52
<i>Резлясова Л.В.</i> Вычисление \mathcal{H}_2 нормы передаточной матрицы системы с запаздыванием	58
<i>Сужач М.П.</i> Пример решения задачи двухуровневой оптимизации	63
<i>Сумачева В.А.</i> Вычисление нормы передаточной функции уравнения с запаздываниями с помощью функций Ляпунова ..	68
<i>Утешев Н.А.</i> Идентификация методом изоляции с помощью сферы	74
2. Математические методы в механике и физике.....	81
<i>Алцубеев В.В.</i> Параметрическая оптимизация динамики пучков в гибридных системах	81
<i>Басинский К.Ю.</i> Распространение капиллярно-гравитационных волн по свободной поверхности слабовязкой жидкости	86
<i>Белова О.А.</i> Об одном семействе траекторий соударения	92
<i>Большаков А.И.</i> Алгоритм выбора шага в методах плотной выдачи	98

<i>Брэгман К.М.</i> Сведение дифференциальных уравнений к полиномиальной системе	103
<i>Викулина Ю.И., Греков М.А.</i> Напряженно-деформированное состояние полуплоскости при учете поверхностного напряжения	109
<i>Витвицкий С.И., Нестеров А.В., Нестеров С.В.</i> Анализ устойчивых решений уравнения Матье при моделировании периодически нестационарного объекта управления	115
<i>Дронь В.О.</i> О минимальных сплайнах третьего порядка на замкнутом промежутке	121
<i>Иванов А.Н.</i> Символьные вычисления в моделировании динамики пучков заряженных частиц	127
<i>Иванов В.А.</i> Двухкомпонентная плоскость с криволинейной трещиной в окрестности межфазной линии	133
<i>Климаков А.А.</i> Моделирование электростатической линзы	139
<i>Кодрик В.В., Кутузов С.А.</i> Моделирование внешнего поля Земли с помощью гравитирующих стержней	145
<i>Котов П.А.</i> О динамических погрешностях диссипативных динамических систем	151
<i>Курушина С.Е., Желнов Ю.В., Максимов В.В.</i> Фазовый переход «беспорядок – порядок – беспорядок» в стохастических системах реакция-диффузия	158
<i>Михеев В.С., Михеев С.Е.</i> Интерполяция по линейчатой таблице	165
<i>Подзвальный Е.А.</i> Метод глобально-оптимального поиска структуры канала согласования	171
<i>Прокопенко А.В.</i> Траектории звезд в модели галактики, допускающей три интеграла движения	177
<i>Пронин А.О.</i> Об устойчивости механического поведения нелинейно-упругих тел	183
<i>Степанова В.А.</i> Нелинейная задача для плоскости с прямолинейной межфазной трещиной	189
<i>Телевный Д.С.</i> Моделирование триодной системы с полевым острием	195
<i>Уланов Е.А., Утешев А.Ю.</i> Аналитическое решение обобщенной задачи Ферма – Торричелли – Штейнера	201
<i>Целищев С.О., Еремин А.С.</i> Об устойчивости структурного метода решения систем ОДУ	207
<i>Шиманчук Д.В., Белова О.А.</i> Стабилизация орбитального	

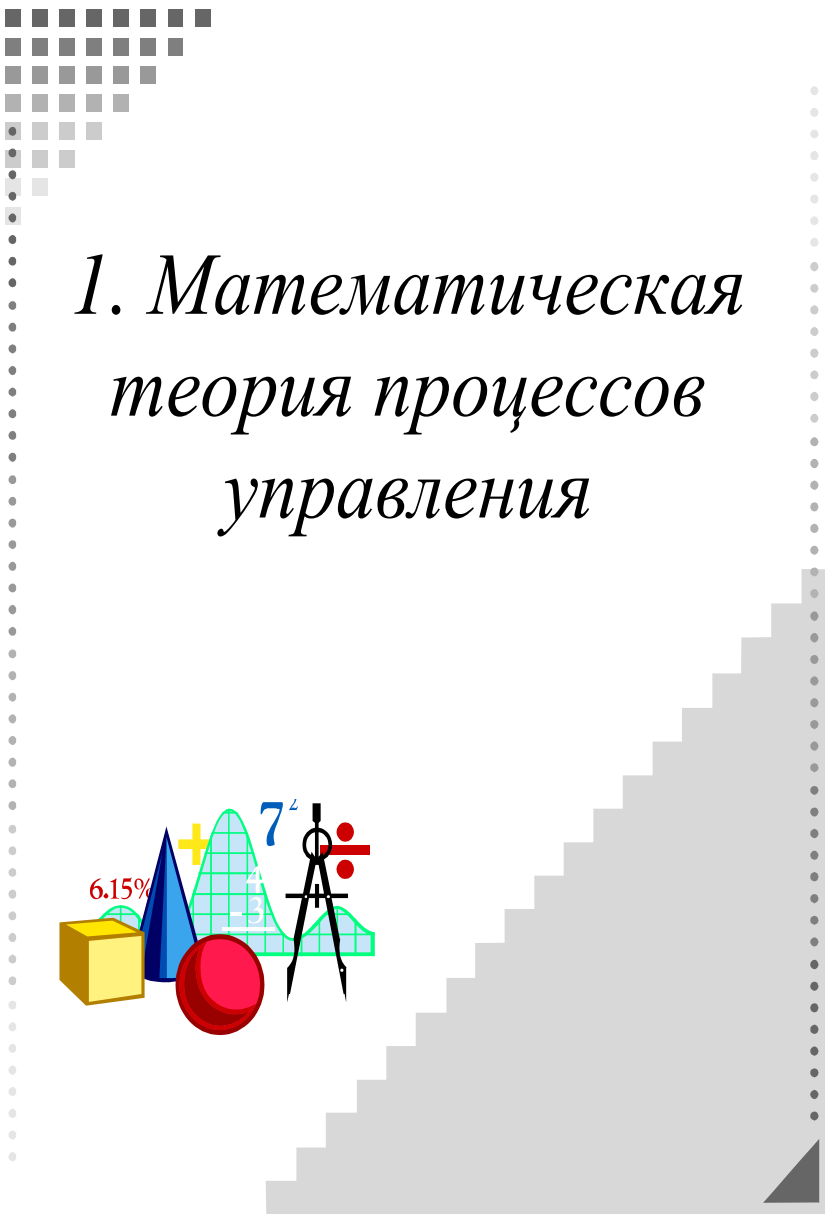
движения космического аппарата в окрестности коллинеарной точки либрации.....	213
<i>Широколов А.Ю.</i> Математическая модель оптимизации параметров ускорителя на бегущей волне в случае кусочно-гладких управлений.....	220
<i>Ярыш Г.В., Пронина Ю.Г.</i> Механохимическая коррозия линейного участка трубопровода.....	226
3. Математические модели медико-биологических систем.....	233
<i>Абдрафиков В.А.</i> Анализ морфологии клеток крови.....	233
<i>Кузнецов С.Р.</i> Пространственная математическая модель развития дисбактериоза на слизистых оболочках человека.....	239
<i>Мамочев В.А.</i> Применение статистического метода в задаче классификации.....	245
<i>Морозов П.Д.</i> Моделирование колебаний базилярной мембраны на основании теории Бекешы.....	249
<i>Морозов П.Д., Михеев В.С.</i> Акустическая модуляция речи ...	255
<i>Чернэуцану Е.К.</i> Строгая h -отделимость двух множеств и линейное программирование.....	259
4. Информационные и компьютерные технологии	265
<i>Антипин И.А., Козыниченко В.А.</i> Об одной модификации нейронной сети Хемминга.....	265
<i>Борисов Е.А.</i> Применение оптического преобразования Фурье к задаче определения ориентации случайно расположенного в кадре текстового блока относительно апертуры объектива..	271
<i>Быков К.Н.</i> Инвариантность калибровочных соотношений для сплайнов эрмитова типа при различном порядке выбора узлов.....	277
<i>Гузницева А.П.</i> Программная реализация позиционного управления нелинейным объектом.....	283
<i>Дрокин И.С.</i> Компьютерный анализ робастных свойств сверхустойчивых замкнутых систем.....	289
<i>Дурновцева С.А.</i> Моделирование исходных сейсмических колебаний грунта для формирования проектных основ.....	295
<i>Журавлев М.М.</i> Применение принципа минимальной длины описания к задаче идентификации механических поврежде-	

ний по изображению	299
<i>Иванова О.А., Пешехонов К.А.</i> Программная реализация метода вычисления эллиптических функций	304
<i>Лобов А.Л.</i> Компьютерные алгоритмы идентификации моделей морских судов	309
<i>Михеев П.А.</i> Индивидуальное быстроедействие абонента беспроводной ЛВС стандарта 802.11	315
<i>Нвохчири А.М.</i> Ранжирование веб-сайтов нигерийских университетов	321
<i>Погожнев С.В., Шарлай А.С.</i> Об определении выделенных формул в изображениях печатного текста	327
<i>Проскура Е.С.</i> Автоматизация получения пользовательского сертификата для мобильного клиента MS Exchange	333
<i>Романовский Л.М.</i> Локальное укрупнение триангуляции и калибровочные соотношения	338
<i>Рябов А.В., Осьминин Л.А.</i> Автоматизация процесса разработки нормативного графика движения поездов	344
<i>Сафонова А.В.</i> Метод машинного обучения с использованием LARS и Lasso в задаче ранжирования поисковой выдачи	350
<i>Севостьянов Р.А.</i> Программное обеспечение для управления движением робота Futaba RBT-1	356
<i>Смирнов М.Н.</i> Алгоритм синтеза управлений, подавляющих ограниченные внешние воздействия на морское судно	362
<i>Федорова М.А.</i> Синтез и компьютерное моделирование астатической системы управления курсом морского судна	368
<i>Хакимьянова Ю.А.</i> Применение подхода BDD к разработке трейдинговых систем	374
<i>Цой О.С.</i> Дискретизирующий алгоритм прокладки курса судна с учетом естественных препятствий	380
<i>Якушкин О.О.</i> Искусственная нейронная сеть для нахождения сегментов линий	385
<i>Li M.H., Klyuev V.V., Wu S.H.</i> Sentence Alignment from Wikipedia as Comparable Corpora by using STF-IDTF	390
<i>Tikhomirov V.A.</i> Spectral image compression using different clustering methods	396
<i>Ueno R., Klyuev V.V.</i> Semantic Search Engine Query Expansion Using WordNet	402

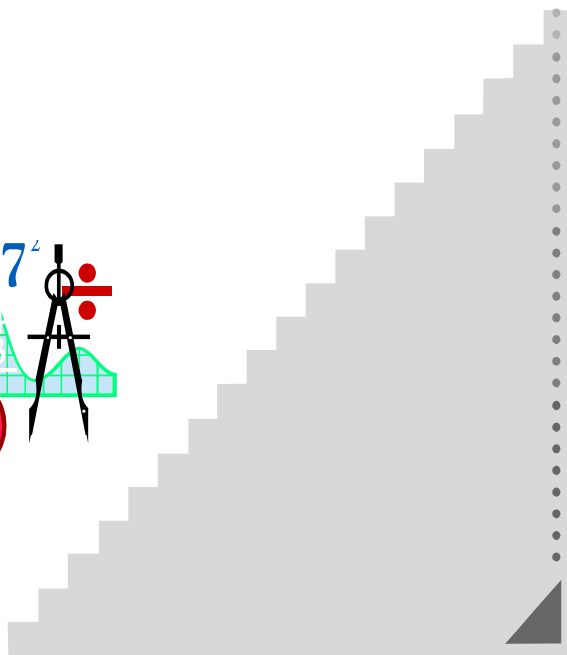
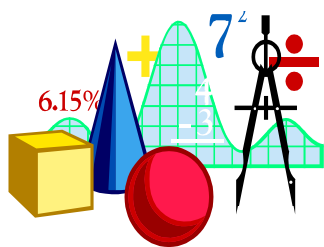
5. Управление социально-экономическими системами 409

<i>Андриенко В.А.</i> Об одной оптимизационной задаче в модели межотраслевого баланса	409
<i>Афанасьева Е.В.</i> Моделирование распределения исследователей по областям науки с помощью вероятностных цепочек ...	416
<i>Барченко М.К.</i> Некрамеровская модель разорения страховой компании	422
<i>Басков О.В.</i> Последовательный алгоритм построения двустепенного конуса и его применение в принятии решений	427
<i>Богданова О.К.</i> Математическое моделирование распространения нового продукта в социальных сетях	432
<i>Бровкина Л.А.</i> Модель прогнозирования объемов перевозок грузов	436
<i>Буре В.М., Сергеева А.А.</i> Оптимальные стратегии выбора логистической фирмы	440
<i>Горбати М.В.</i> Динамическая модель лицензирования на рынке одежды	444
<i>Григорова Т.В.</i> Оптимизация коалиционной структуры в иерархических играх	450
<i>Григорьева К.В., Рединских Н.Д.</i> Абсолютное равновесие в одной игре	456
<i>Громов И.А., Слобожанин Н.М.</i> О распределении пятипараметрического ресурса	461
<i>Захаров А.О.</i> Сужение множества Парето на основе нечеткой замкнутой информации	467
<i>Земцова В.Н.</i> Нечеткая суммарно-разностная функция выбора	473
<i>Зенович О.С., Малафеев О.А.</i> Оптимизационная модель формирования портфеля инвестора на фондовом рынке	479
<i>Зубарева М.Л., Губар Е.А.</i> Применение метода Кларка – Райта для построения инкассаторских маршрутов в сети банкоматов	485
<i>Иванченко Е.В., Колбин В.В.</i> Особенности моделей ценообразования европейского и американского типов опционов в условиях нечетких данных	491
<i>Ильина А.В., Петросян Л.А.</i> Условие Д.В.К. Янга для динамических игр с нетрансферабельными выигрышами	496
<i>Канторович А.В.</i> Теоретико-игровая модель взаимодействия и кооперации в цепи поставок	502

<i>Карпов М.И., Петросян Л.А.</i> Минимизация затрат в транспортной сети	508
<i>Кумачева С.Ш., Чугунникова Н.А.</i> Оценка и прогнозирование валютного риска	514
<i>Мазалова А.В.</i> Парадокс Браесса	519
<i>Мельникова Е.А.</i> Проблема очереди в системе обслуживания: теоретико-игровая модель	522
<i>Мерзлякова Ю.Д., Губар Е.А.</i> Оптимальная политика инкассации банкомата	528
<i>Надеждина А.М.</i> Модели массового обслуживания: call-центры	533
<i>Орлов В.Г.</i> Сети Нэша и максимально эффективные сети	539
<i>Петров А.Г., Слобожанин Н.М.</i> О построении функции потенциала в задаче о кратчайшем пути	545
<i>Свищикова М.В.</i> Один метод прогнозирования ВВП России ..	550
<i>Серяков И.А., Петросян Л.А.</i> Об условно оптимальном перемещении n игроков по непересекающимся путям	555
<i>Степанова О.О.</i> Процесс инвестирования: применение уравнения Гамильтона – Якоби – Беллмана и метода прямых	561
<i>Шижкин В.В.</i> Трехвекторная модель параметрической устойчивости развития предприятия торговли	565
<i>Kostyunin S., Palestini A., Shevkoptyas E.</i> Differential game of resource extraction with random time horizon and different hazard functions	571
Contents	577
Именной указатель	583



1. Математическая теория процессов управления



Долгополик М.В.

Санкт-Петербургский государственный университет

Кодифференцируемые функции в нормированных пространствах¹

Рекомендовано к публикации профессором Демьяновым В.Ф.

1. Введение. В негладком анализе основой для различных аппроксимаций приращения функции как правило является производная по направлениям (см. [1, 2]), либо её аналоги или обобщения, такие как верхняя и нижняя производные Кларка (см. [3]). Существенной проблемой такого подхода является то, что в этом случае аппроксимации являются положительно однородными функциями и не являются непрерывными функциями точки, поскольку производная по направлениям в негладком случае не является непрерывной функцией точки. Непрерывность является весьма важным свойством (в первую очередь, с позиции численных методов) и ради неё можно отказаться от некоторых удобных свойств аппроксимаций, основанных на производной по направлениям. Целесообразным оказалось не вводить какое-либо обобщение производной, а непосредственно аппроксимировать исходную функцию с помощью выпуклых или вогнутых функций. Такой подход, основанный на понятии кодифференцируемости, оказался очень удобным и позволил добиться для достаточно широкого класса негладких функций непрерывности аппроксимации (кодифференциала) как функции точки. Впервые понятие кодифференцируемой функции было введено В.Ф. Демьяновым в работах [4, 5]. Подробное изложение теории кодифференцируемости для случая функций, определённых на \mathbb{R}^n , дано в [2].

В данной работе рассматривается обобщение понятия кодифференцируемости на случай функций, определённых на нормированном пространстве. Описываются исчисление кодифференцируемых функций и необходимые условия экстремума. Изложение опирается на статью [6].

2. Вспомогательные факты. Пусть $(E, \|\cdot\|_E)$ – вещественное нормированное пространство. Обозначим через 0_E нулевой элемент пространства E . Рассмотрим прямое произведение пространств \mathbb{R} и

¹Работа выполнена при финансовой поддержке РФФИ, грант № 09-01-00360

E^* . Пространство $\mathbb{R} \times E^*$, очевидно, является банаховым пространством с нормой

$$\|[a, \varphi]\|_p = (|a|^p + \|\varphi\|_{E^*}^p)^{\frac{1}{p}},$$

где $1 < p < +\infty$. Случай $p = 1$ удобно исключить, поскольку тогда пространство $\mathbb{R} \times E^*$ не будет обладать нужными свойствами. Введём в этом пространстве систему преднорм $\{\|\cdot\|_x \mid x \in E\}$, где

$$\|[a, \varphi]\|_x = |a| + |\varphi(x)|.$$

Это семейство преднорм порождает в пространстве $\mathbb{R} \times E^*$ топологию, которую будем обозначать τ^* . Нетрудно показать, что топология τ^* не сильнее топологии, порождённой нормой. Ясно также, что топология τ^* является хаусдорфовой и операции сложения и умножения на число в $\mathbb{R} \times E^*$ непрерывны относительно этой топологии.

Из определения топологии τ^* легко получить, что топологическое пространство $(\mathbb{R} \times E^*, \tau^*)$ гомеоморфно пространству $(\mathbb{R}, \tau) \times (E^*, w^*)$. Здесь τ – стандартная топология в \mathbb{R} , w^* , как обычно, обозначает $*$ -слабую топологию (эти пространства также топологически изоморфны, как топологические векторные пространства).

Укажем критерий компактности подмножества A пространства $\mathbb{R} \times E^*$ в топологии τ^* .

Теорема 1 [6]. Пусть E – банахово пространство. Для того чтобы подмножество A пространства $\mathbb{R} \times E^*$ было компактным в топологии τ^* , необходимо и достаточно, чтобы оно было ограничено (относительно нормы) и замкнуто в топологии τ^* .

3. Определение кодифференцируемой функции. Пусть везде далее $(E, \|\cdot\|_E)$ – вещественное нормированное пространство, $\Omega \subset E$ – открытое множество.

Определение 1. Функцию $f: \Omega \rightarrow \mathbb{R}$ будем называть кодифференцируемой в точке $x_0 \in \Omega$, если существует пара выпуклых множеств $\underline{df}(x_0) \subset \mathbb{R} \times E^*$ и $\bar{df}(x_0) \subset \mathbb{R} \times E^*$, компактных в топологии τ^* , таких, что для любого допустимого приращения аргумента $\Delta \in E$ (т.е. $\text{co}\{x_0, x_0 + \Delta\} \subset \Omega$) соответствующее приращение функции представимо в виде

$$f(x_0 + \Delta) = f(x_0) + \Phi(\Delta, x_0) + o(\Delta, x_0),$$

где

$$\begin{aligned}\Phi(\Delta, x_0) &= \max_{[a, \varphi] \in \underline{df}(x_0)} (a + \varphi(\Delta)) + \min_{[b, \psi] \in \bar{df}(x_0)} (b + \psi(\Delta)), \\ \Phi(0_E, x_0) &= 0, \quad \frac{o(\alpha\Delta, x_0)}{\alpha} \xrightarrow{\alpha \downarrow 0} 0.\end{aligned}\tag{1}$$

Если (1) имеет место равномерно по $\Delta \in \{\Delta \in E \mid \|\Delta\| = 1\}$, то будем говорить, что функция f кодифференцируема в точке x_0 равномерно по направлениям. Пара множеств $Df(x) = [\underline{df}(x_0), \bar{df}(x_0)]$ называется кодифференциалом функции f в точке x_0 , множество $\underline{df}(x_0)$ – гиподифференциалом, а множество $\bar{df}(x_0)$ – гипердифференциалом. Ясно, что кодифференциал определяется неоднозначно, поскольку, например, пара множеств $[A, B]$, $A = \underline{df}(x_0) - \{[c, \chi]\}$, $B = \bar{df}(x_0) + \{[c, \chi]\}$, где $[c, \chi]$ некоторая фиксированная точка из $\mathbb{R} \times E^*$, также является кодифференциалом функции f в точке x_0 . Если функция f кодифференцируема в некоторой окрестности точки x_0 , то отображение Df назовём кодифференциальным.

Функция f называется непрерывно кодифференцируемой в точке x_0 , если она кодифференцируема в некоторой окрестности точки x_0 и если существует непрерывное (по Хаусдорфу) в этой точке кодифференциальное отображение Df . Функция f называется гиподифференцируемой в точке x_0 , если существует кодифференциал вида $Df(x_0) = [\underline{df}(x_0), \{0_{\mathbb{R} \times E^*}\}]$. Функция f называется гипердифференцируемой в точке x_0 , если существует кодифференциал вида $Df(x_0) = [\{0_{\mathbb{R} \times E^*}\}, \bar{df}(x_0)]$.

Аналогично определяются непрерывно гиподифференцируемые функции и непрерывно гипердифференцируемые функции.

Напомним, что функция $f: \Omega \rightarrow \mathbb{R}$ называется квазидифференцируемой в точке $x \in E$, если она дифференцируема по направлениям в точке x и её производная по направлениям $f'(x, \cdot)$ является конечной величиной и представима в виде суммы субаддитивной и супераддитивной функций.

Утверждение 1 [6]. *Множества кодифференцируемых и квазидифференцируемых в точке x функций совпадают.*

4. Основные формулы кодифференциального исчисления. В этом разделе будем предполагать, что все рассматриваемые функции определены и конечны на открытом множестве $\Omega \subset E$.

Введём операции сложения и умножения на число на множестве, состоящем из пар выпуклых множеств. Пусть $[A, B]$ – пара выпуклых подмножеств вещественного линейного пространства L . Для любого $\lambda \in \mathbb{R}$ положим

$$\lambda[A, B] = \begin{cases} [\lambda A, \lambda B], & \text{если } \lambda \geq 0, \\ [\lambda B, \lambda A], & \text{если } \lambda < 0. \end{cases}$$

Пусть $[C, D]$ – также пара выпуклых множеств из L . Определим

$$[A, B] + [C, D] = [A + C, B + D].$$

Далее будем считать, что везде операции с парами выпуклых множеств, в том числе и с кодифференциалами, осуществляются по этим правилам.

Теорема 2 [6]. Пусть f_1, \dots, f_n – кодифференцируемые в точке $x \in \Omega$ функции, числа $c_1, c_2 \in \mathbb{R}$. Тогда функции $g_1 = c_1 f_1 + c_2 f_2$, $g_2 = f_1 f_2$, $g_3 = \max_{i \in I} f_i$, $g_4 = \min_{i \in I} f_i$, где $I = \{1, \dots, n\}$, кодифференцируемы в точке x , причём их кодифференциалы вычисляются по формулам

$$\begin{aligned} Dg_1(x) &= c_1 Df_1(x) + c_2 Df_2(x), \\ Dg_2(x) &= f_2(x) Df_1(x) + f_1(x) Df_2(x), \\ \underline{d}g_3(x) &= \text{cl co} \left\{ \underline{d}f_k(x) - \sum_{i \in I \setminus \{k\}} \bar{d}f_i(x) + [f_k(x) - f(x), 0_{E^*}] \mid k \in I \right\}, \\ \bar{d}f(x) &= \sum_{i \in I} \bar{d}f_i(x), \quad \underline{d}g(x) = \sum_{i \in I} \underline{d}f_i(x), \\ \bar{d}g(x) &= \text{cl co} \left\{ \bar{d}f_k(x) - \sum_{i \in I \setminus \{k\}} \underline{d}f_i(x) + [f_k(x) - g(x), 0_{E^*}] \mid k \in I \right\}. \end{aligned}$$

Здесь замыкание берётся в топологии τ^* .

Если $f_1(x) \neq 0$ и функция f_1 кодифференцируема в точке x равномерно по направлениям, то и функция $g_5 = 1/f_1$ кодифференцируема в точке x и её кодифференциал вычисляется по формуле

$$Dg_5(x) = -\frac{1}{f_1^2(x)} Df_1(x).$$

Рассмотрим вопрос о кодифференцируемости суперпозиции функций. Пусть E, T – банаховы пространства, точка $y_0 \in T$. Пусть также функция F , действующая из некоторой окрестности точки y_0 в E , дифференцируема по Фреше в точке y_0 . Обозначим $x_0 = F(y_0) \in E$. Предположим, что функция f определена в некоторой окрестности точки x_0 ($f: U(x_0) \rightarrow \mathbb{R}$, где $U(x_0)$ – окрестность точки x_0) и кодифференцируема в этой точке. Пусть в некоторой окрестности точки y_0 имеет смысл суперпозиция $G(\cdot) = f(F(\cdot))$.

Теорема 3 [6]. *Пусть функция f кодифференцируема в точке x_0 равномерно по направлениям. Тогда функция G кодифференцируема в точке y_0 .*

Пусть E_1, \dots, E_n – банаховы пространства, $x_1 \in E_1, \dots, x_n \in E_n$. Предположим, что функция F_i определена в некоторой окрестности точки x_i и кодифференцируема в точке x_i , где $i \in \{1, \dots, n\}$. Обозначим $y_0 = (F_1(x_1), \dots, F_n(x_n)) \in \mathbb{R}^n$. Предположим также, что функция f определена в некоторой окрестности точки y_0 и кодифференцируема в точке y_0 . Пусть в некоторой окрестности точки $x = (x_1, \dots, x_n) \in E_1 \times \dots \times E_n$ имеет смысл суперпозиция $G(\cdot, \dots, \cdot) = f(F_1(\cdot), \dots, F_n(\cdot))$.

Теорема 4 [6]. *Пусть функции F_i и функция f кодифференцируемы равномерно по направлениям. Тогда функция G кодифференцируема в точке (x_1, \dots, x_n) .*

5. Необходимые условия экстремума. Справедлива следующая

Теорема 5. *Пусть $(E, \|\cdot\|_E)$ – банахово пространство, $\Omega \subset E$ – открытое множество, функция $f: \Omega \rightarrow \mathbb{R}$ кодифференцируема на множестве Ω . Для того чтобы точка $x^* \in \Omega$ была точкой локального минимума (максимума) функции f , необходимо, чтобы*

$$\begin{aligned} 0_{\mathbb{R} \times E^*} &\in \underline{df}(x^*) + \{[0, \psi]\} \quad \forall [0, \psi] \in \bar{df}(x^*), \\ (0_{\mathbb{R} \times E^*} &\in \bar{df}(x^*) + \{[0, \varphi]\} \quad \forall [0, \varphi] \in \underline{df}(x^*)). \end{aligned}$$

Замечание 1. В случае, когда функция f гиподифференцируема (гипердифференцируема), необходимое условие минимума (максимума) упрощается и принимает вид

$$0_{\mathbb{R} \times E^*} \in \underline{df}(x^*) \quad (0_{\mathbb{R} \times E^*} \in \bar{df}(x^*)).$$

6. Заключение. Кодифференцируемые функции образуют важное и удобное для практического применения подмножество всех негладких функций, поскольку множество кодифференцируемых функций является линейным пространством, замкнутым относительно основных алгебраических операций, а также операций взятия поточечного максимума и минимума. Для кодифференцируемых функций имеется конструктивное необходимое условие экстремума, с помощью которого можно построить численный метод нахождения стационарных точек.

Разработанная теория может быть с успехом применена к задачам вариационного исчисления и оптимального управления для нахождения новых условий экстремума, а также для построения численных методов решения данных задач.

Литература

1. Демьянов В.Ф., Васильев Л.В. Недифференцируемая оптимизация. М.: Наука, 1981. 384 с.
2. Демьянов В.Ф., Рубинов А.М. Основы негладкого анализа и квазидифференциальное исчисление. М.: Наука, 1990. 432 с.
3. Кларк Ф. Оптимизация и негладкий анализ. М.: Наука, 1988. 280 с.
4. Демьянов В.Ф. О кодифференцируемых функциях // Вестник ЛГУ, 1988. № 2(8). С. 22–26.
5. Demyanov V.F. Continuous generalized gradients for nonsmooth functions // Lecture Notes in Economics and Mathematical Systems. Berlin: Springer-Verlag, 1988. Vol. 304. P. 24–27.
6. Долгополик М.В. Кодифференциальное исчисление в нормированных пространствах // Проблемы математического анализа, 2011. Вып. 54. С. 3–22.

Жабко А.П., Зараник У.П.
Санкт-Петербургский государственный университет

Область асимптотической устойчивости управляемых систем с запаздыванием

Введение. Решение задач управления и стабилизации для систем с запаздыванием аналитическими методами не всегда удается получить. В статье Красовского Н.Н. [1] было показано, что оптимальное стабилизирующее управление является линейным непрерывным функционалом в фазовом пространстве, и выведены соотношения, описывающие параметры оптимального управления и оптимального значения функционала качества для систем с запаздыванием. К настоящему времени теоретические аспекты данной задачи для систем с запаздыванием разработаны с достаточной полнотой. Однако, в силу бесконечномерной природы систем с запаздыванием, практическое применение теории наталкивается на ряд принципиальных трудностей. Проблема разработки эффективных численных методов решения задач и их программной реализации современными вычислительными средствами является особенно актуальной.

Постановка задачи. Рассмотрим дифференциально-разностную систему управления

$$\dot{x} = f(x(t), x(t - H), u(x)), \quad (1)$$

где $f(x(t), x(t - H), u(x))$ – аналитическая функция относительно своих аргументов. Решение системы обозначим $x(t)$ или $x(t, t_0, \varphi)$ с начальной функцией φ , u – допустимое управление. Будем предполагать, что при допустимом управлении $u = u(x)$ система имеет нулевое решение.

Функционал качества имеет вид

$$J(\varphi) = \int_0^\infty W(x(t, \varphi), u(x(t))) dt, \quad (2)$$

где $W(x, u)$ аналитическая функция, $W = \sum_{m=2}^\infty W^m$, где W^m – однородные формы степени m с вещественными, однозначными, ограниченными и непрерывными коэффициентами. Оптимальным

управлением u_0 будем называть такое допустимое управление, при котором функционал (2) принимает минимальное значение.

В этой статье будем строить оптимальную последовательность управления для системы (1) и функционала качества (2), основанную на приближении данной системы разностной системой уравнений. Задача оптимального управления для систем обыкновенных дифференциальных уравнений была рассмотрена Зубовым В.И. в статье [2]. Вопрос существования оптимального управления задачи (1), (2) остается открытым.

Основной результат. Множество начальных функций $\varphi(t + \sigma)$, где $\sigma \in [-H, 0]$, обладающих свойством $x(t, \varphi) \rightarrow 0$ при $t \rightarrow +\infty$, где $x(t, \varphi)$ решение системы (1) при управлении $u = u(x)$, называется областью притяжения нулевого решения $x = 0$ при управлении u [2]. Приближение области асимптотической устойчивости

$$\mathcal{A}_\varepsilon = \left\{ \varphi \in C : \|\varphi\| \leq \frac{1}{\varepsilon}, \|\dot{\varphi}\| < K(\varepsilon) \right\}.$$

Теорема. Если существует $u_0 = u_0(x)$ – оптимальное управление, то верно следующее включение $\mathcal{A}(u_0) \supset \mathcal{A}(u)$.

Доказательство. Множество \mathcal{A}_ε компактно, тогда существует точка x – граничная точка области притяжения при допустимом управлении $u = \mathcal{A}_\varepsilon(u)$. Функционал

$$V(\varphi) = \int_0^\infty W(x(t, \varphi), x(t - H, \varphi), u(x(t))) dt,$$

обладает свойством $V(\varphi) \rightarrow +\infty$ при $\varphi \rightarrow \tilde{\varphi}$, где $\tilde{\varphi} \in \mathcal{A}_\varepsilon(u)$. Покажем справедливость этого свойства. Пусть $\tilde{\varphi} \in \Gamma(\mathcal{A}_\varepsilon(u))$ и $\{\varphi_m\}$ – последовательность точек таких, что

$$\varphi_m \rightarrow \tilde{\varphi} \text{ при } m \rightarrow +\infty,$$

где $\{\varphi_m\}_{m=1}^\infty \in \mathcal{A}_\varepsilon(u)$, $m = 1, 2, \dots$

Обозначим через s_m момент первого пересечения интегральной кривой системы (1) $x(t, \varphi_m)$ с поверхностью сферы $\|x(t, \varphi_m)\| \leq \delta$. Последовательность $\{s_m\}$ такая, что $s_m \rightarrow +\infty$ при $m \rightarrow +\infty$, иначе существует подпоследовательность $s_{m_n} : \|s_{m_n}\| \leq T$. Тогда кривые, начинающиеся в достаточно малой окрестности $\tilde{\varphi}$, остаются

в сколь угодно малой окрестности кривой $x(s, \tilde{\varphi})$ при $\|s\| \leq T$. Следовательно, кривая заходит в δ -окрестность точки $x = 0$ и неограниченно приближается к $x = 0$, а это противоречит тому, что $\tilde{\varphi} \in \Gamma(\mathcal{A}_\varepsilon(u))$. Значит, $s_m \rightarrow +\infty$ при $m \rightarrow +\infty$. Так как W положительно определенная форма своих аргументов, то $V(\varphi_m) \geq \gamma s_m$. Следовательно, $V(\varphi_m) \rightarrow +\infty$ при $m \rightarrow +\infty$.

Предположим, что $\tilde{\varphi} \in \Gamma(\mathcal{A}_\varepsilon(u_0))$ и содержится в некоторой области притяжения $\mathcal{A}_\varepsilon(u)$. Тогда существует $\rho > 0$ такое, что при $\|\tilde{\varphi} - \varphi_0\| < \rho$ будет $V(\varphi_0) < M < +\infty$. Так как $\tilde{\varphi}$ является граничной точкой $\mathcal{A}_\varepsilon(u_0)$, то найдется φ_0 такое, что $V_0(\varphi_0) > M$. Однако, из условия оптимальности управления u_0 следует, что $V_0(\varphi_0) < V(\varphi_0)$. Противоречие доказывает теорему.

Замечание. Существование управления гарантируется теоремой 3.1 [1].

Метод приближения решения дифференциально-разностных систем уравнений разностными был рассмотрен в [3], где также и приведено обоснование построения области асимптотической устойчивости дифференциально-разностных систем.

Утверждение. *Если оптимальное управление существует, то*

$$\mathcal{A}(u_0) \subset A^h(u_0).$$

Доказательство аналогичного утверждения подробно рассмотрено в статье [4].

Литература

1. Красовский Н.Н. Об аналитическом конструировании оптимального регулятора в системе с запаздыванием времени // ПММ, 1962. С. 39–51.
2. Зубов В.И. Лекции по теории управления. М.: Наука, 1975. 496 с.
3. Жабко А.П., Зараник У.П. О приближении решений экспоненциально устойчивых систем дифференциально-разностных уравнений // Вестник СПбГУ. Сер. 10, 2011. Вып. 3.
4. Жабко А.П., Зараник У.П. Об одном способе приближения области асимптотической устойчивости дифференциально-разностных систем // Вестник Мордовского университета. Сер. Физико-математические науки, 2011. Вып. 1. С. 43–48.

Казанцева Т.Е.

Тюменский государственный университет

О поведении одной динамической системы с двумя параметрами

Рекомендовано к публикации доцентом Бариновым В.А.

Рассматривается система

$$\begin{aligned}\dot{x} &= -y + x \left(1 - \frac{f(x, y)}{a}\right), \\ \dot{y} &= x + y \left(1 - \frac{f(x, y)}{a}\right) - \frac{\mu}{\sqrt{1 + \frac{y^2}{x^2}}},\end{aligned}\quad (1)$$

где $f(x, y) = 1 + 1/(1 + (\sqrt{x^2 + y^2} - 1)^2)$ – непрерывная функция, $x = x(t), y = y(t)$ – неизвестные функции, μ и a – действительные параметры, причем $a > 0, \mu \in [0, 1)$. Если $\mu = 0$, то система принимает вид

$$\dot{x} = -y + x \left(1 - \frac{f(x, y)}{a}\right), \quad \dot{y} = x + y \left(1 - \frac{f(x, y)}{a}\right). \quad (2)$$

Систему (1) можно трактовать как систему (2) с воздействием, определяемым слагаемым $-\mu/(\sqrt{1 + y^2/x^2})$. Система (2) упоминалась в [1]. Анализ ее поведения представляется несложным. Система имеет единственное состояние равновесия $(0, 0)$. Линеаризованная в окрестности $(0, 0)$ система (2) имеет характеристические корни вида $\lambda_1 = 1 - f(0, 0)/a + i, \lambda_2 = 1 - f(0, 0)/a - i$. Решение устойчиво, если $f(0, 0) > a$ и неустойчиво, если $f(0, 0) < a$. При $f(0, 0) = a$ корни становятся чисто мнимыми, значит судить о поведении системы по линеаризации в окрестности начала координат нельзя. Однако, очевидно, что при переходе характеристических корней через мнимую ось качественное поведение системы меняется. Выясним характер этого изменения.

Используя указанный выше вид функции $f(x, y)$ и переходя к полярным координатам, приведем систему (2) к виду

$$\dot{r} = 1 - \frac{1}{a}(1 + (r - 1)^2), \quad \dot{\theta} = 1. \quad (3)$$

Кратко рассмотрим, что происходит с системой при возрастании значения параметра a .

1. При $0 < a < 1$ характеристические корни имеют вид $\lambda_{1,2} = 1 - 3/2a \pm i$, их действительная часть отрицательна, т.е. начало координат является устойчивым фокусом. Система не имеет особых траекторий помимо неподвижной точки.
2. При $a = 1$ точка $(0, 0)$ – устойчивый фокус. Тем не менее, значение $a = 1$ является бифуркационным, так как при переходе через него происходит рождение предельного цикла из бесконечности. Действительно, разрешая уравнение $1 - \frac{1}{a}(1 + (r-1)^2) = 0$ относительно r , получаем

$$r_1 = \frac{a - 1 + \sqrt{-2 + 3a - a^2}}{a - 1}, \quad r_2 = \frac{a - 1 - \sqrt{-2 + 3a - a^2}}{a - 1}.$$

Оба корня действительны при $a \in [1, 2]$, причем первый неотрицателен на этом отрезке и монотонно убывает от $+\infty$ до 1. Второй корень неотрицателен при $a \in [3/2, 2]$ и монотонно возрастает от 0 до 1. Таким образом, при $a = a_0 = 1$ в системе (3) возникает замкнутая изолированная траектория (окружность), радиуса r_1 . Этот предельный цикл неустойчив, так как $r' < 0$ при $0 < r < r_1$ и $r' > 0$ при $r > r_1$. Также можно доказать, что цикл является простым.

3. При $1 < a < 3/2$ фокус в начале координат по-прежнему устойчив. На этом промежутке изменения параметра происходит уменьшение радиуса неустойчивого предельного цикла, появившегося из бесконечности, до значения 2.
4. При $a = 3/2$ характеристические корни чисто мнимые. Фокус становится сложным, а система – негрубой. Согласно теореме 2.4 [2, §3.2] при переходе через $a = a_1 = 3/2$ происходит бифуркация Андронова – Хопфа (в мягком режиме). Причем рождающийся при этом предельный цикл устойчив и стягивается к неподвижной точке при $a \rightarrow 3/2 + 0$.
5. При $3/2 < a < 2$ действительная часть характеристических корней положительна, поэтому фокус в начале координат вновь становится простым, но уже неустойчивым. В данном диапазоне изменения параметра система (3) имеет два периодических решения, определяющихся двумя простыми предель-

ными циклами: неустойчивым внешним и устойчивым внутренним. С увеличением a внешний цикл уменьшает радиус, а внутренний увеличивает, причем $r_1, r_2 \rightarrow 1$ при $a \rightarrow 2$.

6. При $a = 2$ фокус в начале координат – простой неустойчивый. Два предельных цикла сливаются в один. Непосредственно по уравнению видно, что он будет сложным и полуустойчивым. Существование сложного предельного цикла означает, что система при данном значении a негрубая. Таким образом, $a = a_2 = 2$ – еще одно бифуркационное значение.
7. При $a > 2$ система имеет единственное состояние равновесия – простой неустойчивый фокус в начале координат. Предельных циклов нет.

Таким образом, невозмущенная система (2) испытывает три бифуркации. Все возникающие в ней предельные циклы легко найти аналитически, причем все сделанные выводы о поведении системы подтверждаются численными расчетами.

Рассмотрим систему (1) при $\mu \in (0, 1)$, т.е. подвергнем систему внешним возмущениям. Система по-прежнему имеет одну неподвижную точку. Но с введением параметра μ задача сразу усложняется. Во-первых, каждой паре значений (μ, a) соответствуют свои координаты неподвижной точки, т.е. при изменении хотя бы одного из параметров неподвижная точка начинает перемещаться по фазовой плоскости. Во-вторых, аналитические выражения для характеристических корней получаются крайне громоздкими. В-третьих, получить простое уравнение, задающее особые траектории, как в случае с невозмущенной системой, не удастся.

Рассмотрим, как перемещается особая точка при изменении параметров. При переходе к полярным координатам для всех неподвижных точек выполняется

$$\mu = \frac{r}{\cos^2\theta}, \quad a = \frac{1 + 1/(1 + (r - 1)^2)}{1 - \sin\theta/\cos\theta}.$$

Понятно, что $0 < r < \mu$. Тогда для любого фиксированного μ параметр a и полярный угол θ , отвечающий неподвижной точке, связаны соотношением

$$a = \frac{1 + 1/(1 + (\mu\cos^2\theta - 1)^2)}{1 - \sin\theta/\cos\theta}. \quad (4)$$

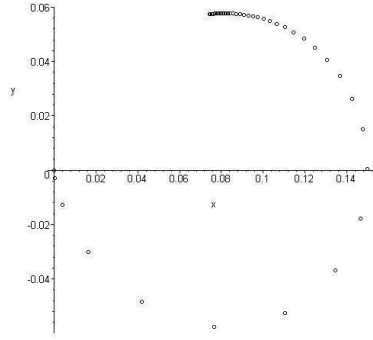


Рис. 1. Изменение положения неподвижной точки при $\mu = 0,2$
 ность достаточно большого радиуса

$$x^2 + y^2 = R^2.$$

Условие касания траектории $x = x(t)$, $y = y(t)$ системы (1) с этой окружностью имеет вид

$$\frac{1}{2} \frac{dr^2}{dt} = \dot{x}x + \dot{y}y. \quad (5)$$

Подставляя в (5) выражения для производных из (1) и переходя к полярным координатам, получим

$$r \left(\left(1 - \frac{1}{a} \left(1 + \frac{1}{1 + (r-1)^2} \right) \right) r - \mu \sin\theta \sqrt{\cos^2\theta} \right). \quad (6)$$

Так как мы предполагаем, что радиус R достаточно велик, нетрудно показать, что при $0 < a < 1$ левая часть (6) отрицательна, а при $a > 1$ положительна. Это означает, что в первом случае все траектории, пересекающие окружности (4), с возрастанием t входят внутрь их, т.е. бесконечность неустойчива. Когда значение a становится больше единицы, бесконечность становится устойчивой. Неподвижная точка при $1 < a < 3/2$ и при любом μ является простым устойчивым фокусом. А это означает, что прямая $a = \varphi(\mu) = 1$ является бифуркационной кривой, при переходе через которую из бесконечности рождается неустойчивый предельный цикл большого

На рис. 1 показано, как при $\mu = 0,2$ изменяется местоположение неподвижной точки на плоскости с ростом параметра a . При ближайшем к нулю значении a полярный угол близок к $-\pi/2$. При $\theta = 0$ параметр $a = 1 + 1/(1 + (\mu - 1)^2)$, а значение $r = \mu = 0,2$ максимально.

Рассмотрим на фазовой плоскости окруж-

радиуса. Это первая бифуркационная кривая на плоскости параметров.

Для отыскания второй бифуркационной кривой использовались численные расчеты. Для каждого фиксированного μ_k , $k = 1, \dots, 20$ исследовалось поведение системы при изменении a . На каждом шаге определялись координаты особой точки (x_k, y_k) , а также характеристические корни. Как и в предыдущем исследовании, они имеют вид $\alpha(\mu, a) \pm i\beta(\mu, a)$, что соответствует особой точке типа фокус. В силу очевидной непрерывности, при каждом заданном μ_k существует такое значение a , при переходе через которое характеристические корни переходят через мнимую ось. Бифуркационные значения соответствуют паре чисто мнимых характеристических корней $\pm i\beta(\mu, a)$, причем β очень близко к единице. Логично предположить, что, как и в случае с системой (2), при таких a происходит бифуркация Андронова – Хопфа. Действительно, для системы (1) выполняются условия теоремы Хопфа [3], следовательно, такая бифуркация имеет место. Однако, в отличие от случая невозмущенной системы, эта теорема не гарантирует единственность возникающего предельного цикла.

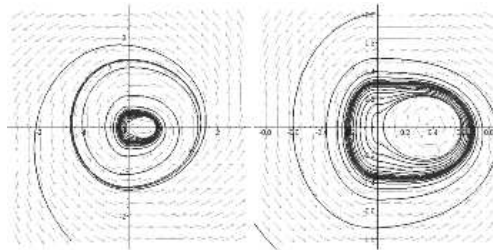


Рис. 2. Рождение предельных циклов

Методом интерполяции получена зависимость $a = \varphi_2(\mu)$ в виде многочлена. График этой зависимости при $\mu \in [0, 1)$ представляет собой вторую бифуркационную кривую. Численные эксперименты подтверждают, что при переходе через нее рождаются как минимум три предельных цикла (рис. 2). Также численными методами удалось определить, что существует некая кривая $a = \varphi_3(\mu)$ такая, что при переходе через нее происходит попарное слияние устойчивых и неустойчивых предельных циклов в полустойчивые, следовательно, наблюдается еще одна бифуркация. Кривую $a = \varphi_3(\mu)$ удалось пока лишь наметить.

Несмотря на то, что некоторые вопросы остались нерешенны-

ми, на основе проведенного исследования удалось разбить плоскость управляющих параметров на качественно различные области (рис. 3).

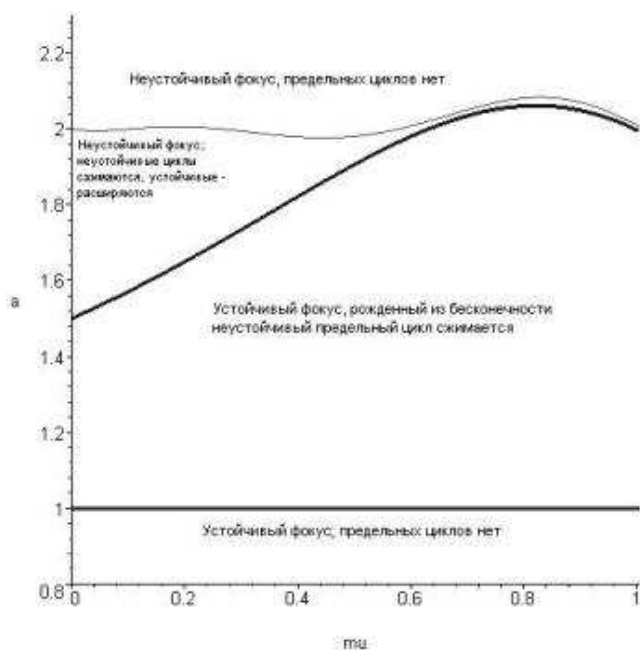


Рис. 3. Бифуркационная диаграмма

Литература

1. Drazin P.G. Nonlinear Systems. Cambridge University Press, 1992.
2. Нормальные формы и бифуркации векторных полей на плоскости / Пер. с англ. Под ред. Ю.С. Ильяшенко. М.: МЦНМО, 2005. 416 с.
3. Марсден Дж. Бифуркация рождения цикла и ее приложения. М.: Мир, 1980. 368 с.

Кузнецова А.С.

Санкт-Петербургский государственный университет

Синхронизация в осцилляторах, соединенных через обратную связь

Рекомендовано к публикации профессором Камачкиным А.М.

Введение. Встречающиеся в технике колебательные системы часто бывают связанными. Они могут быть связаны по-разному: через обратную связь, через управление и другие динамические характеристики [1, 2]. Происходящие при связанных колебаниях процессы чрезвычайно разнообразны. Система связанных осцилляторов представляет собой динамическую систему. Возникает вопрос, какие из известных нам типов движений могут возникнуть в этих системах. В частности, вопросы существования периодических решений и возникающие при этом проблемы синхронизации. Такие задачи рассматривались давно, в частности, в [3]. В этой книге исследовались сложные связанные колебания через представление движения как суперпозиции однопериодических главных колебаний. В данной статье рассматриваются осцилляторы, связанные в кольца, которые были описаны в статье [4]. Исследуется вопрос синхронизации такой системы осцилляторов.

Каждый осциллятор представляет собой преобразователь O_1 с одним переменным входом $u_{in}(t)$ и одним переменным выходом $u_{out}(t)$. Значения функций $u_{in}(t)$ и $u_{out}(t)$ являются скалярами. Внутри осциллятор состоит из системы-интегратора с обратной связью, которая осуществляется через реле. Осциллятор с полупериодом T_a представлен на рис. 1.

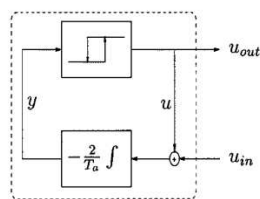


Рис. 1. Модель осциллятора

Преобразователь O_1 работает следующим образом. На вход интегратора подается сумма внешнего сигнала $u_{in}(t)$ и сигнала обратной связи $u(t)$. Эта сумма интегрируется с коэффициентом $-\frac{2}{T_a}$, в результате получается сигнал $y(t)$, который поступает на вход в реле. Реле переходит из включенного состояния (на выходе 1) в выключенное (на выходе -1), если сигнал на входе $y(t)$ дости-

гает -1 сверху, и переходит из выключенного состояния во включенное, если сигнал достигает 1 снизу [4].

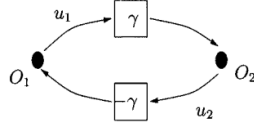


Рис. 2. Система двух осцилляторов

Рассмотрим систему двух одинаковых последовательно связанных осцилляторов, показанную на рис. 2. Осцилляторы O_1 и O_2 имеют собственные полупериоды T_1 и T_2 , и связаны между собой с коэффициентами передачи γ и $-\gamma$ соответственно. Не уменьшая общности, будем предполагать, что $T_2 \geq T_1$.

В статье [5] было установлено необходимое условие синхронизации:

Теорема 1. Система двух осцилляторов при $T_2 \geq T_1$, показанная на рис. 2, синхронизируется, причем осциллятор O_2 опережает O_1 , если $\gamma > \frac{T_2 - T_1}{T_2 + T_1}$. Данное синхронное состояние является устойчивым.

Покажем, что сформулированное условие является не только необходимым, но и достаточным.

Математическая модель системы двух осцилляторов. Запишем математическую модель системы двух осцилляторов, показанную на рис. 2 в виде системы дифференциальных уравнений

$$\begin{cases} \dot{x}(t) = -\frac{2}{T_1}(-\gamma u_2(y(t)) + u_1(x(t))), \\ \dot{y}(t) = -\frac{2}{T_2}(\gamma u_2(y(t)) + u_1(x(t))), \end{cases} \quad (1)$$

где $\gamma, T_1, T_2 \in R_+^1$ – параметры, $u_{10}, u_{20} \in \{1, 0\}$ – начальные условия для реле, $x_0, y_0 \in R^1$ – начальные условия для системы, t_0 – начальный момент времени,

$$u_1(x(t)) = \begin{cases} -1, & \text{if } (x(t) < 1 \wedge u(t-0) = -1) \vee (x(t) = -1), \\ 1, & \text{if } (x(t) > -1 \wedge u(t-0) = 1) \vee (x(t) = 1), \\ u_{10}, & \text{if } x(\tau) \in (-1, 1) \quad \forall \tau \in [t_0, t], \end{cases}$$

$$u_2(y(t)) = \begin{cases} -1, & \text{if } (y(t) < 1 \wedge u(t-0) = -1) \vee (y(t) = -1), \\ 1, & \text{if } (y(t) > -1 \wedge u(t-0) = 1) \vee (y(t) = 1), \\ u_{20}, & \text{if } y(\tau) \in (-1, 1) \quad \forall \tau \in [t_0, t]. \end{cases}$$

Синхронизированному состоянию соответствует такое периодическое решение системы (1), у которого время между двумя последовательными переключениями u_1 и u_2 совпадает.

Рассмотрим правую часть (1). Периодическим решением, которому соответствует синхронизированное состояние, будет являться функция

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{cases} (-1)^{k+1} \begin{pmatrix} x_0 \\ 1 \end{pmatrix} + (t - (k-1)(t_1 + t_2)) \begin{pmatrix} \frac{2+2\gamma}{2\gamma-2} \\ \frac{T_1}{T_2} \end{pmatrix} \\ \text{при } t \in [(k-1)(t_1 + t_2), k(t_1 + t_2) - t_2], \\ (-1)^{k+1} \begin{pmatrix} 1 \\ y_1 \end{pmatrix} + (t - k(t_1 + t_2) + t_2) \begin{pmatrix} \frac{-2+2\gamma}{-2\gamma-2} \\ \frac{T_1}{T_2} \end{pmatrix} \\ \text{при } t \in [k(t_1 + t_2) - t_2, k(t_1 + t_2)], \end{cases}$$

где $k \in N$, $x_0 = \frac{T_2 - T_1 - \gamma^2(T_1 + T_2)}{2\gamma T_1}$, $y_1 = \frac{(T_1 + T_2)\gamma^2 + T_2 - T_1}{2\gamma T_2}$, $t_1 = \frac{T_1 + T_2}{4} + \frac{T_1 - T_2}{4\gamma}$, $t_2 = \frac{T_1 + T_2}{4} - \frac{T_1 - T_2}{4\gamma}$. Данная функция будет периодической и непрерывной при $\gamma > \frac{T_2 - T_1}{T_2 + T_1}$. Следовательно, нашли решение, которое будет являться синхронизированным состоянием.

Теорема 2. Система двух осцилляторов, показанная на рисунке 2, синхронизируется, причем осциллятор O_2 опережает O_1 , тогда и только тогда, когда $\gamma > \frac{T_2 - T_1}{T_2 + T_1}$. Данное синхронное состояние является устойчивым.

Замечание. В фазовом пространстве x, y график этой функции представляет собой параллелограмм с вершинами в точках $(x_0, 1)^T, (1, y_1)^T, (-1, -y_1)^T, (-x_0, -1)^T$.

Система трех осцилляторов. Осцилляторы подобного типа можно соединять в кольца, состоящие не только из двух, но и из n штук. Теперь рассмотрим систему трех осцилляторов. Запишем систему уравнений для данной модели:

$$\begin{cases} \dot{x}(t) = -\frac{2}{T_1}(u_1(x(t)) + \gamma u_3(z(t))), \\ \dot{y}(t) = -\frac{2}{T_2}(u_2(y(t)) + \gamma u_1(x(t))), \\ \dot{z}(t) = -\frac{2}{T_3}(u_3(z(t)) + \gamma u_2(y(t))), \end{cases} \quad (2)$$

где $\gamma, T_1, T_2, T_3 \in R_+^1$ – параметры, $u_{10}, u_{20}, u_{30} \in \{-1, 1\}$ – начальные условия для реле, $x_0, y_0, z_0 \in R^1$ – начальные условия для системы,

$$u_1(x(t)) = \begin{cases} -1, & \text{если } (x(t) < -1 \wedge u(t-0) = -1) \vee (x(t) = -1), \\ 1, & \text{если } (x(t) > -1 \wedge u(t-0) = 1) \vee (x(t) = 1), \\ u_{10}, & \text{если } x(\tau) \in (-1, 1) \quad \forall \tau \in [t_0, t], \end{cases}$$

$$u_2(y(t)) = \begin{cases} -1, & \text{если } (y(t) < 1 \wedge u(t-0) = -1) \vee (y(t) = -1), \\ 1, & \text{если } (y(t) > -1 \wedge u(t-0) = 1) \vee (y(t) = 1), \\ u_{20}, & \text{если } y(\tau) \in (-1, 1) \quad \forall \tau \in [t_0, t], \end{cases}$$

$$u_3(z(t)) = \begin{cases} -1, & \text{если } (z(t) < 1 \wedge u(t-0) = -1) \vee (z(t) = -1), \\ 1, & \text{если } (z(t) > -1 \wedge u(t-0) = 1) \vee (z(t) = 1), \\ u_{30}, & \text{если } z(\tau) \in (-1, 1) \quad \forall \tau \in [t_0, t]. \end{cases}$$

Решением, при котором система будет находиться в синхронизированном состоянии, будет являться функция

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{cases} (-1)^{k+1} \begin{pmatrix} x_0 \\ 1 \\ z_0 \end{pmatrix} + (t - (k-1)(t_1 + t_2)) \begin{pmatrix} \frac{2-2\gamma}{T_1} \\ \frac{2\gamma-2}{T_2} \\ \frac{-2\gamma-2}{T_3} \end{pmatrix} \\ \text{при } t \in [(k-1)(t_1 + t_2 + t_3), (k-1)(t_1 + t_2 + t_3) + t_1], \\ \\ (-1)^{k+1} \begin{pmatrix} x_1 \\ y_1 \\ -1 \end{pmatrix} + (t - k(t_1 + t_2) + t_2) \begin{pmatrix} \frac{2+2\gamma}{T_1} \\ \frac{2\gamma-2}{T_2} \\ \frac{2-2\gamma}{T_3} \end{pmatrix} \\ \text{при } t \in [(k-1)(t_1 + t_2 + t_3) + t_1, \\ (k-1)(t_1 + t_2 + t_3) + t_1 + t_2], \\ \\ (-1)^{k+1} \begin{pmatrix} 1 \\ y_2 \\ z_2 \end{pmatrix} + (t - k(t_1 + t_2) + t_2) \begin{pmatrix} \frac{2\gamma-2}{T_1} \\ \frac{-2\gamma-2}{T_2} \\ \frac{-2\gamma+2}{T_3} \end{pmatrix} \\ \text{при } t \in [(k-1)(t_1 + t_2 + t_3) + t_1 + t_2, k(t_1 + t_2 + t_3)], \end{cases}$$

где $k \in N$,

$$t_3 = \frac{1}{2\gamma^2 - 6\gamma} ((1 - \gamma)(T_1 + T_3) - 2T_2),$$

$$t_2 = \frac{1}{2\gamma^2 - 6\gamma} ((1 - \gamma)(T_2 + T_3) - 2T_1),$$

$$t_1 = \frac{1}{2\gamma^2 - 6\gamma} ((1 - \gamma)(T_1 + T_2) - 2T_3),$$

$$x_0 = \frac{T_3(\gamma - 1)^2 + T_1(1 + \gamma) - 2T_2(1 - \gamma)}{(\gamma^2 - 3\gamma)T_1},$$

$$z_0 = \frac{-T_3(\gamma^2 - \gamma + 2) + (T_1 + T_2)(1 - \gamma^2)}{(\gamma^2 - 3\gamma)T_3},$$

$$\begin{aligned}
x_1 &= \frac{T_1(\gamma^2 - \gamma + 2) + (T_2 + T_3)(\gamma^2 - 1)}{(\gamma^2 - 3\gamma)T_1}, \\
y_1 &= \frac{2T_1(1 - \gamma) - T_2(\gamma + 1) - T_3(\gamma - 1)^2}{(\gamma^2 - 3\gamma)T_2}, \\
y_2 &= \frac{-T_2(\gamma^2 - \gamma + 2) + (T_1 + T_3)(1 - \gamma^2)}{(\gamma^2 - 3\gamma)T_2}, \\
z_2 &= \frac{T_2(\gamma - 1)^2 + T_3(1 + \gamma) - 2T_1(1 - \gamma)}{(\gamma^2 - 3\gamma)T_3}.
\end{aligned}$$

Необходимо, чтобы выполнялось условие $t_1, t_2, t_3 > 0$. Следовательно, $\gamma < 3$. Таким образом, можно сформулировать теорему.

Теорема 3. Система трех осцилляторов, определяемая уравнениями (2), синхронизируется, тогда и только тогда, когда $1 < \gamma < 3$.

Заключение. С помощью аналогичного подхода синхронизированное состояние может быть найдено для любого конечного числа n связанных осцилляторов типа O_1 . Рассмотрение фазового пространства этой системы позволило более полно изучить возможные движения осцилляторов типа O_1 . Рассмотрение этих осцилляторов при других значениях параметра γ показывает, что системы осцилляторов имеют и другие типы движений, а не только периодические. В частности, такие системы осцилляторов, обладают решениями, устойчивыми по Лагранжу.

Литература

1. Фейгин М.И. Вынужденные колебания систем с разрывными нелинейностями. М.: Наука, 1994. 285 с.
2. Мигулин В.В. Основы теории колебаний. М.: Наука, 1978. 391 с.
3. Магнус К. Колебания. М.: Мир, 1982. 303 с.
4. Varigonda S., Georgiou T.T. Dynamics of relay relaxation oscillators // Transactions on automatic control, 1964. Vol. 9, № 1. P. 157–159.
5. Macki J.W., Nistri P., Zecca P. Mathematical models for hysteresis // SIAM Review, 1993. Vol. 35, № 1. P. 94–123.

Любич С.Я.

Санкт-Петербургский государственный университет

Математическая модель оптимального планирования расписания движения поездов¹

Рекомендовано к публикации профессором Соколовым Б.В.

1. Введение. К настоящему времени разработано большое число алгоритмов, предназначенных для решения задач теории расписаний [1]. Однако в том случае, когда размерность исходной задачи велика, то при применении традиционных методов математического программирования возникают значительные вычислительные трудности. В статье предлагается динамическая и статическая модели, алгоритм решения задачи планирования, которые позволяют для ряда ситуаций эффективно решать распределительные задачи рассматриваемого класса.

2. Постановка задачи. Рассмотрим задачу планирования расписания движения на примере трёх поездов, каждый из которых потенциально может доставить универсальный груз на любую из заданных трех станций. Определены топология и характеристики каждого участка железнодорожной сети, задана средняя скорость перемещения каждого поезда, ограничения, связанные с невозможностью нахождения одновременно на одном участке пути и на одной станции нескольких поездов. Кроме того, предполагается, что каждый поезд может назначаться только на одну из трёх станций. Каждая станция в каждый момент времени может принять только один поезд. Необходимо составить такое расписание движения, чтобы каждый из поездов доставил груз на одну из станций, а суммарный штраф за нарушение директивных сроков доставки грузов был минимальным.

3. Декомпозиция задачи планирования. Для решения сформулированной нестационарной задачи теории расписаний большой размерности был рассмотрен следующий вариант ее пространственно-временной декомпозиции. На первом этапе предложено, используя статическую модель планирования движения для каждого поезда, найти его кратчайшие маршруты перемещения из исходного положения до каждой из заданных станций в отдельности. Поиск кратчайших путей осуществлялся с помощью алгоритма Флойда [2].

¹Работа выполнена при финансовой поддержке РФФИ, гранты №10-07-00311, 09-07-00066, 10-08-90027-Бел_а, отделения нанотехнологий и информационных технологий РАН, проект №О-2.3/03

Далее на разработанной авторами динамической модели осуществлялась координация движения рассматриваемых поездов во времени с учетом общих пространственных, технических и технологических ограничений.

4. Динамическая модель планирования движения группы поездов. Проведем формализацию поставленной задачи теории расписаний, используя динамическую интерпретацию процесса выполнения технических операций [3]. Рассмотрим математическую модель выполнения комплекса операций (прохождение участков пути от начального положения A_i до конечного положения – станции B_j) $D^{(i)} = \{D_\chi^{(i)}, \chi = 1, \dots, S_j, i = 1, 2, 3, j = 1, 2, 3\}$. Для этого введем неотрицательную величину $x_i = \sum_{\chi=1}^{S_i} x_{i\chi}^{(j)}, \forall j$, которую будем называть состоянием данной операции. Изменение состояния операции $D_\chi^{(i)}$ (движения поезда A_i по $D_\chi^{(i)}$ участку пути) в каждый момент времени определяется дифференциальным уравнением

$$\dot{x}_i = \sum_{j=1}^3 \sum_{\chi=1}^{S_j} \omega_{i\chi}^{(j)} u_{i\chi}^{(j)}(t), \quad (1)$$

где $\omega_{i\chi}^{(j)}(t)$ – нормированные интенсивности выполнения операций, которые получаются в результате решения задачи с использованием статической модели планирования перемещения каждого поезда (см. пункт 3) и численно равны для каждого текущего участка пути отношению усредненной длины текущего участка движения каждого поезда A_i к фактической длине каждого текущего альтернативного $D_\chi^{(i)}$ участка пути к станции B_j , по которому может двигаться данный поезд. Управляющее воздействие в этом случае принимает значения $u_{i\chi}^{(j)} = 1$, если поезд A_i выполняет операцию $D_\chi^{(i)}$, $u_{i\chi}^{(j)} = 0$ – в противном случае, $t \in \sigma = (t_0, t_f]$ – интервал планирования.

Чтобы учесть требование неразрывного выполнения работы, введем в модель планирования дополнительные дифференциальные уравнения и ограничения

$$\dot{z}_{i\chi}^{(j)} = u_{i\chi}^{(j)}, \quad \dot{h}_{i\chi}^{(j)} = z_{i\chi}^{(j)}, \quad \dot{g}_{i\chi}^{(j)} = v_{i\chi}^{(j)}; \quad (2)$$

$$v_{i\chi}^{(j)}(a_{i\chi}^{(j)} - z_{i\chi}^{(j)}) = 0, \quad \gamma_{i\chi}^{(j)} \in \{0, 1\}. \quad (3)$$

В (3) $a_{i\chi}^{(j)}$ – заданное время выполнения операции $D_\chi^{(i)}$. В соответствии с содержательной постановкой задачи в каждый момент вре-

мени управления должны удовлетворять ограничениям:

$$\begin{aligned} \sum_{j=1}^3 u_{i\chi}^{(j)}(t) &\leq 1, \quad \forall i, \quad \forall \chi; & \sum_{i=1}^3 \sum_{j=1}^3 u_{i\chi}^{(j)} &\leq 1, \quad \forall \chi; \\ \sum_{i=1}^3 u_{i\chi}^{(j)}(t) &\leq 1, \quad \forall j, \quad \forall \chi; & u_{i\chi}^{(j)} &\in \{0, 1\} \quad \forall i, \quad \forall \chi, \quad \forall j. \end{aligned} \quad (4)$$

Ограничения вида

$$u_{i\chi}^{(j)} \left(a_{i(\chi-1)}^{(j)} - x_{i(\chi-1)}^{(j)} \right) = 0 \quad \forall t, \quad (5)$$

определяют очередность выполнения частично-упорядоченных операций $D_\chi^{(i)}$, описывающих возможные маршруты перемещения A_i -го поезда [3]. Введем в рассмотрение вектор управляющих воздействий $\vec{u} = \|u_{11}^{(1)} u_{11}^{(2)} \dots u_{i\chi}^{(j)} \dots u_{3S_3}^{(3)}; v_{11}^{(1)} v_{11}^{(2)} \dots v_{i\chi}^{(j)} \dots v_{3S_3}^{(3)}\|^T$, $\vec{u} \in Q$, где Q – множество допустимых управляющих воздействий, удовлетворяющих заданным ограничениям (3)–(5). Будем предполагать, что они принадлежат классу кусочно-непрерывных функций. Краевые условия имеют в рассматриваемой задаче вид

$$t = t_0 : \quad x_{i\chi}(t_0) = z_{i\chi}^{(j)}(t_0) = h_{i\chi}^{(j)}(t_0) = g_{i\chi}^{(j)}(t_0) = 0, \quad (6)$$

$$t = t_f : \quad \left(a_i^{(j)} - \sum_{\chi=1}^{S_j} z_{i\chi}^{(j)} \right) z_{i\chi}^{(j)} = 0, \quad z_{i\chi}^{(j)}(t_f), h_{i\chi}^{(j)}(t_f), g_{i\chi}^{(j)}(t_f) \in \mathbb{R}^1. \quad (7)$$

Показателями качества выполнения операций, связанных с перемещениями поездов, выберем следующие функционалы:

$$\begin{aligned} J_1 &= \sum_{i=1}^3 \sum_{j=1}^3 \left[\left(a_i^{(j)} - \sum_{\chi=1}^{S_j} z_{i\chi}^{(j)} \right)^2 \left(z_{i\chi}^{(j)} \right)^2 + \right. \\ &+ \left. \sum_{\chi=1}^3 \left[z_{i\chi}^{(j)} g_{i\chi}^{(j)} + \frac{\left(a_{i\chi}^{(j)} \right)^2}{2} - h_{i\chi}^{(j)} \right]^2 \left(z_{i\chi}^{(j)} \right)^2 \right], \end{aligned} \quad (8)$$

J_1 – вспомогательный функционал, с помощью которого оцениваются потери из-за невыполнения краевых условий и ограничений на неразрывность операций;

$$J_2 = \sum_{i=1}^3 \sum_{j=1}^3 \int_{t_0}^{t_f} \eta_{iS_j}^{(j)}(\tau) u_{iS_j}^{(j)}(\tau) d\tau, \quad (9)$$

J_2 – основной функционал, определяющий суммарный штраф за нарушение директивных сроков при выполнении операций, где $\eta_{ij}(\tau)$ – некоторые монотонно возрастающие функции времени, которые выбираются с учетом директивных сроков начала (либо конца) выполнения операций.

Таким образом, задача оптимального планирования работ сводится к поиску допустимого управления \vec{u}_σ , обеспечивающего для уравнений (1), (2) выполнение краевых условий (6), (7), удовлетворяющего заданным ограничениям (3)–(5) и доставляющего показателю качества (9) абсолютный минимум по сравнению со всеми другими управлениями, отвечающими тем же условиям и ограничениям. Анализ показывает, что поставленная задача оптимального программного управления с помощью метода локальных сечений может быть сведена к двухточечной краевой задаче [4, 5].

5. Алгоритм решения. Анализ показывает, что основная трудность решения рассматриваемой двухточечной краевой задачи состоит в том, что в процессе итерационного поиска оптимального плана возможны прерывания работ, которые нужно устранять алгоритмически. Существует несколько эвристических правил разрешения конфликтных ситуаций, связанных с ликвидацией разрывов при выполнении работ [1, 2]. В статье предлагается алгоритм устранения разрывов при выполнении операций, в основу которого положена идея комбинированного использования метода ветвей и границ и метода последовательных приближений (МПП) Крылова – Черноусько [4, 5]. Для этого наряду с основной задачей планирования (обозначим ее буквой Γ) рассматривается релаксированная задача планирования (Р), в которой сняты ограничения, связанные с требованиями неразрывного выполнения работ, а целевая функция имеет вид

$$J_p = J_1 + J_2. \quad (10)$$

Можно показать, что задача (Р) удовлетворяет требованиям, предъявляемым к релаксированным задачам планирования. В основу решения поставленной задачи оптимального планирования работ положена схема комбинированного использования метода последовательных приближений и метода ветвей и границ, подробно изложенная в работе [4]. Вычислительная схема данного алгоритма состоит в следующем.

Шаг 1. Задается диспетчерское решение (произвольное допустимое управление) $\vec{u}_g(t)$, $t \in (t_0, t_f]$.

Шаг 2. Интегрируется основная система уравнений (1), (2) с начальными условиями (6) и $\vec{u}(t) = \vec{u}_g(t)$. Получаем $\vec{x}^{(0)}(t)$ и вычисляем $J_p = J_p^{(0)}$, которое принимаем за рекорд.

Шаг 3. Интегрируется сопряженная система уравнений с соответствующими условиями трансверсальности. Получаем $\vec{\psi}^{(0)}(t)$ (см. [4]). На этом заканчивается итерация с номером $i = 0$.

Шаг 4. Ищется управление $\vec{u}^{(i+1)}(t)$, $i = 0, 1, \dots$, исходя из условия максимизации соответствующего гамильтониана

$$\begin{aligned} & H \left(\vec{\psi}^{(i)}(t), \vec{x}^{(i)}(t), \vec{u}^{(i+1)}(t), t \right) = \\ & = \max_{\vec{u}^{(i)} \in Q(x^{(i)})} H \left(\vec{\psi}^{(i)}(t), \vec{x}^{(i)}(t), \vec{u}^{(i)}(t), t \right). \end{aligned} \quad (11)$$

Одновременно с максимизацией гамильтониана интегрируются уравнения (1), (2), при этом в каждый момент времени решается задача бивалентного программирования с целевой функцией H и ограничениями (3)–(5). Если в процессе максимизации гамильтониана в некоторый момент времени $t_j \in (t_0, t_f]$ более приоритетная операция $D_\xi^{(\omega)}$ прерывает операцию $D_\chi^{(i)}$, то в этот момент проводится ветвление (разбиение) исходной задачи (Γ) на две подзадачи ($P_\chi^{(i)}$, $P_\xi^{(\omega)}$). В задаче $P_\chi^{(i)}$ вводится запрет на прерывание операции $D_\chi^{(i)}$ и далее с использованием метода последовательных приближений решается релаксированная задача планирования. Значение целевой функции (10) в этом случае обозначим $J_{p0}^{(1)}$. В задаче $P_\xi^{(\omega)}$ ранее начавшаяся операция $D_\chi^{(i)}$ снимается с выполнения, а в момент времени t_j начинает выполняться операция $D_\xi^{(\omega)}$. Далее также решается релаксированная задача планирования. Значение целевой функции (10) в этом случае обозначим $J_{p1}^{(1)}$. Если $J_{p0}^{(1)} < J_p^{(0)}$ или $J_{p1}^{(1)} < J_p^{(0)}$, то за рекорд принимается соответственно $J_p = J_{p0}^{(1)}$ или $J_p = J_{p1}^{(1)}$. В этом случае, если выполняются оба неравенства, то за рекорд принимается $J_p = \min \{ J_{p0}^{(1)}, J_{p1}^{(1)} \}$, а конфликтная ситуация разрешается следующим образом: если $J_{p0}^{(1)} < J_{p1}^{(1)}$, то при максимизации (11) вводится запрет на прерывание операции $D_\chi^{(i)}$, в противном случае, если $J_{p0}^{(1)} > J_{p1}^{(1)}$, то операция $D_\chi^{(i)}$ вводится в гамильтониан в момент времени t_j – момент начала выполнения операции $D_\xi^{(\omega)}$.

Далее после разрешения указанной выше конфликтной ситуации продолжается планирование, исходя из (11) с учетом требований неразрывности выполнения операций до новой конфликтной ситуации, которая разрешается по аналогичной схеме. Итерационный

процесс поиска оптимального плана заканчивается при следующих условиях: либо когда в процессе решения релаксированной задачи получаем допустимое решение задачи (Г), либо когда после окончания интегрирования на шаге 4 алгоритма выполнится условие

$$|J_p^{(i+1)} - J_p^{(i)}| < \varepsilon_1, \quad (12)$$

где ε_1 – заданная величина, $i = 0, 1, \dots$. Если условие (12) не выполняется, то происходит повторение шага 3 и т.д. Вычислительная эффективность (трудоемкость) предложенного алгоритма на одной итерации приближенно может быть оценена по формуле: $N \leq C(m_1)k_1k_2 \left(\frac{\sigma}{\Delta t}\right)$, где N – среднее число операций типа сложение на одной итерации; $C(m_1) = \frac{11m_1^3 + 12m_1^2 + 3m_1}{6}$ – максимальное число операций типа сложение при решении в каждый текущий момент времени на интервале планирования задачи о назначениях венгерским методом [2] (см. (12)); m_1 – максимальная размерность указанной задачи о назначениях, которая определяется числом независимых путей в заданном сетевом графике комплекса планируемых операций; σ – длина интервала планирования; Δt – шаг интегрирования уравнений (1), (2) и соответствующей сопряженной системы уравнений; k_1 – максимальное число прерываний работ на одной итерации $k_1 = \frac{\sigma}{\Delta t}$; k_2 – максимальное число ветвлений при каждом прерывании ($k_2 = ns$).

Литература

1. Танаев В.С., Шкурба В.В. Введение в теорию расписаний. М.: Наука, 1975. 256 с.
2. Таха Х.А. Введение в исследование операций, 7-е издание / Пер. с англ. М.: Издат. Дом «Вильямс», 2005. 912 с.
3. Зимин И.Н., Иванилов Ю. П. Решение задач сетевого планирования сведением их к задачам оптимального управления // ЖВМ и МФ, 1971. Т. 11, № 3. С. 632–641.
4. Калинин В.Н., Соколов Б. В. Оптимальное планирование процесса взаимодействия активных подвижных объектов // Дифференц. уравнения, 1985. № 5. С. 752–757.
5. Крылов И.А., Черноусько Ф.Л. Алгоритм метода последовательных приближений для задач оптимального управления // ЖВМ и МФ, 1972. № 1. С. 14–34.

Медведева И.В.

Санкт-Петербургский государственный университет

Модификация алгебраического метода исследования устойчивости дифференциально-разностных уравнений

Рекомендовано к публикации профессором Жабко А.П.

Введение. Рассмотрим задачу исследования экспоненциальной устойчивости линейных стационарных систем с запаздывающим аргументом вида

$$\dot{x}(t) = Ax(t) + Bx(t-h), \quad t \geq t_0, \quad (1)$$

где $x \in R^n$, $A, B \in R^{n \times n}$ – постоянные матрицы, $h \geq 0$. Будем считать, что $t_0 = 0$ – начальный момент времени, $\phi(\theta) \in C([-h, 0])$ – начальная функция. Введем множество

$$S_0 = \{x_t : \|x(t+\sigma)\| \leq \|x(t)\| \quad \forall \sigma \in [-h, 0]\}.$$

В работах [1, 2] был предложен алгебраический подход к анализу экспоненциальной устойчивости системы (1), который основан на следующих утверждениях.

Теорема 1. Пусть система (1) экспоненциально устойчива по Ляпунову. Тогда для любой положительно-определенной матрицы $W_{n \times n}$ существует функционал $v(x_t)$ такой, что

$$1) \left. \frac{dv(x_t)}{dt} \right|_{(1)} = -x^*(t)Wx(t);$$

2) существует $c_1 > 0$ такое, что $v(x_t) \geq c_1 \|x(t)\|^2$ при $x_t \in S_0$.

Теорема 2. Пусть существует функционал $v(x_t)$ такой, что

$$1) \left. \frac{dv(x_t)}{dt} \right|_{(1)} = -x^*(t)Wx(t), \quad \text{где матрица } W \text{ положительно}$$

определена;

2) $v(x_t) \geq c_1 \|x(t)\|^2$ при $x_t \in S_0$ при некотором $c_1 > 0$.

Тогда система (1) экспоненциально устойчива по Ляпунову.

Предложенный подход заключается в следующем: квадратичный функционал Ляпунова – Красовского $v(x_t)$ оценивается на множе-

стве S_0 с помощью приближения функций этого множества некоторыми специальными функциями с учетом погрешности приближения. Таким образом, для функционала $v(x_t)$ строится оценка снизу в виде некоторой квадратичной формы, положительная определенность которой, согласно теореме 2, гарантирует экспоненциальную устойчивость системы (1). В качестве специальных функций в статье [1] были использованы линейные, а в статье [2] – кусочно-линейные функции. В настоящей работе предлагается модификация этого алгоритма с помощью замены множества кусочно-линейных функций множеством некоторых гладких функций.

Описание метода. Рассмотрим уравнение

$$\dot{x}(t) = ax(t) + bx(t-h), \quad a, b = \text{const}, \quad h > 0. \quad (2)$$

Функционал $v(x_t)$, удовлетворяющий первому условию теоремы 2 для уравнения (2), имеет вид [3]

$$\begin{aligned} v(x_t) = & u(0)x^2(t) + 2bx(t) \int_{-h}^0 u(h+\theta)x(t+\theta)d\theta + \\ & + b^2 \int_{-h}^0 \int_{-h}^0 x(t+\theta_1)u(\theta_2-\theta_1)x(t+\theta_2)d\theta_1d\theta_2, \end{aligned} \quad (3)$$

где $u(\tau)$ – матрица Ляпунова, ассоциированная с W . Примем $W = 1$. Будем оценивать функционал (3) на множестве кривых, удовлетворяющих условиям

$$|x^{(k)}(t+\sigma)| \leq (|a| + |b|)^k |x(t)|, \quad k = 0, 1, \dots, \quad \sigma \in [-h, 0]. \quad (4)$$

Разобьем отрезок $[t-h, t]$ на N частей одинаковой длины $\Delta = \frac{h}{N}$ точками

$$t-h = t_N < t_{N-1} < \dots < t_0 = t,$$

где $t_j = t - j\Delta$. Рассмотренное в работе [2] кусочно-линейное приближение функции $x(t+\theta)$ в функционале (3) на этом отрезке имеет вид

$$l_j(t, \theta) = x(t_j) + \left(x(t_j) - x(t_{j+1}) \right) \frac{\theta}{\Delta}, \quad \theta \in [-\Delta, 0], \quad j = \overline{0, N-1}.$$

В результате оценки погрешности такого приближения была получена оценка снизу для функционала (3) в виде

$$v(x_t) \geq v_1(\hat{x}),$$

где $v_1(\hat{x})$ – квадратичная форма относительно вектора \hat{x} , а вектор \hat{x} размерности $N + 1$ определяется соотношениями $\hat{x}_j = x(t_j)$, $j = \overline{0, N}$. Результаты исследования экспоненциальной устойчивости уравнения (2) алгебраическим методом, полученные при кусочно-линейном приближении, приведены в работе [2].

Однако область экспоненциальной устойчивости в пространстве параметров, построенная таким методом, оказалась достаточно далека от известной точной области экспоненциальной устойчивости. Причиной этого во многом является негладкость кусочно-линейных функций, приближающих гладкие функции множества S_0 . Поэтому рассмотрим теперь приближение $x(t + \theta)$ сплайном кубических функций.

Будем искать заданную на отрезке $[t_{j+1}, t_j]$ кубическую функцию $q_j(t, \theta)$, удовлетворяющую условиям

$$\begin{aligned} q_j(t, 0) &= x(t_j), & q_j(t, -\Delta) &= x(t_{j+1}), \\ q'_j(t, 0) &= \dot{x}(t_j), & q'_j(t, -\Delta) &= \dot{x}(t_{j+1}). \end{aligned} \quad (5)$$

При таком задании приближения в виде кубического сплайна условие гладкости сплайна будет выполняться автоматически. Искомое приближение $x(t + \theta)$, удовлетворяющее условиям (5), имеет вид

$$\begin{aligned} q_j(t, \theta) &= g_1(\theta)x(t_j) + g_2(\theta)x(t_{j+1}) + g_3(\theta)\dot{x}(t_j) + g_4(\theta)\dot{x}(t_{j+1}), & (6) \\ \theta &\in [-\Delta, 0], \quad j = \overline{0, N-1}, \end{aligned}$$

где

$$\begin{aligned} g_1(\theta) &= -\frac{2}{\Delta^3}\theta^3 - \frac{3}{\Delta^2}\theta^2 + 1, & g_2(\theta) &= \frac{2}{\Delta^3}\theta^3 + \frac{3}{\Delta^2}\theta^2, \\ g_3(\theta) &= \frac{1}{\Delta^2}\theta^3 + \frac{2}{\Delta}\theta^2 + \theta, & g_4(\theta) &= \frac{1}{\Delta^2}\theta^3 + \frac{1}{\Delta}\theta^2, \quad \theta \in [-\Delta, 0]. \end{aligned}$$

Выражая теперь с помощью формулы Тейлора $x(t_j + \theta)$ как функцию четырех параметров $x(t_j)$, $x(t_{j+1})$, $\dot{x}(t_j)$, $\dot{x}(t_{j+1})$, определяющих значение $q_j(t, \theta)$ на отрезке $[t_{j+1}, t_j]$, и используя (4) при $k = 4$, получим следующую оценку для погрешности приближения (6):

$$\begin{aligned} |\eta_j(t, \theta)| &= |x(t_j + \theta) - q_j(t, \theta)| \leq d_3|x(t)|(\theta^4 - 6\Delta\theta^3 + 7\Delta^2\theta^2), & (7) \\ d_3 &= \frac{(|a| + |b|)^4}{24}, \quad \theta \in [-\Delta, 0], \quad j = \overline{0, N-1}. \end{aligned}$$

Таким образом, подставляя в функционал (3) вместо $x(t + \theta)$ его приближение сплайном кубических функций (6) и учитывая оценку погрешности такого приближения (7), так же, как и в кусочно-линейном случае, получим оценку снизу для функционала (3) в виде квадратичной формы

$$v(x_t) \geq v_3(\hat{x}),$$

в которой вектор \hat{x} определяется соотношениями

$$\hat{x}_j = x(t_j), \quad \hat{x}_{N+j+1} = \dot{x}(t_j), \quad j = \overline{0, N}.$$

Другими словами, вектор \hat{x} состоит теперь не только из самих значений $x(t_0), \dots, x(t_N)$, но также из значений их производных, и имеет размерность $2N + 2$.

В работах [1, 2] далее проводилось исследование положительной определенности квадратичной формы $v_1(\hat{x})$ с помощью критерия Сильвестра, исходя из чего и делался вывод об экспоненциальной устойчивости уравнения (2). При приближении сплайном кубических функций такой подход оказывается несостоятельным. Невозможность применения здесь критерия Сильвестра объясняется тем, что компоненты вектора \hat{x} , относительно которого получена квадратичная форма $v_3(\hat{x})$, зависят друг от друга, их связывают соотношения

$$|x(t + \sigma)| \leq |x(t)|, \quad |\dot{x}(t + \sigma)| \leq (|a| + |b|)|x(t)|, \quad \sigma \in [-h, 0]. \quad (8)$$

Отметим, что в кусочно-линейном случае компоненты вектора \hat{x} также не были независимы, однако исследование положительной определенности квадратичной формы $v_1(\hat{x})$ с помощью критерия Сильвестра все же давало положительные результаты. Добиться положительных результатов и в случае приближения кубическим сплайном, а также улучшить результаты, полученные при кусочно-линейном приближении, позволит следующий подход.

Задача минимизации. Будем находить минимум квадратичной формы $v_1(\hat{x})$ или $v_3(\hat{x})$ на множестве кривых, удовлетворяющих ограничениям (8) при $x(t) \neq 0$. В силу однородности квадратичной формы можно положить $\hat{x}_0 = 1$. Для исследования экспоненциальной устойчивости уравнения (2) необходимо решить задачу квадратичного программирования

$$\min_{\hat{x} \in \hat{S}_i} v_i(\hat{x}) \quad (9)$$

при $i = 1$ для кусочно-линейного приближения и при $i = 3$ для кусочно-кубического приближения, где

$$\widehat{S}_1 = \{\hat{x}_{(N+1) \times 1} : \hat{x}_0 = 1, |\hat{x}_j| \leq 1, j = \overline{1, N}\},$$

$$\widehat{S}_3 = \{\hat{x}_{(2N+2) \times 1} : \hat{x}_0 = 1, |\hat{x}_j| \leq 1, j = \overline{1, N}, |\hat{x}_j| \leq (|a| + |b|), j = \overline{N+1, 2N+1}\}.$$

Согласно теореме 2, положительность минимума (9) является достаточным условием экспоненциальной устойчивости уравнения (2).

Решение задачи минимизации. Сравнение результатов. Выводы. В таблицах 1–3 для некоторых конкретных примеров заведомо устойчивых уравнений приведены такие значения запаздывания h_0 в зависимости от N , для которых при $0 \leq h \leq h_0$ можно гарантировать экспоненциальную устойчивость рассматриваемых уравнений. В таблице 1 такие значения h_0 получены при помощи критерия Сильвестра [1, 2], в таблицах 2 и 3 – решением задачи квадратичного программирования (9) при $i = 1$ и $i = 3$ соответственно встроенными средствами системы MATLAB.

Таблица 1. Случай кусочно-линейного приближения, критерий Сильвестра

	a, b	$N = 1$	$N = 2$	$N = 3$	$N = 5$	$N = 10$	Точн. гр. уст-ти
1	$a = -1$ $b = -1$	0,423	0,478	0,588	0,717	0,942	$+\infty$
2	$a = -0,5$ $b = -1$	0,482	0,514	0,623	0,742	0,938	$\approx 2,417$
3	$a = 1$ $b = -2$	0,148	0,161	0,192	0,224	0,267	$\approx 0,604$
4	$a = 0$ $b = -1$	0,580	0,571	0,672	0,778	0,941	$\approx 1,5708$

Таблица 2. Случай кусочно-линейного приближения, задача минимизации

	a, b	$N = 1$	$N = 2$	$N = 3$	$N = 5$	$N = 10$	Точн. гр. уст-ти
1	$a = -1$ $b = -1$	0,499	0,806	1,013	1,308	1,834	$+\infty$
2	$a = -0,5$ $b = -1$	0,563	0,901	1,108	1,371	1,73	$\approx 2,417$
3	$a = 1$ $b = -2$	0,248	0,341	0,419	0,539	0,604	$\approx 0,604$
4	$a = 0$ $b = -1$	0,658	1,047	1,244	1,449	1,5707	$\approx 1,5708$

Таблица 3. Случай приближения сплайном кубических функций

	a, b	$N = 1$	$N = 3$	$N = 5$	$N = 10$	$N = 20$	Точн. гр. уст-ти
1	$a = -1$ $b = -1$	0,577	1,320	1,849	2,876	4,48	$+\infty$
2	$a = -0,5$ $b = -1$	0,695	1,506	1,924	2,315	2,415	$\approx 2,417$
3	$a = 1$ $b = -2$	0,301	0,604	0,604	0,604	0,604	$\approx 0,604$
4	$a = 0$ $b = -1$	0,900	1,5707	1,5707	1,5707	1,5707	$\approx 1,5708$

Анализируя таблицы 1–3, можно отметить значительное улучшение результатов, полученных в работах [1, 2], за счет перехода к решению задачи минимизации, и особенно за счет перехода к приближению $x(t+\theta)$ в функционале (3) гладкой функцией. При приближении сплайном кубических функций в примерах наблюдается стремление значений h_0 к точной границе устойчивости, что позволяет выдвинуть гипотезу о сходимости алгебраического метода. Жизнеспособность такой гипотезы подтверждает и следующее утверждение, которое вытекает непосредственно из вида коэффициентов $v_3(\dot{x})$.

Утверждение. Пусть уравнение (2) экспоненциально устойчиво. Тогда $|v(x_t) - v_3(\dot{x})| \xrightarrow{N \rightarrow +\infty} 0$ на кривых x_t , удовлетворяющих условиям (4).

Дальнейшие исследования будут направлены на строгое доказательство сходимости алгебраического метода.

Литература

1. Медведева И.В. Обращение прямого метода Ляпунова при анализе устойчивости систем с запаздыванием // Процессы управления и устойчивость: Труды 41-й международной научной конференции аспирантов и студентов / Под ред. Н.В. Смирнова, Г.Ш. Тамасяна. СПб.: Издат. Дом С.-Петербур. гос. ун-та, 2010. С. 33–38.
2. Жабко А.П., Медведева И.В. Алгебраический подход к анализу устойчивости дифференциально-разностных систем // Вестник СПбГУ. Сер. 10, 2011. Вып. 1. С. 9–20.
3. Харитонов В.Л. Функционалы Ляпунова с заданной производной // Вестник СПбГУ. Сер. 10, 2005. Вып. 1. С. 110–117, 199–207.

Морозова Д.Н.

Санкт-Петербургский государственный университет

Синтез линейных управлений с линейным запаздыванием

Рекомендовано к публикации доцентом Чижовой О.Н.

Задача 1. Пусть задана система дифференциальных уравнений

$$\dot{X} = PX + QU. \quad (1)$$

Будем считать, что X – вектор размерности n , U – вектор размерности $m < n$, P, Q – заданные, постоянные, вещественные матрицы размерности $n \times n, n \times m$ соответственно. Требуется построить управление

$$U = U(t, X(\alpha t)) \quad (2)$$

и начальную функцию

$$X = \phi(t) \quad (3)$$

при $t \in [0, \tau]$ так, чтобы система (1) имела при начальной функции (3) и управлении (2) решение $X = X(t, X_0)$, удовлетворяющее условиям

$$X(0) = X_0, \quad X(T) = 0. \quad (4)$$

Здесь $0 < \tau < T$, $0 < \alpha < 1$, X_0 известны. Если подставить управление (2) в систему (1), то получим систему дифференциальных уравнений с запаздывающим аргументом. Для того чтобы построить решение такой системы, необходимо задать начальную функцию. Тогда решение системы (1) на следующем промежутке $t \in [\tau, \frac{\tau}{\alpha}]$ будет определяться как решение системы уравнений

$$\dot{X} = PX(t) + QU(t, \phi(\alpha t)) \quad (5)$$

с начальным условием $X = \phi$ при $t = \tau$. Действуя аналогичным образом, можно построить решение системы (1) при управлении (2)

на любом промежутке $[\frac{\tau}{\alpha^k}, \frac{\tau}{\alpha^{k+1}}]$, $k = 1, 2, \dots$. Следовательно, можно построить решение системы (1) при управлении (2) на всём промежутке $[0, T]$. Обозначим $Y(t)$ фундаментальную матрицу системы линейных уравнений $\dot{X} = PX$ с начальным условием $Y(0) = E$.

Теорема 1. *Если строки матрицы $B(t) = Y^{-1}(t)Q$ линейно независимы на любом промежутке $[t, T]$, $t \in [0, \alpha T]$, то поставленная задача разрешима.*

Доказательство. Если существует управление, переводящее систему (1) из любого начального состояния X_0 в положение $X = 0$ за время $[0, T]$, то по лемме о представлении допустимых управлений [1] его можно представить в следующем виде

$$U(t) = B^*(t)C + V(t), \quad (6)$$

где C – это постоянный вектор, а векторная функция $V(t)$ удовлетворяет условию ортогональности

$$\int_0^T B(t)V(t)dt = 0. \quad (7)$$

Подставим управление (6) в систему (1) и запишем решение полученной системы в форме Коши

$$X(t, 0, X_0) = Y(t) \left(X_0 + \int_0^t B(s)B^*(s)dsC + \int_0^t B(s)V(s)ds \right). \quad (8)$$

В момент T выполняется второе граничное условие $X(T) = 0$. Поэтому

$$0 = Y(t) \left(X_0 + \int_0^T B(s)B^*(s)dsC \right).$$

Введём матрицу $A(t, T) = \int_t^T B(s)B^*(s)ds$. По критерию линейной независимости при выполнении условий теоремы 1 матрица A будет неособой.

$$\begin{aligned} 0 &= Y(T)\{X_0 + A(0, T)C\}, \quad Y(T) \neq 0, \\ X_0 + A(0, T)C &= 0. \end{aligned} \quad (9)$$

Для того чтобы получить управление вида (6), проинтегрируем замкнутую систему ещё раз на $[t, T]$

$$-Y^{-1}(t)X(t) = A(t, T)C + \int_t^T B(s)V(s)ds.$$

Затем, домножая слева на $-Y(t)$, получим

$$X(t) = -Y(t)A(t, T)C - Y(t) \int_t^T B(s)V(s)ds. \quad (10)$$

Из формулы (9) найдём $C = -A^{-1}(0, T)X_0$. Подставляя это выражение в (10), имеем

$$X(t) = Y(t)A(t, T)A^{-1}(0, T)X_0 - Y(t) \int_t^T B(s)V(s)ds. \quad (11)$$

Формула (11) даёт решение системы (1) при управлении (6), причём это решение должно удовлетворять граничным условиям (4).

$$X(0) = Y(0)A(0, T)A^{-1}(0, T)X_0 - Y(0) \int_0^T B(s)V(s)ds = X_0,$$

$$X(T) = Y(T)A(T, T)A^{-1}(0, T)X_0 - Y(T) \int_T^T B(s)V(s)ds = 0.$$

Оба граничных условия выполнены. Тогда формула (11), рассматриваемая на $[0, \tau]$, может быть взята в качестве начальной функции

$$\phi(t) = Y(t)A(t, T)A^{-1}(0, T)X_0 - Y(t) \int_t^T B(s)V(s)ds.$$

Рассмотрим выражение (10) в момент αt :

$$X(\alpha t) = -Y(\alpha t)A(\alpha t, T)C - Y(\alpha t) \int_{\alpha t}^T B(s)V(s)ds.$$

В условиях теоремы матрица $A(\alpha t, T)$ будет не вырождена для любых рассматриваемых значений t .

$$C = -A^{-1}(\alpha t, T)Y^{-1}(\alpha t) \left(X(\alpha t) + Y(\alpha t) \int_{\alpha t}^T B(s)V(s)ds \right).$$

Затем подставляем найденное C в управление (6) и получаем

$$U(t, X(\alpha t)) = -B^*(t)A^{-1}(\alpha t, T)Y^{-1}(\alpha t)X(\alpha t) - \\ -B^*(t)A^{-1}(\alpha t, T) \int_{\alpha t}^T B(s)V(s)ds + V(t) = M(t, \alpha, T)X(\alpha t) + N(t, \alpha, T),$$

$$M(t, \alpha, T) = -B^*(t)A^{-1}(\alpha t, T)Y^{-1}(\alpha t),$$

$$N(t, \alpha, T) = -B^*(t)A^{-1}(\alpha t, T) \int_{\alpha t}^T B(s)V(s)ds + V(t).$$

Теорема доказана.

Задача 2. Для системы

$$\dot{X} = P_1X(t) + P_2X(\alpha t) + QU \quad (12)$$

построить управление вида (2) и начальную функцию (3) так, чтобы соответствующее решение удовлетворяло граничным условиям (4). Будем считать, что X – вектор размерности n , U – вектор размерности $m < n$, P_1, P_2, Q – заданные, постоянные, вещественные матрицы размерности $n \times n, n \times n, n \times m$ соответственно. Будем искать управление в виде

$$U(t, X(\alpha t)) = C_1X(\alpha t) + \bar{U}(t), \quad (13)$$

где матрицу C_1 будем строить из условия

$$QC_1 = -P_2. \quad (14)$$

Займёмся разрешением системы (14). Обозначим Q^1, \dots, Q^m столбцы матрицы Q и C_1^1, \dots, C_1^n – столбцы матрицы C_1 . Тогда произведения этих матриц запишутся как

$$\begin{aligned} QC_1 &= \begin{pmatrix} q_{11}c_{11} + \dots + q_{1m}c_{m1} & \dots & q_{11}c_{1n} + \dots + q_{1m}c_{mn} \\ \dots & \dots & \dots \\ q_{n1}c_{11} + \dots + q_{nm}c_{m1} & \dots & q_{n1}c_{1n} + \dots + q_{nm}c_{mn} \end{pmatrix} \\ &= (QC_1^1, \dots, QC_1^n) = \\ &= ((c_{11}Q^1 + \dots + c_{m1}Q^m), \dots, (c_{1n}Q^1 + \dots + c_{mn}Q^m)), \end{aligned}$$

$$QC_1^1 = \begin{pmatrix} q_{11} & q_{12} \dots q_{1m} \\ q_{21} & q_{22} \dots q_{2m} \\ \vdots & \vdots \dots \vdots \\ q_{n1} & q_{n2} \dots q_{nm} \end{pmatrix} \begin{pmatrix} c_{11} \\ c_{21} \\ \dots \\ c_{m1} \end{pmatrix} = \begin{pmatrix} q_{11}c_{11} + \dots + q_{1m}c_{m1} \\ q_{21}c_{11} + \dots + q_{2m}c_{m1} \\ \dots \\ q_{n1}c_{11} + \dots + q_{nm}c_{m1} \end{pmatrix}.$$

$$QC_1^1 = c_{11}Q^1 + \dots + c_{m1}Q^m = (QC_1)^1.$$

Применим векторизацию [2]

$$\begin{aligned} \text{vec}(QC_1) &= \begin{pmatrix} c_{11}Q^1 + \dots + c_{m1}Q^m \\ c_{12}Q^1 + \dots + c_{m2}Q^m \\ \dots \\ c_{1n}Q^1 + \dots + c_{mn}Q^m \end{pmatrix} = \begin{pmatrix} (QC_1)^1 \\ (QC_1)^2 \\ \dots \\ (QC_1)^n \end{pmatrix} = \\ &= \begin{pmatrix} QC_1^1 \\ QC_1^2 \\ \dots \\ QC_1^n \end{pmatrix} = \underbrace{\begin{pmatrix} Q0 \dots \dots 0 \\ 0Q \dots \dots 0 \\ \dots \\ 00 \dots \dots Q \end{pmatrix}}_{=\tilde{Q}} \begin{pmatrix} C_1^1 \\ C_1^2 \\ \dots \\ C_1^n \end{pmatrix} = \tilde{Q}\text{vec}(C_1). \end{aligned}$$

Будем рассматривать систему вида

$$\tilde{Q}\text{vec}(C_1) = \text{vec}(-P_2). \quad (15)$$

Известна следующая [3]

Теорема 2. Система (15) разрешима тогда и только тогда, когда $\text{rank}(\tilde{Q}) = \text{rank}(\tilde{Q}, \text{vec}(-P_2))$.

Замкнём систему (12) управлением (13)

$$\dot{X} = P_1 X(t) + Q\bar{U}(t). \quad (16)$$

Для построения управления в системе (16) можно пользоваться алгоритмом, рассмотренным в задаче 1.

Рассмотрим ещё один вид управления

$$U = H_1 X(t) + H_2 X(\alpha t) + \bar{U}. \quad (17)$$

Наложим ограничение $QH_2 = -P_2$. Замкнём систему (12) управлением (17)

$$\dot{X} = (P_1 + QH_1)X(t) + Q\bar{U}.$$

Пусть $P = P_1 + QH_1$. Это квадратная матрица. Перепишем систему с учётом введённого обозначения

$$\dot{X} = PX(t) + Q\bar{U}. \quad (18)$$

Для построения управления в системе (18) можно пользоваться алгоритмом В.И. Зубова, рассмотренным в задаче 1. Далее все рассуждения аналогичны изложенным выше.

Литература

1. Зубов В.И. Лекции по теории управления. М.: Наука, 1975. 495 с.
2. Ланкастер П. Теория матриц. М.: Наука, 1978. 280 с.
3. Гантмахер Ф.Р. Теория матриц. М.: Наука, 2004. 560 с.

Нгуен Динь Хуен
Санкт-Петербургский государственный университет

**Об асимптотической устойчивости некоторого
класса нелинейных неавтономных
разностных систем**

Рекомендовано к публикации профессором Александровом А.Ю.

1. Введение. В широком классе случаев при исследовании математических моделей реальных процессов и явлений непрерывные системы приближенно заменяют дискретными [1]. В частности, большинство численных методов решения дифференциальных уравнений основано на сведении их к уравнениям в конечных разностях [2].

Одно из направлений исследований, возникающих в указанных приложениях разностных уравнений, связано с анализом устойчивости их решений.

Основным методом анализа устойчивости нелинейных разностных систем является прямой метод Ляпунова. Главная трудность, возникающая при применении этого метода, заключается в проблеме нахождения функций Ляпунова. К сожалению, общих способов их построения не существует. Эта проблема является особенно трудной для сложных систем [3, 4].

Сложные системы имеют составную структуру и представляют собой объединение нескольких более простых подсистем, взаимосвязанных между собой. Характерной чертой сложных систем является многомерность, т.е. высокая размерность описывающих эти системы уравнений. Многомерность приводит к трудностям, как аналитическим, так и вычислительным, и вынуждает искать специальные пути, позволяющие понизить размерность на отдельных этапах исследования.

В настоящей работе рассматривается сложная система, состоящая из двух взаимодействующих подсистем. Предполагается, что изолированные подсистемы существенно нелинейны, а связи являются нелинейным и неавтономным. Достаточные условия асимптотической устойчивости нулевого решения такой системы были получены в статье [5]. Цель настоящей работы – показать, что если связи между подсистемами удовлетворяют некоторым дополнитель-

ным ограничениям, то найденные в [5] условия устойчивости можно ослабить.

2. Постановка задачи. Пусть задана сложная система

$$\begin{aligned}x_1(k+1) &= x_1(k) + hF_1(x_1(k)) + hB_1(k)Q_1(x_2(k)), \\x_2(k+1) &= x_2(k) + hF_2(x_2(k)) + hB_2(k)Q_2(x_1(k)),\end{aligned}\quad (1)$$

которая описывает взаимодействие двух изолированных подсистем

$$x_1(k+1) = x_1(k) + hF_1(x_1(k)), \quad (2)$$

$$x_2(k+1) = x_2(k) + hF_2(x_2(k)). \quad (3)$$

Здесь $k = 0, 1, \dots$; векторы $x_1(k) \in E^{n_1}$, $x_2(k) \in E^{n_2}$, $x(k) = (x_1^*(k), x_2^*(k))^*$; элементы векторов $F_1(x_1)$, $F_2(x_2)$ – непрерывно дифференцируемые при $x_1 \in E^{n_1}$, $x_2 \in E^{n_2}$ однородные функции порядка $\mu_1 > 1$ и $\mu_2 > 1$ соответственно; элементы векторов $Q_1(x_2)$, $Q_2(x_1)$ – непрерывно дифференцируемые при $x_1 \in E^{n_1}$, $x_2 \in E^{n_2}$ однородные функции порядка $\alpha_1 > 1$ и $\alpha_2 > 1$ соответственно; матрицы $B_1(k)$ и $B_2(k)$ ограничены при $k = 0, 1, \dots$; h – шаг дискретизации ($h > 0$).

Будем считать, что нулевые решения изолированных систем дифференциальных уравнений

$$\dot{x}_1(t) = F_1(x_1(t)), \quad (4)$$

$$\dot{x}_2(t) = F_2(x_2(t)) \quad (5)$$

асимптотически устойчивы.

В работе [6] доказано, что если нулевые решения систем дифференциальных уравнений (4), (5) асимптотически устойчивы, то при любом шаге дискретизации h нулевые решения разностных систем (2), (3) также являются асимптотически устойчивыми. Определим условия устойчивости нулевого решения сложной системы (1) по нелинейному приближению (2), (3).

В статье [5] с помощью метода функций Ляпунова было доказано, что при выполнении неравенства

$$\mu_1\mu_2 < \alpha_1\alpha_2 \quad (6)$$

нулевое решение системы (1) асимптотически устойчиво.

Цель настоящей работы – показать, что в случае, когда матрицы $B_1(k), B_2(k)$ удовлетворяют некоторым дополнительным условиям, ограничение (6) на значения параметров $\mu_1, \mu_2, \alpha_1, \alpha_2$ можно ослабить.

3. Условия асимптотической устойчивости. По матрицам $B_1(k)$ и $B_2(k)$ построим вспомогательные матрицы $\Phi_1(k)$ и $\Phi_2(k)$. Пусть

$$\Phi_i(0) = 0, \Phi_i(k) = \Phi_i(k-1) + B_i(k-1), k = 1, 2, \dots, i = 1, 2.$$

Будем считать, что матрицы $\Phi_1(k)$ и $\Phi_2(k)$ ограничены при $k = 0, 1, \dots$

Теорема. *Если*

$$\varepsilon_1 \varepsilon_2 < 1, \quad (7)$$

где

$$\varepsilon_1 = \max \left\{ \frac{\mu_2 + 1}{2\alpha_2}, \frac{\mu_2}{\alpha_2 - 1 + \mu_1}, \frac{\mu_2 - \alpha_1}{\alpha_2 - 1} \right\},$$

$$\varepsilon_2 = \max \left\{ \frac{\mu_1 + 1}{2\alpha_1}, \frac{\mu_1}{\alpha_1 - 1 + \mu_2}, \frac{\mu_1 - \alpha_2}{\alpha_1 - 1} \right\},$$

то при любом шаге дискретизации h нулевое решение системы (1) асимптотически устойчиво.

Доказательство. Функцию Ляпунова выбираем в виде

$$\begin{aligned} \tilde{V} = & V_1(x_1) + V_2(x_2) - \\ & - h \left(\frac{\partial V_1}{\partial x_1} \right)^* \Phi_1(k) Q_1(x_2) - h \left(\frac{\partial V_2}{\partial x_2} \right)^* \Phi_2(k) Q_2(x_1), \end{aligned} \quad (8)$$

где $V_1(x_1), V_2(x_2)$ – дважды непрерывно дифференцируемые, положительно-определенные однородные функции порядка $\gamma_1 - \mu_1 + 1$ и $\gamma_2 - \mu_2 + 1$ соответственно, причем функции $(\partial V_i / \partial x_i)^* F_i(x_i)$, $i = 1, 2$, отрицательно определены. В работе [7] доказано, что если нулевые решения систем (4) и (5) асимптотически устойчивы, то такие функции обязательно существуют, а в качестве γ_1 и γ_2 можно выбирать любые рациональные числа с четными числителями и нечетными знаменателями, $\gamma_1 > \mu_1, \gamma_2 > \mu_2$.

Функция (8) положительно определена. Вычислим ее приращение на решениях системы (1). Имеем

$$\begin{aligned}\Delta \tilde{V} &= V_1(x_1(k+1)) - V_1(x_1(k)) + V_2(x_2(k+1)) - V_2(x_2(k)) - \\ &\quad - h \left(\frac{\partial V_1}{\partial x_1}(x_1(k+1)) \right)^* \Phi_1(k+1) Q_1(x_2(k+1)) - \\ &\quad - h \left(\frac{\partial V_2}{\partial x_2}(x_2(k+1)) \right)^* \Phi_2(k+1) Q_2(x_1(k+1)) + \\ &+ h \left(\frac{\partial V_1}{\partial x_1}(x_1(k)) \right)^* \Phi_1(k) Q_1(x_2(k)) + h \left(\frac{\partial V_2}{\partial x_2}(x_2(k)) \right)^* \Phi_2(k) Q_2(x_1(k)).\end{aligned}$$

При всех $k = 0, 1, \dots$ и $x \in E^n$ справедлива оценка

$$\begin{aligned}\Delta \tilde{V} &\leq -A_1 \|x_1\|^{\gamma_1} - A_2 \|x_2\|^{\gamma_2} + A_3 \|x_1\|^{\gamma_1 - \mu_1 - 1} \|x_2\|^{2\alpha_1} + \\ &+ A_4 \|x_1\|^{2\alpha_2} \|x_2\|^{\gamma_2 - \mu_2 - 1} + A_5 \|x_2\|^{\alpha_1(\gamma_1 - \mu_1 + 1)} + A_6 \|x_1\|^{\alpha_2(\gamma_2 - \mu_2 + 1)} + \\ &+ A_7 \|x_1\|^{\gamma_1 - \mu_1} \|x_2\|^{\mu_2 + \alpha_1 - 1} + A_8 \|x_1\|^{\mu_1 + \alpha_2 - 1} \|x_2\|^{\gamma_2 - \mu_2} + \\ &+ A_9 \|x_1\|^{\gamma_1 - \mu_1 + \alpha_2} \|x_2\|^{\alpha_1 - 1} + A_{10} \|x_1\|^{\alpha_2 - 1} \|x_2\|^{\gamma_2 - \mu_2 + \alpha_1} + \\ &+ A_{11} \|x_1\|^{\gamma_1 - \mu_1 + \alpha_2(\alpha_1 - 1)} \|x_2\|^{\mu_2} + A_{12} \|x_1\|^{\mu_1} \|x_2\|^{\gamma_2 - \mu_2 + \alpha_1(\alpha_2 - 1)} + \\ &+ A_{13} \|x_1\|^{\gamma_1 - \mu_1 + \alpha_1 \alpha_2} + A_{14} \|x_2\|^{\gamma_2 - \mu_2 + \alpha_1 \alpha_2} + \\ &+ A_{15} \|x_2\|^{\alpha_1(\gamma_1 - \mu_1 + 1) + \mu_2 - 1} + A_{16} \|x_1\|^{\alpha_2(\gamma_2 - \mu_2 + 1) + \mu_1 - 1} + \\ &+ A_{17} \|x_1\|^{\alpha_2} \|x_2\|^{\alpha_1(\gamma_1 - \mu_1 + 1) - 1} + A_{18} \|x_1\|^{\alpha_2(\gamma_2 - \mu_2 + 1) - 1} \|x_2\|^{\alpha_1} + \\ &+ A_{19} \|x_1\|^{\alpha_1 \alpha_2 - \alpha_2} \|x_2\|^{\alpha_1(\gamma_1 - \mu_1) + \mu_2} + A_{20} \|x_1\|^{\alpha_2(\gamma_2 - \mu_2) + \mu_1} \|x_2\|^{\alpha_1 \alpha_2 - \alpha_1} + \\ &+ A_{21} \|x_1\|^{\alpha_1 \alpha_2} \|x_2\|^{\alpha_1(\gamma_1 - \mu_1)} + A_{22} \|x_1\|^{\alpha_2(\gamma_2 - \mu_2)} \|x_2\|^{\alpha_1 \alpha_2} + \\ &+ A_{23} \|x_1\|^{\gamma_1 - 1} \|x_2\|^{\alpha_1} + A_{24} \|x_1\|^{\alpha_2} \|x_2\|^{\gamma_2 - 1} + \\ &+ A_{25} \|x_1\|^{\mu_1} \|x_2\|^{\alpha_1(\gamma_1 - \mu_1)} + A_{26} \|x_1\|^{\alpha_2(\gamma_2 - \mu_2)} \|x_2\|^{\mu_2},\end{aligned}$$

где A_1, A_2, \dots, A_{26} – положительные постоянные.

Используя свойства однородных функций [7], получаем, что если

$$\begin{aligned}\frac{\gamma_1 - \mu_1 - 1}{\gamma_1} + \frac{2\alpha_1}{\gamma_2} &> 1, & \frac{2\alpha_2}{\gamma_1} + \frac{\gamma_2 - \mu_2 - 1}{\gamma_2} &> 1, \\ \frac{\gamma_1 - \mu_2}{\gamma_2} + \frac{\mu_2 + \alpha_1 - 1}{\gamma_2} &> 1, & \frac{\mu_1 + \alpha_2 - 1}{\gamma_1} + \frac{\gamma_2 - \mu_2}{\gamma_2} &> 1, \\ \frac{\gamma_1 - \mu_2 + \alpha_2}{\gamma_2} + \frac{\alpha_1 - 1}{\gamma_2} &> 1, & \frac{\alpha_2 - 1}{\gamma_1} + \frac{\gamma_2 - \mu_2 + \alpha_1}{\gamma_2} &> 1,\end{aligned}\quad (9)$$

то при достаточно малых значениях $\|x(k)\|$ и при всех $k = 0, 1, \dots$ имеем

$$\Delta \tilde{V} \leq -\frac{A_1}{2} \|x_1\|^{\gamma_1} - \frac{A_2}{2} \|x_2\|^{\gamma_2}.$$

Нетрудно проверить, что если справедливо неравенство (7), то параметры γ_1 и γ_2 можно выбрать так, чтобы выполнялись условия (9).

Теорема доказана.

Заключение. В настоящей работе для исследуемого класса сложных систем предложен способ построения неавтономной функции Ляпунова. Показано, что применение этой функции позволяет получить более широкую область значений параметров системы, для которых можно гарантировать асимптотическую устойчивость нулевого решения, по сравнению с областью, задаваемой известными условиями.

Литература

1. Халанай А., Векслер Д. Качественная теория импульсных систем. М.: Мир, 1971. 312 с.
2. Деккер К., Вервер Я. Устойчивость методов Рунге – Кутты для жестких нелинейных дифференциальных уравнений. М.: Мир, 1988. 334 с.
3. Метод векторных функций Ляпунова в теории устойчивости. / Под ред. А.А. Воронова и В.М. Матросова М.: Наука, 1987. 312 с.
4. Руш Н., Абетс П., Лалуа М. Прямой метод Ляпунова в теории устойчивости. М.: Мир, 1980. 300 с.
5. Александров А.Ю., Жабко А.П. О сохранении устойчивости при дискретизации систем обыкновенных дифференциальных уравнений // Сибирский математический журнал, 2010. Т. 51, № 3. С. 481–497.
6. Александров А.Ю., Жабко А.П. Устойчивость разностных систем. СПб.: НИИ Химии СПбГУ, 2003. 112 с.
7. Зубов В.И. Устойчивость движения. М.: Высшая школа, 1973. 272 с.

Пономарев А.А.

Санкт-Петербургский государственный университет

О выборе параметров метода «предиктор-корректор»

Рекомендовано к публикации профессором Жабко А.П.

1. Введение. Метод «предиктор-корректор» (*Model predictive control*, MPC) в различных модификациях успешно применяется для управления подвижными объектами и технологическими процессами (примеры можно найти в [1]). Одним из достоинств метода является возможность его настройки с помощью изменения ряда параметров. Тем не менее, вопрос зависимости качества стабилизации от подобных модификаций исследован недостаточно.

Следует отметить, что существует ряд результатов, касающихся достаточных условий асимптотической устойчивости метода MPC. Некоторые из них приведены в монографии [2] и обзорной статье [3]. Эти способы стабилизации накладывают достаточные, но довольно строгие ограничения на выбор параметров метода, что может приводить даже к неразрешимости оптимизационной задачи. Имеется потребность в том, чтобы при более общих предположениях описать множество параметров метода, гарантирующих устойчивость замкнутой системы, а также в том, чтобы стремиться к наилучшему выбору параметров, исходя из качества стабилизации и оптимизации.

В данной работе аналитически исследуется движение линейной стационарной системы под управлением регулятора MPC. Из аналитического описания непосредственно вытекают необходимые и достаточные условия экспоненциальной устойчивости замкнутой системы, а также ряд условий, обеспечивающих возможность выбора подходящих значений параметров. Эти последние условия затем уточнены оценками, полученными с помощью второго метода Ляпунова.

2. Метод «предиктор-корректор» и задача стабилизации. Рассмотрим управляемую линейную стационарную систему

$$\dot{x}(t) = Ax(t) + Bu(t), \quad (1)$$

где $x \in \mathbb{R}^n$, $u \in \mathbb{R}^m$, $t \geq 0$, с наблюдением полного вектора состояния $x(t)$. Начальное условие для этой системы задаётся при $t = 0$.

Конкретизируем схему метода «предиктор-корректор», о которой будет идти речь в дальнейшем. Применение МРС в некоторый момент времени t_0 начинается с решения оптимизационной задачи

$$\mathcal{I}(x(\cdot, t_0, x^{(0)}, u(\cdot)), u(\cdot), t_0) \xrightarrow{u(\cdot) \in \mathcal{U}_{t_0}} \inf.$$

Здесь $x(\cdot, t_0, x^{(0)}, u(\cdot))$ – решение задачи Коши для системы (1) с начальным условием $x(t_0) = x^{(0)}$ и управлением $u(\cdot)$, выбранным из множества \mathcal{U}_{t_0} допустимых на отрезке $[t_0, t_0 + T]$ управлений, T – параметр метода, называемый горизонтом прогноза, $T > 0$. Функционал качества $\mathcal{I}(x(\cdot), u(\cdot), t_0)$ есть квадратичный интегральный функционал вида

$$\mathcal{I}(x(\cdot), u(\cdot), t_0) = \int_{t_0}^{t_0+T} (x^\top(t)Mx(t) + u^\top(t)Nu(t)) dt,$$

где $M \geq 0$ и $N > 0$ – симметрические матрицы. Выберем некоторое оптимальное в указанном смысле управление, если таковое существует, и применим его на отрезке $[t_0, t_0 + h]$, где h – второй параметр метода, называемый шагом, $0 < h \leq T$. По достижении времени $t_0 + h$ из системы наблюдения извлекается значение $x(t_0 + h)$, и процесс повторяется с теми же параметрами T и h , но новым начальным условием.

Мы ограничимся теми ситуациями, когда множество \mathcal{U}_{t_0} не зависит от t_0 . Система (1) и функционал качества \mathcal{I} стационарны, поэтому достаточно исследовать первый шаг метода МРС, который соответствует $t_0 = 0$. При исследовании этого первого шага множество допустимых управлений обозначим \mathcal{U} .

Под стабилизацией понимается достижение экспоненциальной устойчивости нулевого положения равновесия замкнутой системы. Именно, пусть $x^{(0)}$ – начальное значение вектора $x(t)$, т.е. $x(0) = x^{(0)}$. Будем говорить, что при этом начальном условии функция $\bar{u}(t, x^{(0)})$, заданная при всех $t \geq 0$, является стабилизирующим управлением, если при $u(t) \equiv \bar{u}(t, x^{(0)})$ система (1) имеет решение $\bar{x}(t)$ с начальным условием $\bar{x}(0) = x^{(0)}$, и это решение допускает оценку

$$\|\bar{x}(t)\| \leq \alpha \|x^{(0)}\| e^{\gamma t},$$

где $\alpha \geq 1$ и $\gamma > 0$ – некоторые постоянные.

В нашем случае задача состоит в том, чтобы исследовать, при каких значениях шага h и горизонта прогноза T и при каких начальных условиях $x^{(0)}$ регулятор MPC гарантирует стабилизирующее управление. Рассмотрим наиболее простые из уже известных результатов, касающихся этой проблемы. Оба они относятся только к случаю, когда матрица M положительно определена.

Утверждение 1 [2]. Пусть множество \mathcal{U} управлений, допустимых на отрезке $[0, T]$, есть множество таких управлений, которые переводят систему (1) из положения $x(0) = x^{(0)}$ в положение $x(T) = 0$. Возможно, что это пространство сужено дополнительными ограничениями. Если задача оптимизации функционала \mathcal{I} разрешима на каждом шаге метода MPC, то система (1), замкнутая MPC с таким пространством допустимых управлений, асимптотически устойчива при любых h и T .

Утверждение 2 [2]. Не конкретизируя пространство допустимых управлений \mathcal{U} , предположим, что горизонт прогноза T бесконечен, т.е. в момент времени t_0 минимизируется функционал

$$\mathcal{I}(x(\cdot), u(\cdot), t_0) = \int_{t_0}^{\infty} (x^T(t)Mx(t) + u^T(t)Nu(t)) dt.$$

Тогда если поставленная оптимизационная задача разрешима на каждом шаге метода MPC, то система (1), замкнутая данным MPC, асимптотически устойчива при любом h .

Оба утверждения достаточно просто доказываются с помощью прямого метода Ляпунова. Однако очевидно, что при делаемых предположениях условие разрешимости оптимизационной задачи, которое является необходимым условием применимости MPC, выполняется далеко не обязательно. Тем не менее, утверждение 2 позволяет предположить, что при достаточно большом, но конечном значении горизонта T метод MPC будет стабилизирующим. Это предположение не лишено основания, как видно из дальнейших результатов.

3. Условия устойчивости. В качестве класса допустимых управлений \mathcal{U} рассмотрим класс кусочно непрерывных функций, заданных на отрезке $[0, T]$. Предположим, что оптимальное управление в задаче минимизации функционала существует.

Теорема 1 [4]. Управление $u_{opt}(\cdot)$ оптимально тогда и только тогда, когда вместе с соответствующим движением $x_{opt}(\cdot)$ оно

удовлетворяет условию

$$u_{opt}(t) = -N^{-1}B^T \int_t^T e^{A^T(\tau-t)} M x_{opt}(\tau) d\tau \quad (2)$$

при всех значениях t из отрезка $[0, T]$. Это оптимальное управление единственно.

Следствие. Если оптимальное управление $u_{opt}(\cdot)$ существует при любом выборе начального вектора $x^{(0)}$, то оно линейно зависит от компонент этого вектора:

$$u_{opt}(t) = S(t)x^{(0)},$$

причём элементы матрицы $S(t)$ бесконечное число раз непрерывно дифференцируемы на отрезке $[0, T]$.

Вопрос существования оптимального управления при всяком начальном векторе равносильен проблеме существования матрицы $S(t)$, которая это оптимальное управление порождает. Подставим управление $u_{opt}(t) = S(t)x^{(0)}$ в систему (1) с условием $x(0) = x^{(0)}$. Получим оптимальную траекторию

$$x_{opt}(t) = R(t)x^{(0)}, \quad R(t) = e^{At} + \int_0^t e^{A(t-\tau)} BS(\tau) d\tau.$$

Из условия (2) получается выражение матрицы $S(t)$ через $R(t)$, которое вместе с указанным только что выражением $R(t)$ через $S(t)$ образует систему уравнений относительно $2n^2$ неизвестных компонент матриц $S(t)$ и $R(t)$. Эта система сводится к системе линейных обыкновенных дифференциальных уравнений, решая которую, получим матрицы $S(t)$ и $R(t)$ в виде

$$\begin{pmatrix} S(t) \\ R(t) \end{pmatrix} = \begin{pmatrix} \mathbb{O} & N^{-1}B^T \\ \mathbb{E} & \mathbb{O} \end{pmatrix} e^{\Phi t} \begin{pmatrix} \mathbb{E} \\ K_1 \end{pmatrix}.$$

Здесь

$$\Phi = \begin{pmatrix} A & BN^{-1}B^T \\ M & -A^T \end{pmatrix},$$

а K_1 — $(n \times n)$ -матрица, элементы которой удовлетворяют системе

$$\begin{pmatrix} \mathbb{O} & \mathbb{E} \end{pmatrix} e^{\Phi T} \begin{pmatrix} \mathbb{E} \\ K_1 \end{pmatrix} = \mathbb{O}. \quad (3)$$

Очевидно следующее утверждение.

Теорема 2. Пусть существует решение K_1 системы (3). Система (1), замкнутая МРС, экспоненциально устойчива тогда и только тогда, когда все собственные числа матрицы

$$R(h) = \begin{pmatrix} \mathbb{E} & \mathbb{O} \end{pmatrix} e^{\Phi h} \begin{pmatrix} \mathbb{E} \\ K_1 \end{pmatrix}$$

по модулю меньше единицы.

Из разложения

$$R(h) = \mathbb{E} + R'(0)h + o(h), \text{ где } R'(0) = A + BN^{-1}B^TK_1,$$

можно получить условие того, что при фиксированном T и некотором достаточно малом h достигается экспоненциальная устойчивость. Точнее говоря, справедлива следующая теорема.

Теорема 3. Пусть система (3) имеет решение K_1 . Если матрица $A + BN^{-1}B^TK_1$ гурвицева, то существует такое число $h_* > 0$, что при всех значениях шага h из промежутка $(0, h_*)$ система (1), замкнутая МРС, экспоненциально устойчива.

Оценка числа h_* получается с помощью квадратичных функций Ляпунова. Именно, в предположениях теоремы 3 спектр матрицы $R'(0)$ лежит в открытой левой полуплоскости, поэтому для произвольной положительно определённой $(n \times n)$ -матрицы W существует положительно определённая матрица V , удовлетворяющая матричному уравнению Ляпунова

$$(R'(0))^TV + VR'(0) = -W.$$

Можно гарантировать, что спектр матрицы $R(h)$ лежит внутри единичного круга, если матрица $R^T(h)VR(h) - V$ отрицательно определена. Используя разложение $R(h)$ в ряд по h , получаем следующее утверждение.

Теорема 4. При выполнении условия теоремы 3 в качестве числа h_* достаточно взять такое число, что при всех значениях h из промежутка $(0, h_*)$ справедливо неравенство

$$-\lambda_{\min}(W)h + (1 + \|K_1\|)^2(e^{2\|\Phi\|h} - 1 - 2\|\Phi\|h)\lambda_{\max}(V) < 0, \quad (4)$$

где $\lambda_{\min}(\cdot)$ и $\lambda_{\max}(\cdot)$ – наименьшее и наибольшее собственные числа.

Что касается выбора T , то с помощью непрерывности утверждение 2, которое гарантирует стабилизацию при $T = \infty$ и *положительно определённой* матрице M , можно распространить на случай достаточно большого, но конечного горизонта прогноза. Действительно, если система (3) имеет решение K_∞ при $T = \infty$, то она имеет решение и при конечном, но достаточно большом значении T , причём соответствующее решение стремится к K_∞ , когда $T \rightarrow \infty$.

Теорема 5. *Пусть при $T = \infty$ оптимальное в смысле минимизации функционала \mathcal{I} управление существует и является стабилизирующим. Тогда для каждого данного значения шага h найдётся такой горизонт прогноза T , что система (1), замкнутая кусочно непрерывным МРС, экспоненциально устойчива.*

Очевидно, что если оптимальное управление при $T = \infty$ существует, то при условии положительной определённости матрицы M оно заведомо является стабилизирующим. Это приводит к непосредственному обобщению утверждения 2 на достаточно большие значения T . Кроме того, можно заметить, что при условиях теоремы 5 условие теоремы 3 заведомо выполняется.

4. Заключение. Доказанные теоремы 2–5 дают некоторое представление об аналитических свойствах метода МРС в линейном стационарном случае при кусочно непрерывном управлении. Аналогичные результаты могут быть получены и в кусочно постоянном случае. Их непосредственное использование затруднительно, однако при изучении нелинейных систем полученная информация о поведении МРС в применении к линейным системам небесполезна.

Литература

1. Camacho C.F., Bordons C. Model predictive control. London: Springer-Verlag, 1999. 280 p.
2. Maciejowski J.M. Predictive control with constraints. Harlow: Prentice Hall, 2002. 331 p.
3. Mayne D.Q., Rawlings J.B., Rao C.V., Scolaert P.O.M. Constrained model predictive control: Stability and optimality // Automatica, 2000. Vol. 36, № 6. P. 789–814.
4. Зубов В.И. Лекции по теории управления. М.: Наука, 1975. 496 с.

Рехлясова Л.В.

Санкт-Петербургский государственный университет

Вычисление \mathcal{H}_2 нормы передаточной матрицы системы с запаздыванием

Рекомендовано к публикации профессором Харитоновым В.Л.

1. Основные понятия. Рассмотрим систему с запаздыванием в следующем виде

$$\begin{aligned}\dot{x}(t) &= A_0x(t) + A_1x(t-h) + B_0u(t) + B_1u(t-h), \\ y(t) &= C_0x(t) + C_1x(t-h),\end{aligned}\quad (1)$$

где $A_0, A_1 \in \mathbb{R}^{n \times n}$, $B_0, B_1 \in \mathbb{R}^{n \times p}$, $C_0, C_1 \in \mathbb{R}^{m \times n}$, $u(t) \in \mathbb{R}^p$, $y(t) \in \mathbb{R}^m$, $t \geq 0$, $h > 0$.

Будем считать, что система (1) экспоненциально устойчива при нулевом управлении $u(t)$. Введем нулевую начальную функцию $x(\theta) = 0$, $\theta \in [-h, 0]$. Пусть матрица $K(t)$ является фундаментальной матрицей [1] системы (1) с начальными условиями $K(0) = \mathbb{E}$ и $K(\theta) = \mathbb{O}$ при $\theta < 0$.

Определение 1 [2]. Передаточной матрицей системы (1) называется матрица, устанавливающая связь между образом Лапласа вектора наблюдения $y(t)$ и образом Лапласа вектора управления $u(t)$ при нулевой начальной функции.

Будем считать, что все условия для использования преобразования Лапласа [3] выполнены и для функций $x(t, 0)$, $u(t)$ и $y(t)$ существуют их образы $x(t, 0) \Rightarrow X(s)$, $u(t) \Rightarrow U(s)$, $y(t) \Rightarrow Y(s)$. Применяя преобразование Лапласа к системе (1), получим

$$\begin{aligned}sX(s) &= A_0X(s) + A_1X(s)e^{-sh} + B_0U(s) + B_1U(s)e^{-sh}, \\ Y(s) &= C_0X(s) + C_1X(s)e^{-sh}.\end{aligned}$$

Из преобразованной системы находим

$$Y(s) = (C_0 + C_1e^{-sh})(s\mathbb{E} - A_0 - A_1e^{-sh})^{-1}(B_0 + B_1e^{-sh})U(s).$$

Следовательно, передаточная матрица

$$G(s) = (C_0 + C_1e^{-sh})(s\mathbb{E} - A_0 - A_1e^{-sh})^{-1}(B_0 + B_1e^{-sh}).$$

Можно установить, что преобразование Лапласа фундаментальной матрицы имеет вид $\widehat{K}(s) = (s\mathbb{E} - A_0 - A_1 e^{-sh})^{-1}$, тогда получим $G(s) = (C_0 + C_1 e^{-sh})\widehat{K}(s)(B_0 + B_1 e^{-sh})$.

2. Норма \mathcal{H}_2 передаточной матрицы. Введем определение пространства \mathcal{H}_2 , названного так в честь английского математика Г.Х. Харди.

Определение 2 [2]. Пространством \mathcal{H}_2 называется Гильбертово пространство комплекснозначных матричных функций F , аналитичных в открытой правой полуплоскости, и таких, что интеграл $\int_{-\infty}^{\infty} Tr(F^*(i\omega)F(i\omega))d\omega$ существует и конечен. Здесь Tr – след матрицы, а «*» – знак сопряженного транспонирования.

Определение 3 [2]. Квадрат \mathcal{H}_2 нормы матрицы $G(s)$ есть

$$\|G\|_2^2 = \frac{1}{2\pi} \int_{-\infty}^{\infty} Tr(G^*(i\omega)G(i\omega))d\omega.$$

Замечание. Так как система (1) является экспоненциально устойчивой, то \mathcal{H}_2 норма передаточной матрицы $G(s)$ существует и конечна.

Утверждение [4]. *Норму \mathcal{H}_2 матрицы $G(s)$ можно находить по формуле*

$$\|G\|_2^2 = \int_0^{\infty} Tr(H^T(t)H(t))dt, \quad (2)$$

где $H(t)$ является оригиналом преобразования Лапласа $G(s)$.

3. Вспомогательные матрицы. Введем матрицу $U(\tau)$, которая удовлетворяет условиям

$$U'(\tau) = U(\tau)A_0 + U(\tau - h)A_1, \quad \tau \geq 0, \quad (3)$$

$$U(-\tau) = U^T(\tau), \quad (4)$$

$$U(0)A_0 + U(-h)A_1 + A_0^T U(0) + A_1^T U(h) = -C_0^T C_0 - C_1^T C_1. \quad (5)$$

Матрица $U(\tau)$ является матрицей Ляпунова и для нее справедливо следующее

Утверждение 1 [5]. Если система (1) экспоненциально устойчива, то матрица $U(\tau) = \int_0^\infty K^T(t)[C_0^T C_0 + C_1^T C_1]K(t + \tau)dt$ является единственным решением (3), удовлетворяющим условиям (4) и (5).

Введем еще одну матрицу $V(\tau)$, удовлетворяющую следующим условиям

$$V'(\tau) = V(\tau)A_0 + V(\tau - h)A_1, \quad \tau \geq 0, \quad (6)$$

$$V(-\tau) = \tilde{V}^T(\tau), \quad (7)$$

$$V(0)A_0 + V(-h)A_1 + A_0^T V(0) + A_1^T V(h) = -C_0^T C_1, \quad (8)$$

где $\tilde{V}(\tau)$ есть решение (6), удовлетворяющее условию

$$\tilde{V}(0)A_0 + \tilde{V}(-h)A_1 + A_0^T \tilde{V}(0) + A_1^T \tilde{V}(h) = -C_1^T C_0.$$

Для матрицы $V(\tau)$ верно

Утверждение 2. Если система (1) является экспоненциально устойчивой, то матрица $V(\tau) = \int_0^\infty K^T(t)C_0^T C_1 K(t + \tau)dt$ является единственным решением (6), удовлетворяющим (7) и (8).

4. Теорема об \mathcal{H}_2 норме передаточной матрицы. Здесь представлен основной результат данной работы в виде формулы для вычисления \mathcal{H}_2 нормы передаточной матрицы $G(s)$, с помощью которой ее можно будет вычислять аналитически.

Теорема. Пусть система (1) экспоненциально устойчива. Тогда \mathcal{H}_2 норма ее передаточной матрицы есть

$$\begin{aligned} \|G\|_2^2 = & Tr\left(B_0^T U(0)B_0 + B_0^T U(-h)B_1 + B_1^T U(h)B_0 + B_1^T U(0)B_1\right) + \\ & + 2Tr\left(B_0^T V(-h)B_0 + B_0^T V(-2h)B_1 + B_1^T V(0)B_0 + B_1^T V(-h)B_1\right), \end{aligned}$$

где $U(\tau)$ удовлетворяет (3)–(5), а $V(\tau)$ удовлетворяет (6)–(8).

Доказательство. Чтобы найти \mathcal{H}_2 норму матрицы $G(s)$, пользуясь формулой (2), нужно сначала вычислить ее оригинал по Лапласу. Нетрудно убедиться, что

$$H(t) = C_0 K(t)B_0 + C_0 K(t - h)B_1 + C_1 K(t - h)B_0 + C_1 K(t - 2h)B_1.$$

Представим $H(t)^T H(t) = I_1(t) + I_2(t)$, где

$$\begin{aligned} I_1(t) = & B_0^T K^T(t) C_0^T C_0 K(t) B_0 + B_0^T K^T(t) C_0^T C_0 K(t-h) B_1 + \\ & + B_1^T K^T(t-h) C_0^T C_0 K(t) B_0 + B_1^T K^T(t-h) C_0^T C_0 K(t-h) B_1 + \\ & + B_0^T K^T(t-h) C_1^T C_1 K(t-h) B_0 + B_0^T K^T(t-h) C_1^T C_1 K(t-2h) B_1 + \\ & + B_1^T K^T(t-2h) C_1^T C_1 K(t-h) B_0 + B_1^T K^T(t-2h) C_1^T C_1 K(t-2h) B_1, \end{aligned}$$

$$\begin{aligned} I_2(t) = & + B_0^T K^T(t) C_0^T C_1 K(t-h) B_0 + B_0^T K^T(t) C_0^T C_1 K(t-2h) B_1 + \\ & + B_1^T K^T(t-h) C_0^T C_1 K(t-h) B_0 + B_1^T K^T(t-h) C_0^T C_1 K(t-2h) B_1 + \\ & + B_0^T K^T(t-h) C_1^T C_0 K(t) B_0 + B_0^T K^T(t-h) C_1^T C_0 K(t-h) B_1 + \\ & + B_1^T K^T(t-2h) C_1^T C_0 K(t) B_0 + B_1^T K^T(t-2h) C_1^T C_0 K(t-h) B_1. \end{aligned}$$

Тогда

$$\|G\|_2^2 = \int_0^\infty \text{Tr}(H^T(t)H(t)) dt = \int_0^\infty \text{Tr}(I_1(t)) dt + \int_0^\infty \text{Tr}(I_2(t)) dt.$$

Рассмотрим первый интеграл $\int_0^\infty \text{Tr}(I_1(t)) dt = \text{Tr}\left(\int_0^\infty I_1(t) dt\right)$.

Пользуясь тем, что $K(\theta) = \mathbb{O}$ при $\theta < 0$, получим

$$\begin{aligned} \int_0^\infty I_1(t) dt = & B_0^T \int_0^\infty K^T(t) W K(t) dt B_0 + B_0^T \int_0^\infty K^T(t) W K(t-h) dt B_1 + \\ & + B_1^T \int_0^\infty K^T(t) W K(t+h) dt B_0 + B_1^T \int_0^\infty K^T(t) W K(t) dt B_1, \end{aligned}$$

где $W = C_0^T C_0 + C_1^T C_1$.

По утверждению 1 матрица $U(\tau) = \int_0^\infty K^T(t) W K(t+\tau) dt$ является единственным решением (3), удовлетворяющим (4) и (5), поэтому

$$\begin{aligned} & \int_0^\infty \text{Tr}(I_1(t)) dt = \\ & = \text{Tr}\left(B_0^T U(0) B_0 + B_0^T U(-h) B_1 + B_1^T U(h) B_0 + B_1^T U(0) B_1\right). \end{aligned}$$

Теперь рассмотрим второй интеграл. В выражении для $I_2(t)$ входят пары слагаемых, которые отличаются друг от друга, с точностью до транспонирования. Тогда

$$\begin{aligned} \text{Tr}(I_2(t)) = & 2\text{Tr}\left(B_0^T K^T(t) Q K(t-h) dt B_0 + B_1^T K^T(t) Q K(t-h) B_1 + \right. \\ & \left. + B_1^T K^T(t) Q K(t) dt B_0 + B_0^T K^T(t) Q K(t-2h) B_1\right), \end{aligned}$$

где $Q = C_0^T C_1$. По утверждению 2 матрица $V(\tau)$ является единственным решением (6), удовлетворяющим (7), (8) и, следовательно,

$$\begin{aligned} \int_0^\infty \text{Tr}(I_2(t)) dt = & 2\text{Tr}\left(B_0^T V(-h) B_0 + B_1^T V(-h) B_1 + \right. \\ & \left. + B_1^T V(0) B_0 + B_0^T V(-2h) B_1\right). \end{aligned}$$

Таким образом, суммируя два выражения для $\int_0^\infty \text{Tr}(I_1(t)) dt$ и $\int_0^\infty \text{Tr}(I_2(t)) dt$, получим, что норма \mathcal{H}_2 передаточной матрицы $G(s)$ удовлетворяет условию теоремы.

Теорема доказана.

Литература

1. Харитонов В.Л. Функционалы Ляпунова с заданной производной. Функционалы полного типа // Вестник СПбГУ, 2005. Сер. 10, Вып 1. С. 110–117.
2. Лаврентьев М.А., Шабат Б.В. Методы теории функций комплексного переменного. М.: Наука, 1973. 736 с.
3. Zhou K., Doyle J., Glover K. Robust and optimal control. Prentice Hall, 1996. 586 p.
4. Jarlebring E., Vanbiervliet J., Michiels W. Explicit expressions for the \mathcal{H}_2 norm of time-delay systems based on the delay Lyapunov equation // 49th IEEE Conference on Decision and Control, 2010.
5. Харитонов В.Л. Функционалы Ляпунова с заданной производной. Матрицы Ляпунова // Вестник СПбГУ. Сер. 10, 2005. Вып 2. С. 200–209.

Сукач М.П.

Санкт-Петербургский государственный университет

Пример решения задачи двухуровневой оптимизации¹

Рекомендовано к публикации профессором Демьяновым В.Ф.

В данной статье представлен пример решения задачи двухуровневой оптимизации методом точных штрафов [1].

Постановка задачи. Пусть $f, f_1 : R^n \rightarrow R$ – заданные непрерывные функции, $G \subset R^n$ – заданное множество. Положим

$$\Omega = \{x \in G \mid f_1(x) \leq f_1(y) \forall y \in G\},$$

где $\Omega \subset G$ – множество точек минимума функции f_1 на множестве G . Предположим, что $\Omega \neq \emptyset$. Задача

$$f(x) \longrightarrow \min_{x \in \Omega}$$

называется задачей двухуровневой оптимизации.

Вспомогательные сведения. Рассмотрим задачу минимизации функции $f(x)$ на множестве

$$\Omega = \{x \in R^n \mid \varphi(x) = 0\}.$$

Зафиксируем $\lambda \geq 0$ и введём функцию

$$F_\lambda(x) = f(x) + \lambda\varphi(x).$$

Функция $F_\lambda(x)$ называется штрафной функцией (для заданных f и φ), число λ – штрафным параметром.

Определение 1. Функция $F_\lambda(x)$ называется точной штрафной функцией для функции f на множестве Ω , если существует такой штрафной параметр λ^* , что для любых $\lambda \geq \lambda^*$ множество точек минимума функции $F_\lambda(x)$ на всем пространстве R^n совпадает с множеством точек минимума функции f на множестве Ω .

¹Работа выполнена при финансовой поддержке РФФИ, грант № 09-01-00360

Положим

$$\Omega_\delta = \{x \in R^n \mid \varphi(x) \leq \delta\}, \quad \varphi^\downarrow(x) = \liminf_{y \rightarrow x} \frac{\varphi(y) - \varphi(x)}{\|x - y\|},$$

и предположим, что выполнено условие $\inf_{x \in R^n} f(x) = f^* > -\infty$.

Имеет место следующая теорема о достаточных условиях существования точной штрафной функции.

Теорема [2]. Пусть

1) существует $\lambda^* < \infty$ такое, что для каждого $\lambda > \lambda^*$ найдётся $x_\lambda \in R^n$, для которого $F_\lambda(x_\lambda) = F_\lambda^* = \inf_{x \in R^n} F_\lambda(x)$;

2) найдутся $\delta > 0$ и $a > 0$ такие, что $\varphi^\downarrow(x) \leq -a < 0 \forall x \in \Omega_\delta \setminus \Omega$;

3) функция f липшицева на $\Omega_\delta \setminus \Omega$.

Тогда существует $\lambda^* \geq \lambda$ такое, что $\varphi(x_\lambda) = 0 \forall \lambda > \lambda^*$, $f_\lambda(x) = f_\Omega^* = \inf_{x \in \Omega} f(x)$, т.е. точка x_λ – решение задачи.

Штрафные функции в двухуровневой задаче. Множество Ω представим в виде $\Omega = \{x \in G \mid \varphi(x) = 0\}$, где

$$\varphi(x) = \sup_{y \in G} (f_1(x) - f_1(y)) = f_1(x) - f_{1G}^*, \quad f_{1G}^* = \inf_{y \in G} F_1(y).$$

Рассмотрим случай, когда множество G представимо в виде

$$G = \{x \in R^n \mid \varphi_1(x) = 0\},$$

где $\varphi_1(x) \geq 0 \forall x \in R^n$.

Пусть $\lambda_1 \geq 0, \lambda_2 \geq 0$ фиксированы. Получаем штрафную функцию

$$F_\lambda(x) = f(x) + \lambda_1[\varphi(x) + \lambda_2\varphi_1(x)].$$

Пример. Вычисления производились методом усечённого кодифференциального спуска, который подробно описан в [1], на программном обеспечении, написанном в среде MATLAB [3].

Рассмотрим задачу минимизации функции $f(x)$ на множестве точек минимума функции $f_1(x) = ||x_1| + |x_2| - x_1^2 - x_2^2|$ на множестве G , где

$$G = \{x \in R^2 \mid \max(0, |x_1| + |x_2| - 1) = 0\}.$$

Штрафная функция для данной задачи примет вид:

$$F_\lambda(x) = f(x) + \lambda_1[||x_1| + |x_2| - x_1^2 - x_2^2| + \lambda_2 \max(0, |x_1| + |x_2| - 1)],$$

где $f(x) = \psi_k(x)$, $\psi_1(x) = x_1$, $\psi_2(x) = x_2$, $\psi_3(x) = |x_1|$, $\psi_4(x) = |x_2|$.
Оценим точный штрафной параметр, воспользуемся:

$$h^\perp(x) = \sqrt{2}(-1 - |x_1| - |x_2| - \lambda_2) \quad \forall x \in \Omega_\delta \setminus \Omega$$

Найдём a из неравенства $h^\perp(x) < -a$. Тогда, если использовать оценку $|x_1| + |x_2| < \frac{\sqrt{2}}{2} + \delta + \frac{\sqrt{2}}{2} + \delta = \sqrt{2} + 2\delta$, то получим значение $a = \sqrt{2}(1 + \sqrt{2} + 2\delta + \lambda_2)$.

Найдя константу Липшица из условия $\|f'(x)\| \leq L$ и воспользовавшись неравенством

$$\lambda_1^* \geq \frac{2L}{a},$$

имеем оценку точного штрафного параметра

$$\lambda_1^* \geq \frac{2(1 + \delta)}{\sqrt{2}(1 + \sqrt{2} + 2\delta + \lambda_2)}. \quad (1)$$

Рассмотрим множество, на котором ищется минимум $f(x)$.

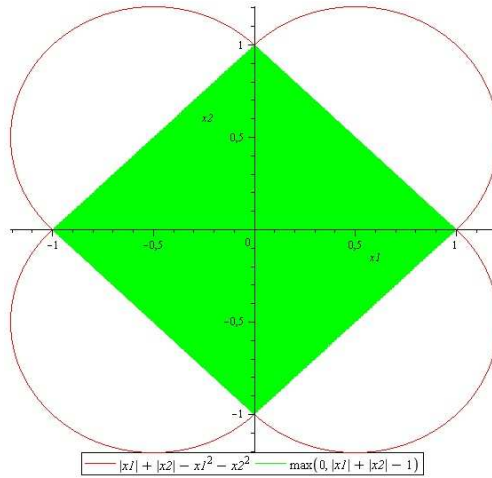


Рис. 1. Штрафные функции

Из рис. 1 видно, что минимум функции $F_\lambda(x)$ будет достигаться в точках пересечения этих двух множеств $(-1, 0)$, $(0, -1)$, $\text{co}\{(0, -1), (0, 1)\}$, $\text{co}\{(-1, 0), (1, 0)\}$ для $f(x) = \psi_k(x)$, $k = 1, \dots, 4$ соответственно. Также из вида штрафной функции очевидно, что точка $(0, 0)$ — точка минимума для функций $f(x) = |x_1|$

и $f(x) = |x_2|$, в зависимости от того, из какой точки мы будем начинать алгоритм.

Примеры вычисления. Рассмотрим, как ведёт себя решение в зависимости от выбора параметров.

1) $f(x) = x_1, \lambda_2 = 10, x = (10, 1)$. Из (1) имеем $\lambda_1^* \geq \frac{2(1+\delta)}{\sqrt{2(11+\sqrt{2+2\delta})}}$.

Таблица 1. Вычисления для $f(x) = x_1$

λ_1	Кол-во шагов	Значение	Полученная точка
0, 1	4	-1	(-1; 0)
0, 091	4	-1	(-1; 0)
0, 09	4	-1, 0003	(-1, 0556; 0)
0, 077	4	-1, 076	(-1, 9935; 0)
0, 076	3	-1, 0885	(-2, 0789; 0)
0, 073	4	-1, 1329	(-2, 3493; 0)
0, 0712	4	-1, 165	(-2, 5225; 0)

Из вычислений видно, что если штрафной параметр мал, то получаем неверное решение.

2) $f(x) = x_2, \lambda_2 = 6, x = (1, 5)$. Тогда получаем $\lambda_1^* \geq \frac{2(1+\delta)}{\sqrt{2(7+\sqrt{2+2\delta})}}$.

Таблица 2. Вычисления для $f(x) = x_2$

λ_1	Кол-во шагов	Значение	Полученная точка
0, 2	4	-1	(0; -1)
0, 15	4	-1	(0; -1)
0, 14	5	-1, 0007	(0; -1, 0714)
0, 13	4	-1, 0156	(0; -1, 3462)
0, 11	4	-1, 12	(0; -2, 0455)
0, 1	4	-1, 225	(0; -2, 5)

Как и в случае 1, при выборе достаточно малого штрафного параметра, получаем неверное решение.

3) $f(x) = |x_1|, \lambda_2 = 3, x = (-3, 7)$. Из (3) имеем $\lambda_1^* \geq \frac{2(1+\delta)}{\sqrt{2(4+\sqrt{2+2\delta})}}$.

Таблица 3. Вычисления для $f(x) = |x_1|$

λ_1	Кол-во шагов	Значение	Полученная точка
0, 167	3	0	(0; 0)
0, 15	3	0	(0; 0)
0, 1	3	0	(0; 0)
0, 02	3	0	(0; -1)
0, 01	3	0	(0; 0)
0, 002	3	0	(0; -1)
0, 0002	3	0	(0; 0)
0, 000002	3	0	(0; 1)

Здесь мы имеем такие результаты в силу того, что функция $f(x) = |x_1|$ имеет свой минимум там, где штрафная функция принимает значение 0.

4) $f(x) = |x_2|, \lambda_2 = 2, x = (-10, 3)$. Тогда имеем $\lambda_1^* \geq \frac{2(1+\delta)}{\sqrt{2(3+\sqrt{2+2\delta})}}$.

Таблица 4. Вычисления для $f(x) = |x_2|$

λ_1	Кол-во шагов	Значение	Полученная точка
0,3	3	0	(0; 0)
0,2	3	0	(-1; 0)
0,19	3	0	(0; 0)
0,15	4	0	(0; 0)
0,0001	3	0	(0; 0)

Аналогично случаю в примере 3, здесь функция $f(x) = |x_2|$ имеет глобальный минимум там, где штраф обнуляется.

Заключение. Как видно из результатов вычислений, если правильно составить штрафную функцию и исследовать значение параметров, можно решать данный класс задач.

Литература

1. Demyanov V.F., Bagirov A.M., Rubinov A.M. A method of truncated codifferential with application to some problems of cluster analysis // Journal of Global Optimization, 2002. Vol. 23. № 1. P. 63–80.
2. Демьянов В.Ф. Условия экстремума и вариационное исчисление. М.: Высшая школа, 2005. 336 с.
3. Андрамонов М.Ю., Тамасян Г.Ш. Реализация аналитического ко-дифференцирования в пакете MATLAB // Вычислительные методы и программирование, 2007. т. 8. с. 1–5.
4. Демьянов В.Ф., Факкиной Ф. Задачи двухуровневой оптимизации и штрафные функции // Известия высших учебных заведений, 2003. № 12. с. 49–61.

Сумачева В.А.

Санкт-Петербургский государственный университет

Вычисление нормы передаточной функции уравнения с запаздываниями с помощью функций Ляпунова

Рекомендовано к публикации профессором Харитоновым В.Л.

Введение. Норма передаточной функции играет важную роль в исследовании динамических систем. В частности, важной сферой ее применения является синтез оптимального управления, при котором должен достигаться минимум \mathcal{H}_2 нормы передаточной функции между внешними воздействиями и выходными сигналами. Известно, что для вычисления \mathcal{H}_2 нормы линейных стационарных систем может быть использована функция Ляпунова. В статье [1] эта теория расширена для случая систем с запаздываниями, где функция Ляпунова вычисляется по-другому, но является ее полным аналогом для систем без запаздываний. Однако результат получен только для систем с одним запаздыванием у входного и выходного сигналов, мы же будем исследовать более общий случай.

Рассмотрим экспоненциально устойчивое линейное стационарное уравнение с несколькими запаздываниями

$$\begin{cases} \dot{x}(t) = a_0x(t) + a_1x(t-h) + \dots + a_mx(t-mh) + \\ \quad + b_0v(t) + b_1v(t-h) + \dots + b_rv(t-rh), \\ y(t) = c_0x(t) + c_1x(t-h) + \dots + c_nx(t-nh), \end{cases} \quad (1)$$

где $a_0, \dots, a_m, b_0, \dots, b_r, c_0, \dots, c_n$ – вещественные коэффициенты; $r \leq m, n \leq m; h > 0$ – положительное запаздывание. Функция $v(t)$ является входным сигналом системы, $y(t)$ – выходным, $x(t)$ – состояние системы.

Передаточная функция. Применим преобразование Лапласа к системе. Предположим, что для входного сигнала оно существует и начальные условия нулевые. Пусть образами $v(t), y(t), x(t)$ являются $V(s), Y(s), X(s)$ соответственно. Тогда система примет вид

$$\begin{cases} (s - a_0 - a_1 e^{-sh} - \dots - a_m e^{-smh})X(s) = \\ = (b_0 + b_1 e^{-sh} + \dots + b_r e^{-srh})V(s), \\ Y(s) = (c_0 + c_1 e^{-sh} + \dots + c_n e^{-snh})X(s). \end{cases}$$

Выразим изображение выхода через изображение входа

$$Y(s) = G(s)V(s).$$

Функция $G(s)$ называется передаточной и позволяет, зная образ входного сигнала, найти образ выходного. Для данной системы она имеет вид

$$G(s) = \frac{(b_0 + b_1 e^{-h} + \dots + b_r e^{-rh})(c_0 + c_1 e^{-h} + \dots + c_n e^{-h})}{s - a_0 - a_1 e^{-h} - \dots - a_m e^{-mh}}. \quad (2)$$

Норма передаточной функции. Мерой зависимости выходного сигнала от входного является норма передаточной функции системы. Найдем \mathcal{H}_2 норму этой функции. По [2] для экспоненциально устойчивой системы она определяется как

$$\|G\|_2^2 = \frac{1}{2\pi} \int_{-\infty}^{\infty} G^2(i\omega) d\omega.$$

Импульсной характеристикой $h(t)$ называется выходной сигнал системы с нулевыми начальными условиями при подаче на вход дельта-функции Дирака $\delta(\tau)$, которая может быть определена следующим образом

$$\delta(t) = \begin{cases} \infty, & t = 0, \\ 0, & t \neq 0, \end{cases}$$

при условии $\int_{-\infty}^{\infty} \delta(t) dt = 1$. Наша система воспринимает импульсный входной сигнал как $m + 1$ импульсов, произошедших в моменты времени $t = 0, t = h, \dots, t = mh$.

Образом дельта-функции является единица, поэтому передаточная функция – образ импульсной характеристики. Тогда по теореме Парсевалья

$$\|G\|_2^2 = \int_0^{\infty} h^2(t) dt. \quad (3)$$

Импульсная характеристика. Найдем импульсную характеристику. Согласно [3] общее решение уравнения (1) имеет вид

$$x(t, \phi) = k(t)\phi(0) + \sum_{j=1}^m a_j \int_{-jh}^0 k(t-jh-\theta)\phi(\theta)d\theta + \\ + \sum_{i=0}^s b_i \int_0^t k(t-\tau)v(\tau-ih)d\tau, \quad t \geq 0,$$

где $k(t)$ – фундаментальное решение (1), являющееся решением уравнения

$$\dot{k}(t) = a_0k(t) + a_1k(t-h) + \dots + a_mk(t-mh)$$

с начальной функцией

$$k(0) = 1, \quad k(\theta) = 0, \quad \theta < 0.$$

Пусть входной сигнал – дельта-функция и начальные условия нулевые $\phi(\theta) = 0, \theta \in [-mh, 0]$. Тогда по свойству дельта-функции

$$\int_0^t k(t-\tau)v(\tau-ih)d\tau = k(t-ih),$$

и состояние системы $x(t)$ будет иметь вид

$$x(t) = b_0k(t) + b_1k(t-h) + \dots + b_rk(t-rh),$$

а соответствующий выходной сигнал, т.е. импульсная характеристика

$$h(t) = c_0[b_0k(t) + \dots + b_s k(t-sh)] + \dots + c_n[b_0k(t-nh) + \dots + \\ + b_r k(t-rh-nh)] = c_0b_0k(t) + \dots + c_0b_r k(t-rh) + \dots + \\ + c_nb_0k(t-nh) + \dots + c_nb_r k(t-(n+r)h).$$

Используем функцию Ляпунова, равную по определению

$$u(\tau) = \int_0^\infty k(t)k(t+\tau)dt.$$

Тогда выражение для нормы окончательно примет вид

$$\|G\|_2^2 = u(0)[c_0^2b_0^2 + (c_0b_1 + c_1b_0)^2 + (c_0b_2 + c_1b_1 + c_1b_0)^2 + \dots + c_n^2b_r^2] + \\ + 2u(h)[c_0b_0(c_0b_1 + c_1b_0) + (c_0b_1 + c_1b_0)(c_0b_2 + c_1b_1 + c_2b_0) + \dots + \\ + (c_{n-1}b_r + c_nb_{r-1})c_nb_r] + \dots + 2u((r+n)h)c_0c_nb_0b_r,$$

что коротко может быть записано

$$\|G\|_2^2 = \sum_{i=0}^{n+r} u(ih) \sum_{\substack{l,k=0 \\ |l-k|=i}}^{n+r} \left(\sum_{\substack{j_1=0,\dots,n \\ j_2=0,\dots,r \\ j_1+j_2=l}} c_{j_1} b_{j_2} \sum_{\substack{j_3=0,\dots,n \\ j_4=0,\dots,r \\ j_3+j_4=k}} c_{j_3} b_{j_4} \right). \quad (4)$$

Например, для одного запаздывания $n = 1, r = 1$ формула будет иметь вид

$$\begin{aligned} \|G\|_2^2 = & u(0)[c_0^2 b_0^2 + (c_0 b_1 + c_1 b_0)^2 + c_1^2 b_1^2] + \\ & + 2u(h)[c_0 b_0(c_0 b_1 + c_1 b_0) + \\ & + (c_0 b_1 + c_1 b_0)c_1 b_1] + 2u(2h)c_0 c_1 b_0 b_1. \end{aligned}$$

Таким образом, вычисление нормы передаточной функции сводится к нахождению функции Ляпунова.

Функция Ляпунова. В работе [4] показано, что $u(\tau)$ можно определить по-другому: как функцию, удовлетворяющую следующим свойствам:

1. $u'(\tau) = a_0 u(\tau) + a_1 u(\tau-h) + \dots + a_m u(\tau-h), \tau \geq 0$ (динамическое свойство).
2. $u(-\tau) = u(\tau), \tau \geq 0$ (свойство симметрии).
3. $a_0 u(0) + a_1 u(h) + \dots + a_m u(mh) = -\frac{1}{2}$ (алгебраическое свойство).

Введем вектор $z(\tau) = (u(\tau + (m-1)h), \dots, u(\tau), \dots, u(\tau - mh))^T$ на промежутке $\tau \in [0, h]$. Продифференцировав его, учитывая динамическое свойство и свойство симметрии, получим что $z(\tau)$ удовлетворяет системе обыкновенных дифференциальных уравнений

$$z' = Az,$$

где матрица A имеет вид

$$A = \begin{pmatrix} a_0 & \cdots & a_{m-1} & a_m & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & a_0 & a_1 & \cdots & a_m \\ -a_m & \cdots & -a_1 & -a_0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & -a_m & -a_{m-1} & \cdots & -a_0 \end{pmatrix}.$$

Решение этой системы находится по формуле Коши

$$z(\tau) = e^{A\tau} z(0).$$

Из алгебраического свойства, определения вектора z и $z(h) = e^{Ah} z(0)$ получаем систему линейных алгебраических уравнений для определения $z(0)$

$$(M + Ne^{Ah})z(0) = (-1, 0, \dots, 0)^T, \quad (5)$$

где матрицы M и N имеют вид

$$M = \begin{pmatrix} 0 & \cdots & 0 & a_0 & \cdots & a_{m-1} & a_m \\ 1 & \cdots & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 1 & 0 & \cdots & 0 & 0 \\ 0 & \cdots & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 1 & 0 \end{pmatrix},$$

$$N = \begin{pmatrix} a_m & a_{m-1} & \cdots & a_0 & 0 & \cdots & 0 \\ 0 & -1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & -1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & -1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & -1 \end{pmatrix}.$$

Определив $z(0)$, можем найти $z(\tau)$ на промежутке $[0, h]$, т.е. знаем $u(\tau)$ при $\tau \in [-mh, mh]$, откуда для формулы нормы получаем $u(0), \dots, u(mh)$.

Чтобы найти $u(\tau)$ в последующих промежутках $[mh, (n+r)h]$, воспользуемся динамическим свойством и методом интегрирования по шагам.

При $\tau \in [mh, (m+1)h]$ имеем

$$u'(\tau) = a_0 u(\tau) + a_1 u(\tau - h) + \dots + a_m u(\tau - mh) = a_0 u(\tau) + f(\tau).$$

Так как в этом случае $\tau - h \in [(m - 1)h, mh], \dots, \tau - mh \in [0, h]$, и функция u на этих промежутках известна, получаем линейное стационарное неоднородное уравнение. Находя его решение по формуле Коши

$$u(\tau) = e^{a_0\tau} \left(e^{a_0mh} u(mh) + \int_{mh}^{\tau} e^{-a_0\xi} f(\xi) d\xi \right),$$

получаем $u(\tau), \tau \in [mh, (m + 1)h]$. Таким образом знаем функцию Ляпунова на промежутке $[0, (m + 1)h]$. Аналогично можно продолжать решение на весь требуемый отрезок $[0, (n + r)h]$ и найти $u(0), u(h), \dots, u((n + r)h)$, что полностью определяет выражение для нормы передаточной функции.

Литература

1. Беллман Р., Кук К. Дифференциально-разностные уравнения. М.: Мир, 1967. 548 с.
2. Jarlebring E., Vanbierviet J., Michiels W. Explicit expression for the \mathcal{H}_2 norm of time-delay system based on the delay Lyapunov equation // Proceedings of the 49th IEEE Conference on Decision and Control. Atlanta. USA, 2010. P. 164–169.
3. Kharitonov V.L. Lyapunov-Krasovskii functional for scalar time delay equation // System & Control Letters, 2004. Vol. 51. P. 133–149.
4. Zhou K., Doyle J.C., Glover K. Robust and Optimal Control. Englewood Cliffs, NJ: Prentice Hall, 1996. 586 p.

Утешев Н.А.

Санкт-Петербургский государственный университет

Идентификация методом изоляции с помощью сферы ¹

Рекомендовано к публикации профессором Демьяновым В.Ф.

Частным случаем задачи идентификации является задача разделения двух множеств сферой. Именно, есть два множества $A = \{a_i | i \in I\}$ и $B = \{b_j | j \in J\}$, $A, B \in \mathbb{R}^n$, и требуется найти такую сферу $S_r(x)$, чтобы идентификационное правило

$$c \in \text{co } S_r(x) \Rightarrow c \in A, \quad c \notin \text{co } S_r(x) \Rightarrow c \in B$$

описывало множества A и B наилучшим в определенном смысле образом. Для выбора оптимального шара вводятся различные функционалы. Наиболее естественным выглядит функционал, который сопоставляет каждой паре (x, r) общее количество неверно определенных точек $F(x, r) = |\{i : \|a_i - x\| > r\}| + |\{j : \|b_j - x\| \leq r\}|$; или функционал, равный максимальному по A и B количеству неверно определенных точек $F(x, r) = \max\{|\{i : \|a_i - x\| > r\}|, |\{j : \|b_j - x\| \leq r\}|\}$. Такие функционалы называют натуральными. Однако нахождение минимума подобных функционалов является проблемой из-за того, что такие функционалы даже не непрерывны. Поэтому для удобства анализа вместо задачи минимизации натурального функционала решают задачу минимизации суррогатного функционала. Суррогатный функционал вводят таким образом, чтобы а) он был неотрицателен и обращался в 0 в том случае, когда все точки идентифицированы правильно; б) к нему можно было бы прилагать аппарат квазидифференциального исчисления (т.е. он должен быть, по крайней мере, непрерывен). Здесь будем рассматривать функционалы вида

$$F_1(x, r) = \sum_{i \in I} \psi_1(a_i, x, r) + \sum_{j \in J} \psi_2(b_j, x, r),$$
$$F_2(x, r) = \max\left\{ \sum_{i \in I} \psi_1(a_i, x, r), \sum_{j \in J} \psi_2(b_j, x, r) \right\},$$

¹Работа выполнена при финансовой поддержке РФФИ, грант № 09-01-00360

где

$$\begin{aligned}\psi_1(c, x, r) &= \max\{0; -f(c, x, r)\}, \\ \psi_2(c, x, r) &= \max\{0; f(c, x, r)\}.\end{aligned}$$

Функция f может задаваться различным образом. Далее рассмотрим необходимые условия минимума для двух вариантов задания f .

$$1. f(c, x, r) = r^2 - (c - x)^2.$$

Введем обозначения

$$\begin{aligned}I^- &= \{i \in I \mid f(a_i, x, r) < 0\}, \\ I^0 &= \{i \in I \mid f(a_i, x, r) = 0\}, \\ J^+ &= \{j \in J \mid f(b_j, x, r) > 0\}, \\ J^0 &= \{j \in J \mid f(b_j, x, r) = 0\}.\end{aligned}$$

Выищем квазидифференциалы функций ψ_1 (при $c = a_i$) и ψ_2 (при $c = b_j$), используя правила вычисления квазидифференциала, приведенные в [1]:

$$\begin{aligned}\bar{\partial}\psi_1 &= \bar{\partial}\psi_2 = \{0_{n+1}\}, \\ \underline{\partial}\psi_1 &= \begin{cases} \{0_{n+1}\}, & \text{если } a_i \in \text{int co } S; \\ [2(x - a_i)], -2r], & \text{если } a_i \notin \text{co } S; \\ \text{co}\{0_{n+1}, [2(x - a_i)], -2r]\}, & \text{если } a_i \in S. \end{cases} \\ \underline{\partial}\psi_2 &= \begin{cases} \{0_{n+1}\}, & \text{если } b_j \notin S; \\ [2(b_j - x)], 2r], & \text{если } b_j \in \text{int co } S; \\ \text{co}\{0_{n+1}, [2(b_j - x)], 2r]\}, & \text{если } b_j \in S. \end{cases}\end{aligned}$$

Отсюда

$$\begin{aligned}\underline{\partial}F_1 &= \sum_{i \in I^-} [2(x - a_i)], -2r] + \sum_{j \in J^+} [2(b_j - x)], 2r] + \\ &+ \sum_{i \in I^0} \text{co}\{0_{n+1}, [2(x - a_i)], -2r]\} + \sum_{j \in J^0} \text{co}\{0_{n+1}, [2(b_j - x)], 2r]\} = \\ &= \{2[x|I^-| - x|J^+| - \sum_{i \in I^-} a_i + \sum_{j \in J^+} b_j + \sum_{i \in I^0} \alpha_i(x - a_i) + \sum_{j \in J^0} \beta_j(b_j - x), \\ &\quad -r|I^-| + r|J^+| - r \sum_{i \in I^0} \alpha_i + r \sum_{j \in J^0} \beta_j] \mid \alpha_i, \beta_j \in [0, 1]\},\end{aligned}$$

$$\bar{\partial}F_1 = \{0_{n+1}\}.$$

Из необходимого условия минимума $-\bar{\partial}F_1 \subset \underline{\partial}F_1$ следует, что при некоторых α_i, β_j последний компонент в выражении для $\underline{\partial}F_1$ обращается в 0. Это возможно в двух случаях. Если $r = 0$, то $|J^+| = 0$, для I^0 и J^0 будет $x - a_i = b_j - x = 0_n$, и условие минимума сводится к совокупности

$$\begin{cases} r = 0, \\ x = \frac{\sum_{i \in I^-} a_i}{|I^-|}. \end{cases}$$

Минимум функционала F_1 будет достигаться в этой точке только в случае, если $|I| = 1$. В самом деле, если рассмотреть шар $S_\varepsilon(x)$, где ε достаточно мало так, чтобы в шар не попало новых точек, то получим

$$\begin{aligned} F_1(x, \varepsilon) &= \sum_{i \in I^-} (-\varepsilon^2 + (a_i - x)^2) + \sum_{j \in J^+} (-(b_j - x)^2 + \varepsilon^2) = \\ &= \sum_{i \in I^-} (a_i - x)^2 - \sum_{j \in J^+} (b_j - x)^2 + \varepsilon^2(|J^+| - |I^-|) = \\ &= F_1(x, 0) + \varepsilon^2(|J^+| - |I^-|) \end{aligned}$$

Множество J^+ может содержать либо 0, либо одну точку, совпадающую с x . В первом случае, так как $|I| > 1$, найдется хотя бы одна точка из I^- . Поэтому $|J^+| - |I^-| < 0$. Во втором случае из условия $x = \frac{\sum_{i \in I^-} a_i}{|I^-|}$ следует, что $|I^-| \geq 2$ (в противном случае точка x принадлежала бы к A и B одновременно), поэтому снова $|J^+| - |I^-| < 0$. Отсюда $F_1(x, \varepsilon) < F_1(x, 0)$, т.е. на исходном шаре не может достигаться минимум. Поэтому рассмотрим второй случай.

Введем обозначения: $C = A \cup B, \bar{c}_i = [c_i, 1]$. Докажем несколько лемм.

Лемма 1. Пусть $\bar{c}_k, k \in \overline{1, l}$, линейно независимы. Множество точек, равноудаленных от c_k , образуют линейное многообразие размерности $n + 1 - l$.

Лемма 2. Пусть $m < l$; $\bar{c}_k, k \in \overline{1, m}$, образуют базис $\{\bar{c}_k, k \in \overline{1, l}\}$. Если некоторая сфера проходит через $\{\bar{c}_k, k \in \overline{1, l}\}$, то все остальные сферы, проходящие через $\{\bar{c}_k, k \in \overline{1, m}\}$, тоже содержат все эти точки.

Лемма 3. Если $\|x_1 - c_1\| = \|x_1 - c_2\|$ и $\|x_2 - c_1\| = \|x_2 - c_2\|$, то $x_1 - x_2 \perp c_1 - c_2$.

Теорема 1. Пусть \overline{C} – матрица, составленная из векторов типа \overline{c}_i . Если ранг матрицы \overline{C} равен k и если минимум функционала F_1 существует, то он достигается на сфере, проходящей как минимум через k точек из C , причем соответствующие столбцы \overline{C} ранговые.

Следствие 1. Если минимум функционала F_1 существует, то он достигается на сфере с центром в линейном многообразии L , образованном точками из C .

Таким образом, если минимум функционала F_1 существует, алгоритм решения задачи следующий:

- строим матрицу \overline{C} ;
- выбираем набор векторов $\overline{a}_i, \overline{b}_j$, ранговый для \overline{C} . Проводим через соответствующие векторы a_i, b_j шар и решаем для него уравнение: $-\sum_{i \in I^-} \overline{a}_i + \sum_{j \in J^+} \overline{b}_j - \sum_{i \in I^0} \alpha_i \overline{a}_i + \sum_{j \in J^0} \beta_j \overline{b}_j = 0 \Leftrightarrow \sum_{j \in J^0} \beta_j \overline{b}_j - \sum_{i \in I^0} \alpha_i \overline{a}_i = \sum_{i \in I^-} \overline{a}_i - \sum_{j \in J^+} \overline{b}_j$. Уравнение заведомо решается, поскольку среди $\overline{b}_j, j \in J^0, \overline{a}_i, i \in I^0$, имеются ранговые векторы для матрицы \overline{C} , а выражение $\sum_{i \in I^-} \overline{a}_i - \sum_{j \in J^+} \overline{b}_j$ является линейной комбинацией столбцов матрицы \overline{C} ;
- если получившиеся числа $\alpha_i, \beta_j \in [0, 1]$, то вычисляем на построенном шаре функционал F_1 ;
- среди построенных шаров (их не больше $C_{|I|+|J|}^k$, где k – ранг матрицы \overline{C}) выбираем тот, на котором значение функционала F_1 минимально.

Рассмотрим теперь функционал

$$F_2(x, r) = \max\left\{\sum_{i \in I} \psi_1(a_i, x, r), \sum_{j \in J} \psi_2(b_j, x, r)\right\}.$$

$$\frac{\partial F_2}{\partial x} = \sum_{i \in I^-} [2(x - a_i), -2r] + \sum_{i \in I^0} \text{co}\{0_{n+1}, [2(x - a_i), -2r]\},$$

если $\sum \psi_1 > \sum \psi_2$;

$$\begin{aligned}
& \sum_{j \in J^+} [2(b_j - x), 2r] + \sum_{i \in I^0} \text{co}\{0_{n+1}, [2(b_j - x), 2r]\}, \\
& \text{если } \sum \psi_1 < \sum \psi_2; \\
& \text{co}\left\{ \sum_{i \in I^-} 2[x - a_i, -r] + \sum_{i \in I^0} \text{co}\{0_{n+1}, 2[x - a_i, -r]\}, \right. \\
& \left. \sum_{j \in J^+} 2[b_j - x, r] + \sum_{j \in J^0} \text{co}\{0_{n+1}, 2[b_j - x, r]\} \right\}, \\
& \text{если } \sum \psi_1 = \sum \psi_2 \\
& = 2[x(|I^-| + \sum_{i \in I^0} \alpha_i) - \sum_{i \in I^-} a_i - \sum_{i \in I^0} \alpha_i a_i, -r(|I^-| + \sum_{i \in I^0} \alpha_i)], \\
& \text{если } \sum \psi_1 > \sum \psi_2; \\
& 2[\sum_{j \in J^+} \beta_j + \sum_{j \in J^0} \beta_j b_j - x(|J^+| + \sum_{j \in J^0} \beta_j), r(|J^+| + \sum_{j \in J^0} \beta_j)], \\
& \text{если } \sum \psi_1 < \sum \psi_2; \\
& \text{co}\{2[x(|I^-| + \sum_{i \in I^0} \alpha_i) - \sum_{i \in I^-} a_i - \sum_{i \in I^0} \alpha_i a_i], -r(|I^-| + \sum_{i \in I^0} \alpha_i)\}, \\
& 2[\sum_{j \in J^+} \beta_j + \sum_{j \in J^0} \beta_j b_j - x(|J^+| + \sum_{j \in J^0} \beta_j), r(|J^+| + \sum_{j \in J^0} \beta_j)], \\
& \text{если } \sum \psi_1 = \sum \psi_2.
\end{aligned}$$

Здесь $\alpha, \beta \in [0, 1]$. В первых двух случаях минимум достигаться не может. В самом деле, допустим $r = 0$. В первом случае здесь могут быть применены те же рассуждения, что и для функционала F_1 . Во втором получаем $\sum \psi_2 = 0$, отсюда $\sum \psi_1 < 0$, что невозможно. Поэтому имеем в первом случае $|I^-| + \sum_{i \in I^0} \alpha_i = 0$; во втором

$$|J^+| + \sum_{j \in J^0} \beta_j = 0 \Rightarrow \sum \psi_2 = 0 \Rightarrow \sum \psi_1 < 0 \Rightarrow \text{false}.$$

Остается третий случай. Ноль принадлежит выпуклой комбинации двух множеств в том случае, если среди точек этих множеств есть пропорциональные. Отсюда

$$\begin{aligned}
& \sum_{i \in I^-} \psi_1(a_i, x, r) = \sum_{j \in J^+} \psi_2(b_j, x, r) \Leftrightarrow \\
& \Leftrightarrow \sum_{i \in I^-} (a_i - x)^2 - |I^-| r^2 = |J^+| r^2 - \sum_{j \in J^+} (b_j - x)^2;
\end{aligned}$$

$$\begin{aligned}
& (x(|I^-| + \sum_{i \in I^0} \alpha_i) - \sum_{i \in I^-} a_i - \sum_{i \in I^0} \alpha_i a_i) \frac{|J^+| + \sum_{j \in J^0} \beta_j}{|I^-| + \sum_{i \in I^0} \alpha_i} = \\
& = x(|J^+| + \sum_{j \in J^0} \beta_j) - \sum_{j \in J^+} \beta_j - \sum_{j \in J^0} \beta_j b_j \Leftrightarrow \\
& \Leftrightarrow \frac{\sum_{i \in I^-} a_i + \sum_{i \in I^0} \alpha_i a_i}{|I^-| + \sum_{i \in I^0} \alpha_i} = \frac{\sum_{j \in J^+} b_j + \sum_{j \in J^0} \beta_j b_j}{|J^+| + \sum_{j \in J^0} \beta_j}.
\end{aligned}$$

Таким образом, получаем необходимые условия минимума:

$$\begin{cases} \sum_{i \in I^-} (a_i - x)^2 - |I^-| r^2 = |J^+| r^2 - \sum_{j \in J^+} (b_j - x)^2; \\ \frac{\sum_{i \in I^-} a_i + \sum_{i \in I^0} \alpha_i a_i}{|I^-| + \sum_{i \in I^0} \alpha_i} = \frac{\sum_{j \in J^+} b_j + \sum_{j \in J^0} \beta_j b_j}{|J^+| + \sum_{j \in J^0} \beta_j}. \end{cases}$$

Эти условия, однако, не дают ясного алгоритма построения оптимальной сферы. В частном случае, если на сфере нет точек, получаем условие $\frac{\sum_{i \in I^-} a_i}{|I^-|} = \frac{\sum_{j \in J^+} b_j}{|J^+|}$, т.е. центры тяжести неверно определенных точек из A и неверно определенных точек из B совпадают. При любом непрерывном распределении точек a_i и b_j , однако, вероятность такого события будет нулевой.

2. $f(c, x, r) = r - \|c - x\|$.

Обозначения примем таким же образом, как в предыдущем пункте. Квазидифференциалы функций ψ_1 (при $c = a_i$) и ψ_2 (при $c = b_j$) будут иметь вид:

$$\begin{aligned}
& \bar{\partial}\psi_1 = \bar{\partial}\psi_2 = \{0_{n+1}\}; \\
& \underline{\partial}\psi_1 = \begin{cases} \left\{ \begin{array}{l} \{0_{n+1}\}, \\ \left[\frac{x - a_i}{\|x - a_i\|}, -1 \right], \\ \text{co} \left\{ 0_{n+1}, \left[\frac{x - a_i}{\|x - a_i\|}, -1 \right] \right\}, \end{array} \right. & \begin{array}{l} \text{если } a_i \in \text{int co } S; \\ \text{если } a_i \notin \text{co } S; \\ \text{если } a_i \in S. \end{array} \end{cases};
\end{aligned}$$

$$\underline{\partial}\psi_2 = \begin{cases} \{0_{n+1}\}, & \text{если } b_j \notin \text{co } S; \\ \left[\frac{b_j - x}{\|b_j - x\|}, 1 \right], & \text{если } b_j \in \text{int co } S; \\ \text{co} \left\{ 0_{n+1}, \left[\frac{b_j - x}{\|b_j - x\|}, 1 \right] \right\}, & \text{если } b_j \in S. \end{cases}$$

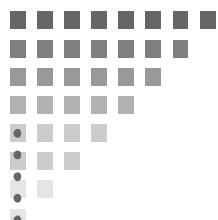
Отсюда

$$\begin{aligned} \underline{\partial}F_1 &= \sum_{i \in I^-} \left[\frac{x - a_i}{\|x - a_i\|}, -1 \right] + \sum_{j \in J^+} \left[\frac{b_j - x}{\|b_j - x\|}, 1 \right] + \\ &+ \sum_{i \in I^0} \text{co} \left\{ 0_{n+1}, \left[\frac{x - a_i}{\|x - a_i\|}, -1 \right] \right\} + \sum_{j \in J^0} \text{co} \left\{ 0_{n+1}, \left[\frac{b_j - x}{\|b_j - x\|}, 1 \right] \right\} = \\ &= \left\{ \left[\sum_{i \in I^-} \frac{x - a_i}{\|x - a_i\|} + \sum_{i \in I^0} \frac{\alpha_i(x - a_i)}{r} + \sum_{j \in J^+} \frac{b_j - x}{\|b_j - x\|} + \sum_{j \in J^0} \frac{\beta_j(b_j - x)}{r}, \right. \right. \\ &\quad \left. \left. -|I^-| + |J^+| - \sum_{i \in I^0} \alpha_i + \sum_{j \in J^0} \beta_j \mid \alpha_i, \beta_j \in [0, 1] \right\}; \\ \bar{\partial}F_1 &= \{0_{n+1}\}. \end{aligned}$$

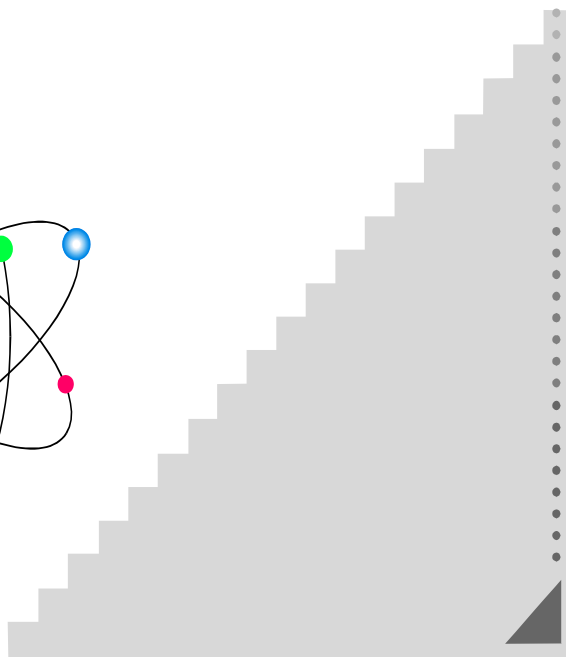
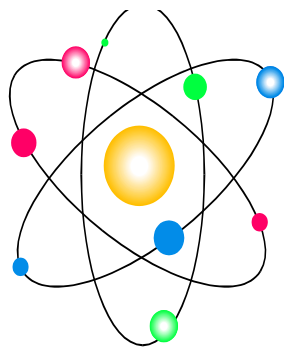
Здесь тоже не удастся построить алгоритма поиска оптимальной сферы. Таким образом, исследование этих функционалов остается делом будущего.

Литература

1. Демьянов В.Ф., Васильев Л.В. Недифференцируемая оптимизация. М.: Наука, 1981. 384 с.



*2. Математические
методы в механике и
физике*



Алцыбеев В.В.

Санкт-Петербургский государственный университет

Параметрическая оптимизация динамики пучков в гибридных системах

Рекомендовано к публикации профессором Овсянниковым Д.А.

1. Постановка задачи. В статье рассмотрена задача оптимального управления пучком траекторий в случае, когда динамика управляемого процесса описывается гибридной системой уравнений, имеющей непрерывную и дискретную часть. Задачи управления для систем с переменной структурой исследовались ранее в работах Емельянова [1], Мироновой [2], Розовой [5] и в других. Например, такая задача может возникнуть при расчёте и оптимизации динамики пучков заряженных частиц в ускорителях с трубками дрейфа [3, 4]. В таком случае непрерывная часть системы соответствует движению частиц в зазорах ускорителя, а дискретная – в трубках дрейфа. Пусть динамика пучка описывается следующей системой уравнений:

$$\begin{aligned} \dot{x} &= f_1(t, x, u), \quad t \in [\mu_{i-1}, \mu_i], \\ x(\mu_{i+1}) &= f_2(\mu_i, x(\mu_i), v_i), \quad t \in [\mu_i, \mu_{i+1}), \\ x(0) &= x_0 \in \bar{M}_0, \\ i &= 2k + 1, \quad k = 0, \dots, N. \end{aligned} \tag{1}$$

Здесь $f_1(t, x, u)$ и $f_2(\mu_i, x(\mu_i), v_i)$ – n -мерные вектор-функции, определённые и непрерывные по совокупности аргументов (t, x, u) и (x, v) на $[\mu_{i-1}, \mu_i] \times \Omega \times U$ и $\Omega \times V(i)$, соответственно, вместе с частными производными по этим переменным до второго порядка включительно. W – открытая область в R^n , U и $V(i)$ – компактные множества в R^r , $u(t)$ – r -мерная вектор-функция управления. Управления $u(t)$ составляют класс D кусочно-непрерывных функций со значениями в U , $v_i \in V(i)$ – параметры управления, M_0 – открытое ограниченное множество, содержащееся в Ω . Введем функционал

$$\int_{M_{T,u}} g(x_T) dx_T, \tag{2}$$

характеризующий состояние пучка на выходе системы. Здесь $g(x_T)$ – непрерывно-дифференцируемая функция, определенная в Ω , $M_{T,u}$ – сечение пучка траекторий в момент времени T , т.е. $x_T \in M_{T,u}$. Будем рассматривать задачу минимизации функционала (2) по всем допустимым управлениям.

2. Вспомогательные соотношения. Рассмотрим уравнение в вариациях для системы (1) [3] при вариации управления $\tilde{u} = u + \Delta u$

$$\begin{aligned} \frac{d\delta x}{dt} - \frac{\partial f_1(t, x, u)}{\partial x} \delta x - \Delta_u f_1(t, x, u) &= 0, \quad t \in [\mu_{i-1}, \mu_i], \\ \delta x(\mu_{i+1}) - \frac{\partial f_2(\mu_i, x, v_i)}{\partial x} \delta x(\mu_i) - \\ - \Delta_u f_2(\mu_i, x, v_i) &= 0, \quad t \in [\mu_i, \mu_{i+1}]. \end{aligned} \quad (3)$$

с начальными условиями

$$\delta x(0) = 0.$$

Далее будем рассматривать случай, когда $\text{div}_x f_1(t, x, u) = 0$ и $\left| \frac{\partial}{\partial x} f_2(\mu_i, x(\mu_i), v_i) \right| = 1$. Выпишем вариацию функционала (2) [3]

$$\delta I = \int_{M_{T,u}} \frac{\partial g(x_T)}{\partial x} \delta x_T dx_T. \quad (4)$$

3. Вариация функционала. Преобразуем вариацию (4) функционала (2). Для этого введем вспомогательную систему

$$\begin{aligned} \frac{d\psi}{dt} &= - \left(\frac{\partial}{\partial x} f_1(t, x, u) \right)^* \psi, \quad t \in [\mu_{i-1}, \mu_i], \\ \psi(\mu_i) &= \left(\frac{\partial}{\partial x} f_2(\mu_i, x(\mu_i), v_i) \right)^* \psi(\mu_{i+1}), \quad t \in [\mu_i, \mu_{i+1}]. \end{aligned} \quad (5)$$

с конечными условиями

$$\psi(T, x(T)) = - \left(\frac{\partial g(x_T)}{\partial x} \right)^*, \quad (6)$$

где «*» означает транспонирование. Умножим функцию $(\psi(t))^*$ на уравнение в вариациях (3), проинтегрируем это выражение по сечениям пучка траекторий. Затем проинтегрируем по времени непрерывные части полученного выражения и просуммируем по дискретным. Полученное в итоге выражение, очевидно, равно нулю. Прибавив его к вариации (4), получим

$$\begin{aligned} \delta I = & - \sum_{k=0}^N \left[\int_{\mu_{i-1}}^{\mu_i} \int_{M_{t,u}} \psi^*(t) \Delta_u f_1(t, x, u) dx_t dt + \right. \\ & \left. + \int_{M_{\mu_{i+1},u}} \psi^*(\mu_{i+1}) \Delta_u f_2(\mu_i, x, v_i) dx_{\mu_{i+1}} \right], \quad (7) \\ & i = 2k + 1. \end{aligned}$$

4. Градиент функционала. Управление u непрерывной части системы зададим кусочно-постоянной функцией, зависящей от параметров μ_i , и $\alpha_i = \text{const}$, т.е. $u = \alpha_i$, $t \in [\mu_{i-1}, \mu_i]$.

Для дискретной части системы управляющими параметрами будем считать параметры μ_{i+1} . Для ускорителя с трубками дрейфа эти параметры будут соответствовать величине напряженности в зазорах ускорителя и координатам начал и концов трубок дрейфа. Следовательно, функционал (2) можно рассматривать как функцию параметров μ_i , μ_{i+1} и α_i . Таким образом, можно определить частные производные

$$\frac{\partial I}{\partial \alpha_i} = - \int_{\mu_{i-1}}^{\mu_i} \int_{M_{t,u}} \psi^*(t) \frac{\partial f_1(t, x(t), \alpha_i)}{\partial \alpha} dx_t dt, \quad (8)$$

$$\begin{aligned} \frac{\partial I}{\partial \mu_i} = & - \int_{M_{\mu_i,u}} \psi^*(\mu_i) f_1(\mu_i, x(\mu_i), \alpha_{i+1}) dx_{\mu_i} - \\ & - \int_{M_{\mu_{i+1},u}} \psi^*(\mu_{i+1}) \frac{\partial f_2(\mu_i, x(\mu_i), \mu_{i+1})}{\partial \mu_i} dx_{\mu_{i+1}}, \quad (9) \end{aligned}$$

$$\begin{aligned} \frac{\partial I}{\partial \mu_{i+1}} = & \int_{M_{\mu_{i+1}, u}} \psi^*(\mu_{i+1}) f_1(\mu_{i+1}, x(\mu_{i+1}), \alpha_{i+2}) dx_{\mu_{i+1}} - \\ & - \int_{M_{\mu_i, u}} \psi^*(\mu_{i+1}) \frac{\partial f_2(\mu_i, x(\mu_i), \mu_{i+1})}{\partial \mu_{i+1}} dx_{\mu_{i+1}}. \end{aligned} \quad (10)$$

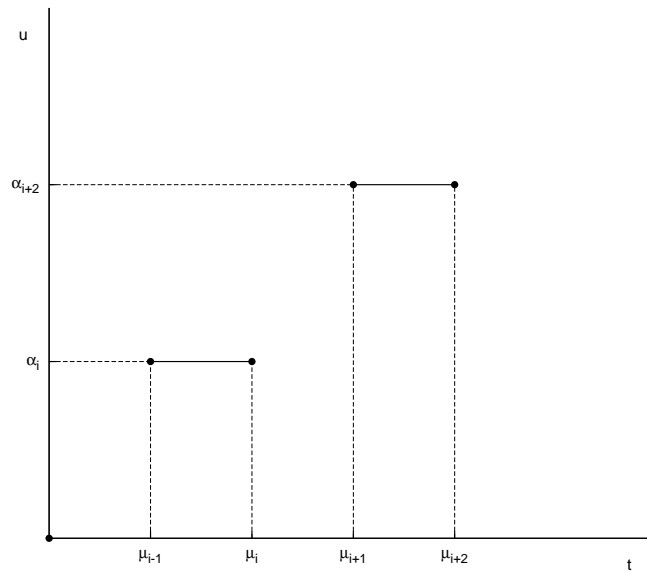


Рис. 1. Структура функции $u(t)$

Введём градиент функционала $l(t, u)$, имеющий компонентами частные производные, определенные в (8)–(10). На основе полученных результатов сформулируем следующее утверждение.

Теорема. Пусть управление u_0 – оптимальное. Тогда

$$(l(t, u_0), (u - u_0)) \geq 0$$

для любых u , таких, что $u - u_0 \in U$.

Доказательство проводится стандартным образом, аналогично тому, как в работе [3].

5. Заключение. В статье получены вариация и градиент функционала качества в случае, когда динамика управляемого процесса описывается гибридной системой уравнений. Сформулированы необходимые условия оптимальности. Полученные формулы (8)–(10) могут быть использованы для алгоритма градиентной минимизации функционала (2) по управляющим параметрам.

Литература

1. Емельянов С.В. Системы с переменной структурой. М.: Наука, 1967. 336 с.
2. Миронова Р.С. Об одной задаче оптимального управления с переменной структурой. В кн. Численные методы в математической физике. М.: Изд-во МГУ, 1986. С. 86–87.
3. Овсянников Д.А. Математические методы управления пучками. Л.: Изд-во ЛГУ, 1990. 228 с.
4. Овсянников Д.А, Дривотин О.И. Моделирование интенсивных пучков заряженных частиц. СПб.: Изд-во СПбГУ, 2003. 175 с.
5. Розова В.Н. Оптимальное управление ступенчатыми системами с неинтегральным функционалом // Вестник РУДН. Сер. 1, 2002. Вып. 1. С. 131–136.

Басинский К.Ю.

Тюменский государственный университет

Распространение капиллярно-гравитационных волн по свободной поверхности слабовязкой жидкости

Рекомендовано к публикации доцентом Бариновым В.А.

Рассматривается распространение капиллярно-гравитационной волны длиной λ в направлении оси x^* по свободной поверхности $z^* = \xi^*(t^*, x^*)$ бесконечно глубокого слоя вязкой несжимаемой жидкости. Ось z^* противоположно направлена вектору силы тяжести \mathbf{g} , плоскость $z^* = 0$ совпадает с невозмущенной свободной поверхностью. Волновое движение жидкости происходит в плоскости x^*z^* со скоростью $\mathbf{u}^* = (u^*(t^*, x^*, z^*), 0, v^*(t^*, x^*, z^*))$. Звездочкой обозначены, где это необходимо, физические (размерные) величины. В работе [1] показано, что в случае пологих волн x^* -ой компонентой нормальной проекции тензора скачка напряжения можно пренебречь по сравнению с z^* -ой компонентой. Кроме того, если отношение вязкой частоты к частоте в идеальной жидкости не превышает 0,4, то решение задачи можно искать в виде затухающих потенциальных волн. Эти ограничения определяют модель волнового движения слабовязкой жидкости. Чтобы решать соответствующую нелинейную краевую задачу первым методом Стокса [2], частоту волны полагаем неизвестной функцией времени. В работе [3] эта задача решена для гравитационных волн. В данной работе рассматриваются капиллярно-гравитационные волны, т.е. распространение волн с учетом силы поверхностного натяжения.

Нелинейная краевая задача о волновом движении слабовязкой жидкости в безразмерном виде запишется

$$\operatorname{div} \mathbf{u} = 0, \quad \frac{\alpha^2}{\alpha - t\alpha'} \frac{\partial \mathbf{u}}{\partial t} - \nu_0 \Delta \mathbf{u} + \nabla p = -\varepsilon (\mathbf{u} \nabla) \mathbf{u}, \quad (1)$$

$$v - \frac{\alpha^2}{\alpha - t\alpha'} \frac{\partial \xi}{\partial t} = \varepsilon u \frac{\partial \xi}{\partial x}, \quad z = \varepsilon \xi, \quad (2)$$

$$p - \gamma_0^2 \xi + \kappa_0^2 \frac{\partial^2 \xi}{\partial x^2} \left[1 + \varepsilon^2 \left(\frac{\partial \xi}{\partial x} \right)^2 \right]^{-3/2} - 2\nu_0 \frac{\partial v}{\partial z} = \quad (3)$$

$$= -\varepsilon\nu_0 \left(\frac{\partial u}{\partial z} + \frac{\partial v}{\partial x} \right) \frac{\partial \xi}{\partial x}, \quad z = \varepsilon\xi, \quad \mathbf{u} \rightarrow 0, \quad z \rightarrow -\infty.$$

Здесь $\alpha = c/c_0 = \omega/\omega_0$, $\alpha' = d\alpha/dt$, $c = c(t)$, $\omega = \omega(t)$ – истинная фазовая скорость и частота волны соответственно, ω_0 , c_0 – частота и фазовая скорость волны линейной задачи для идеальной жидкости.

Безразмерные и физические величины связаны равенствами:

$$\begin{aligned} \mathbf{u}^* &= \varepsilon c_0 \mathbf{u}, \quad p^* = \varepsilon \rho c_0^2 p, \quad p^* = P - P_a + \rho g z, \quad \xi^* = \varepsilon \xi / k, \\ \nu_0 &= \nu^* k / c_0, \quad t = k c t^*, \quad x = k x^*, \quad z = k z^*, \quad k = 2\pi / \lambda, \\ c_0^2 &= c_g^2 + c_\sigma^2, \quad c_g^2 = g/k, \quad c_\sigma^2 = \sigma k / \rho, \\ \gamma_0^2 &= c_g^2 / c_0^2, \quad \kappa_0^2 = c_\sigma^2 / c_0^2, \quad \gamma_0^2 + \kappa_0^2 = 1, \end{aligned}$$

где ρ – плотность, P – давление, P_a – атмосферное давление, ξ^* – форма свободной поверхности, ν^* – коэффициент кинематической вязкости, λ – длина волны, σ – коэффициент поверхностного натяжения.

В силу малости волнового параметра ε условия (2), (3) разложением в ряд Маклорена входящих в них функций можно свести к условиям на фиксированной поверхности $z = 0$.

Скорость волнового движения будем искать в виде затухающих прогрессивных волн

$$\begin{aligned} \mathbf{u} &= \sum_{n=1}^{\infty} \varepsilon^{n-1} e^{-n\beta t/\alpha} \mathbf{u}_n, \quad \mathbf{u}_n = (u_n, 0, v_n), \quad v_n = e^{nz} \{A_n \cos n\bar{x} + \\ &+ B_n \sin n\bar{x}\}, \quad u_n = e^{nz} \{B_n \cos n\bar{x} - A_n \sin n\bar{x}\}, \quad \bar{x} = x - t. \end{aligned}$$

Здесь $\beta(t)$ – безразмерный декремент затухания ($\beta^* = \beta k c_0$ – размерный). Остальные неизвестные функции будем также искать в виде рядов по малому параметру ε :

$$\begin{aligned} p &= \sum_{n=1}^{\infty} \varepsilon^{n-1} e^{-n\beta t/\alpha} p_n, \quad \xi = \sum_{n=1}^{\infty} \varepsilon^{n-1} e^{-n\beta t/\alpha} \xi_n, \\ \alpha &= \alpha_0 \left(1 + \sum_{n=1}^{\infty} \varepsilon^n a_n(t) \right), \quad \beta = \beta_0 \left(1 + \sum_{n=1}^{\infty} \varepsilon^n b_n(t) \right), \\ e^{-\beta t/\alpha} &= e^{-\beta_0 t/\alpha_0} [1 + \varepsilon \beta_0 t/\alpha_0 (a_1 - b_1) + \varepsilon^2 \beta_0 t/\alpha_0 (a_2 - b_2 + \end{aligned}$$

$$+a_1 b_1 - a_1^2 + \beta_0 t / (2\alpha_0) (b_1 - a_1)^2 + \dots].$$

Подставив в уравнения (1) вместо \mathbf{u} , p , ξ , α_0 , β_0 ряды и разложенные в окрестности $z = 0$ граничные условия (2), (3), приравнявая коэффициенты при одинаковых степенях малого параметра, получаем задачи соответствующих приближений. В первом (линейном) приближении задача имеет вид: при ε^0

$$\begin{aligned} \operatorname{div} \mathbf{u}_1 &= 0, \quad \alpha_0 \frac{\partial \mathbf{u}_1}{\partial t} - \beta_0 \mathbf{u}_1 - \nu_0 \Delta \mathbf{u}_1 + \nabla p_1 = 0, \quad (4) \\ v_1 &= \alpha_0 \frac{\partial \xi_1}{\partial t} - \beta_0 \xi_1, \quad p_1 - \gamma_0^2 \xi_1 + \kappa_0^2 \frac{\partial^2 \xi_1}{\partial x^2} - 2\nu_0 \frac{\partial v_1}{\partial z} = 0, \quad z = 0; \end{aligned}$$

для второго приближения: при ε^1

$$\begin{aligned} \operatorname{div} \mathbf{u}_2 &= 0, \quad \alpha_0 \frac{\partial \mathbf{u}_2}{\partial t} - 2\beta_0 \mathbf{u}_2 - \nu_0 \Delta \mathbf{u}_2 + \nabla p_2 = -(\mathbf{u}_1 \nabla) \mathbf{u}_1 + \\ &+ e^{\beta_0 t / \alpha_0} \left[\beta_0 \mathbf{u}_1 (tb_1)' - \alpha_0 \frac{\partial \mathbf{u}_1}{\partial t} (ta_1)' \right], \\ v_2 - \alpha_0 \frac{\partial \xi_2}{\partial t} + 2\beta_0 \xi_2 &= \frac{\partial}{\partial x} (u_1 \xi_1) + e^{\beta_0 t / \alpha_0} \left[\alpha_0 (ta_1)' \frac{\partial \xi_1}{\partial t} - \right. \\ &\left. - \beta_0 \xi_1 (tb_1)' \right], \quad z = 0, \\ p_2 - \gamma_0^2 \xi_2 + \kappa_0^2 \frac{\partial^2 \xi_2}{\partial x^2} - 2\nu_0 \frac{\partial v_2}{\partial z} &= \xi_1 \frac{\partial}{\partial z} \left(2\nu_0 \frac{\partial v_1}{\partial z} - p_1 \right) - \\ &- \nu_0 \left(\frac{\partial u_1}{\partial z} + \frac{\partial v_1}{\partial x} \right) \frac{\partial \xi_1}{\partial x}, \quad z = 0; \end{aligned}$$

для третьего приближения: при ε^2

$$\begin{aligned} \operatorname{div} \mathbf{u}_3 &= 0, \quad \alpha_0 \frac{\partial \mathbf{u}_3}{\partial t} - 3\beta_0 \mathbf{u}_3 - \nu_0 \Delta \mathbf{u}_3 + \nabla p_3 = -[(\mathbf{u}_2 \nabla) \mathbf{u}_1 + \\ &+ (\mathbf{u}_1 \nabla) \mathbf{u}_2] + e^{2\beta_0 t / \alpha_0} \left\{ -\frac{\partial \mathbf{u}_1}{\partial t} [\alpha_0 (ta_2)' + \beta_0 (ta_1)' (tb_1 - ta_1) + \right. \\ &+ \alpha_0 t^2 (a_1')^2] + \beta_0 \mathbf{u}_1 \left[(tb_2)' + \frac{\beta_0}{\alpha_0} (tb_1)' (tb_1 - ta_1) + t^2 a_1' b_1' \right] \left. \right\} + \\ &+ e^{\beta_0 t / \alpha_0} \left[2\beta_0 \mathbf{u}_2 (tb_1)' - \alpha_0 \frac{\partial \mathbf{u}_2}{\partial t} (ta_1)' \right], \end{aligned}$$

$$\begin{aligned}
v_3 = & \alpha_0 \frac{\partial \xi_3}{\partial t} - 3\beta_0 \xi_3 + \frac{\partial}{\partial x} \left[u_1 \xi_2 + u_2 \xi_1 + \frac{1}{2} \frac{\partial u_1}{\partial z} \xi_1^2 \right] + e^{2\beta_0 t / \alpha_0} \times \\
& \times \left\{ \frac{\partial \xi_1}{\partial t} \left[\alpha_0 (ta_2)' + \alpha_0 (a_1')^2 t^2 + \beta_0 (ta_1)' (tb_1 - ta_1) \right] - \right. \\
& \left. - \beta_0 \xi_1 \left[(tb_2)' + \frac{\beta_0}{\alpha_0} (tb_1)' (tb_1 - ta_1) + t^2 a_1' b_1' \right] \right\} + \\
& + e^{\beta_0 t / \alpha_0} \left\{ \alpha_0 (ta_1)' \frac{\partial \xi_2}{\partial t} - 2\beta_0 \xi_2 (tb_1)' \right\}, \quad z = 0, \\
p_3 - \gamma_0^2 \xi_3 + \kappa_0^2 \frac{\partial^2 \xi_3}{\partial x^3} - 2\nu_0 \frac{\partial v_3}{\partial z} = & \xi_1 \frac{\partial}{\partial z} \left(2\nu_0 \frac{\partial v_2}{\partial z} - p_2 \right) + \frac{\partial}{\partial z} \left(\xi_2 + \right. \\
& + \frac{1}{2} \xi_1^2 \frac{\partial}{\partial z} \left. \right) \left(2\nu_0 \frac{\partial v_1}{\partial z} - p_1 \right) + \frac{3}{2} \kappa_0^2 \left(\frac{\partial \xi_1}{\partial x} \right)^2 \frac{\partial^2 \xi_1}{\partial x^2} - \nu_0 \left(\frac{\partial \xi_2}{\partial x} + \right. \\
& \left. + \xi_1 \frac{\partial \xi_1}{\partial x} \frac{\partial}{\partial z} \right) \left(\frac{\partial u_1}{\partial z} + \frac{\partial v_1}{\partial x} \right) - \nu_0 \left(\frac{\partial u_2}{\partial z} + \frac{\partial v_2}{\partial x} \right) \frac{\partial \xi_1}{\partial x}, \quad z = 0.
\end{aligned}$$

Первое приближение (4) совпадает с линейной задачей, рассмотренной в [1]. Ее решение имеет вид

$$\begin{aligned}
\mathbf{u}_1 = & A e^z (\cos \chi, \sin \chi), \quad p_1 = A e^z (\alpha_0 \cos \chi + \nu_0 \sin \chi), \\
\xi_1 = & A (\alpha_0 \cos \chi - \nu_0 \sin \chi), \quad \chi = x - t + \theta, \quad A = \sqrt{A_1^2 + B_1^2}, \\
\theta = & \operatorname{arctg} (A_1 / B_1), \quad \alpha_0^2 + \nu_0^2 = 1, \quad \beta_0 = \nu_0.
\end{aligned}$$

Подставляя решения предыдущих приближений в последующие, получаем задачи в виде системы линейных неоднородных уравнений. Разрешая их, получаем выражения для относительной фазовой скорости, декремента затухания, скорости волнового движения, динамического давления и формы свободной поверхности с точностью до третьего приближения. Здесь приведены только выражения для частоты и декремента затухания (дисперсионные соотношения)

$$\begin{aligned}
\alpha = \alpha_0 \left\{ 1 + \varepsilon^2 A^2 \alpha_0 \frac{1 - e^{-2\nu_0 t / \alpha_0}}{864 \gamma_0^6 \nu_0 t} \right. & \left. [32 - 72 \gamma_0^2 + 108 \gamma_0^4 + 27 \gamma_0^6 + \right. \quad (5) \\
& \left. + 81 \gamma_0^8 + 12 \gamma_0^2 \nu_0^2 (27 \gamma_0^6 + 9 \gamma_0^4 - 12 \gamma_0^2 - 16) + 288 \gamma_0^4 \nu_0^4 (3 \gamma_0^2 + 4) + \right.
\end{aligned}$$

$$\begin{aligned}
& +8(3\gamma_0^2 - 2)(4 - 3\gamma_0^2)(9\gamma_0^4 - 6\gamma_0^2 + 2) \left((3\gamma_0^2 - 2)^2 + 24\gamma_0^2\nu_0^2 \right)^{-1} \Big\}, \\
\beta = \nu_0 & \left\{ 1 + \alpha_0 A^2 \frac{1 - e^{-2\nu_0 t/\alpha_0}}{864\gamma_0^6\nu_0 t} [81\gamma_0^8 + 135\gamma_0^6 - 36\gamma_0^4 - 24\gamma_0^2 - \right. \\
& - 32 + 12\gamma_0^2\nu_0^2(1 + 3\gamma_0^2)(16 + 12\gamma_0^2 - 9\gamma_0^4) - 288\gamma_0^4\nu_0^4(3\gamma_0^2 + 4) + \\
& \left. + 8(4 - 3\gamma_0^2)(3\gamma_0^2 + 2)(9\gamma_0^4 - 6\gamma_0^2 + 2) \left((3\gamma_0^2 - 2)^2 + 24\gamma_0^2\nu_0^2 \right)^{-1} \right\}.
\end{aligned}$$

Из найденных выражений для относительной частоты и декремента затухания волны следует, что свое максимальное конечное значение они принимают в начальный момент и с течением времени стремятся к значениям, соответствующим линейной задаче. Если коэффициент поверхностного натяжения σ положить равным нулю ($\kappa_0^2 = 0$, $\gamma_0^2 = 1$), то выражения (5) переходят в выражения, полученные в работе [3].

Графики зависимости фазовой скорости и декремента затухания волны длины 1 м и относительной высоты 0,1 ($\varepsilon = 0,1$), распространяющейся по поверхности воды ($\nu_0 = 5 \cdot 10^{-6}$), представлены на рисунках 1 и 2 соответственно.

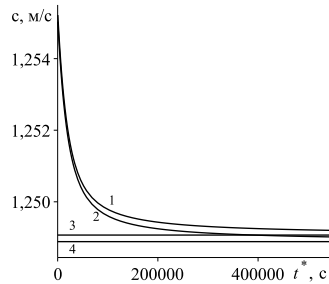


Рис. 1. Зависимость $c(t^*)$ для воды при $\lambda = 1$ м, $\varepsilon = 0,1$, $A=1$: линии 1, 2 – $c(t^*)$ при $\sigma = 73 \cdot 10^{-3}$ Н/м и $\sigma = 0$ соответственно, линии 3, 4 – фазовые скорости линейной задачи при $\sigma = 73 \cdot 10^{-3}$ Н/м и $\sigma = 0$ соответственно

Фазовая скорость и декремент принимают максимальные значения в начальный момент времени. Из графиков видно, что нелинейная фазовая скорость и декремент стремятся к своим линейным значениям. Фазовая скорость волны при $\sigma = 73 \cdot 10^{-3}$ Н/м имеет боль-

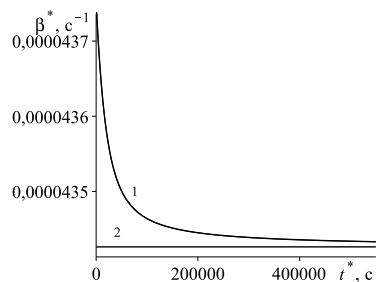


Рис. 2. Зависимость $\beta^*(t^*)$ для воды при $\lambda = 1$ м, $\varepsilon = 0,1$, $A=1$: линия 1 – $\beta^*(t^*)$ при $\sigma = 73 \cdot 10^{-3}$ Н/м и $\sigma = 0$, линия 2 – декремент затухания линейной задачи при $\sigma = 73 \cdot 10^{-3}$ Н/м и $\sigma = 0$

шие значения, чем фазовая скорость при $\sigma = 0$, и по мере затухания эта разница возрастает. Соответствующие графики для декремента затухания совпадают. Таким образом, капиллярно-гравитационная волна движется быстрее гравитационной, но при этом амплитуды их убывают с одинаковой скоростью.

Литература

1. Баринов В.А. Распространение волн по свободной поверхности вязкой жидкости // Вестник СПбГУ. Сер. 10, 2010. Вып. 2. С. 18–31.
2. Сретенский Л.Н. Теория волновых движений жидкости. М.: Наука, 1977. 816 с.
3. Баринов В.А., Басинский К.Ю. Решение нелинейной задачи о волнах на поверхности слабовязкой жидкости // Вестник СПбГУ. Сер. 10, 2011. Вып. 2.

Белова О.А.

Санкт-Петербургский государственный университет

Об одном семействе траекторий соударения

Рекомендовано к публикации профессором Шмыровым А.С.

Введение. В работе численным методом строятся траектории небесного тела, опасно сближающегося с Землей и в то же время проходящего через окрестность коллинеарной точки либрации L_1 .

Такие траектории являются интересными при разработке космических проектов по противодействию астероидно-кометной опасности, поскольку в окрестности коллинеарной точки либрации могут находиться крупные космические станции, которые могут быть использованы для изменения орбит астероида или кометы путем ударного воздействия.

Ограниченная задача трех тел. Одна из частных задач небесной механики – это задача о движении трёх тел, взаимно притягивающихся по закону тяготения Ньютона. Если притягивающиеся тела рассматривать как материальные точки (что выполняется, например, в первом приближении для Солнца, Земли и Луны или для Солнца, Юпитера и какого-либо из астероидов-троянцев), то для ряда случаев могут быть получены простые решения.

Например, в движении астероидов-троянцев реализуются так называемые треугольные решения Лагранжа для случая движения тела малой массы (астероида) в поле тяготения двух тел большой массы (Солнца и Юпитера). Астероид-троянец, находясь в так называемой точке либрации, движется по такой орбите, что Солнце, Юпитер и он сам находятся в трёх вершинах равностороннего треугольника. В общем случае устойчивые траектории трёх гравитационно взаимодействующих тел могут быть очень сложными.

Существует общее аналитическое решение ограниченной задачи трёх тел в виде рядов, сходящихся для любого момента времени. Однако из-за медленной сходимости этих рядов вместо аналитического метода пользуются численными методами решения этой задачи [1].

Точки либрации. Со многих точек зрения удобно изучать движение тела P в системе координат, вращающейся вместе с телами S и E , выбрав в качестве единицы длины расстояние между коллинеарной точкой либрации L_1 и телом E (см. рис. 1). В этой системе координат Лагранжевым решениям задачи трех тел соответствуют фиксированные точки – положения равновесия тела P . Точки, лежащие на прямой, проходящей через S и E , обозначают через L_1 , L_2 и L_3 , а точки, образующие равносторонние треугольники с телами S и E , обозначают через L_4 и L_5 . Эти точки в небесной механике называют точками либрации или либрационными центрами; L_1 , L_2 , L_3 – прямолинейные (коллинеарные) точки либрации, а L_4 и L_5 – треугольные. Если тело P поместить в одну из точек либрации с нулевой (во вращающейся системе координат) скоростью, то оно останется неподвижным [1].

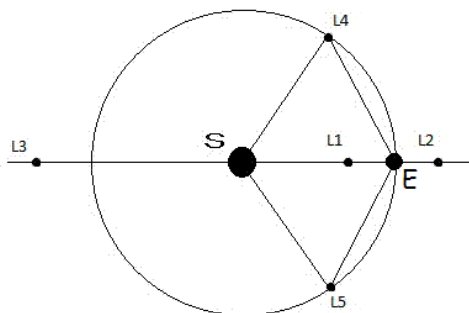


Рис. 1. Точки либрации

Уравнения движения и их характеристики. Уравнения движения астероида в околоземном пространстве во вращающейся системе координат $Ox_1x_2x_3$, предназначенные для описания орбитального движения небесных тел и полученные методом Хилла, имеют вид

$$\begin{cases} \dot{x}_1 = x_2 + y_1, \\ \dot{x}_2 = -x_1 + y_2, \\ \dot{x}_3 = y_3, \\ \dot{y}_1 = -\frac{3x_1}{\|x\|^3} + 2x_1 + y_2, \\ \dot{y}_2 = -\frac{3x_2}{\|x\|^3} - x_2 - y_1, \\ \dot{y}_3 = -\frac{3x_3}{\|x\|^3} - x_3, \end{cases} \quad (1)$$

где центр инерции Земли, совпадает с началом системы координат, ось x_1 соединяет центры масс Земли и Солнца. Здесь (x_1, x_2, x_3) – вектор координат астероида, (y_1, y_2, y_3) – вектор скоростей. Точка либрации L_1 во вращающейся системе координат неподвижна и имеет координаты $x = (1, 0, 0)$, $y = (0, 1, 0)$. За единицу расстояния принята величина равная 10^{-2} а.е. (примерно = $1,5 \times 10^6$ км), за единицу времени – 58,0916 суток [2].

Неустойчивость в L_1 дает возможность эффективного управления орбитальным движением искусственных небесных тел достаточно большой массы с помощью относительно небольшого управляющего воздействия. Этот факт можно использовать в проблеме противодействия астероидно-кометной опасности [3].

Построение траектории движения. Для построения траектории изучаемого класса начальная точка траектории помещается в коллинеарную точку либрации L_1 и подбирается нормальная компонента скорости, обеспечивающая проход небесного тела на достаточно близком (заданном) расстоянии от центра Земли.

На рис. 2–5 приведены различные варианты построения траекторий.

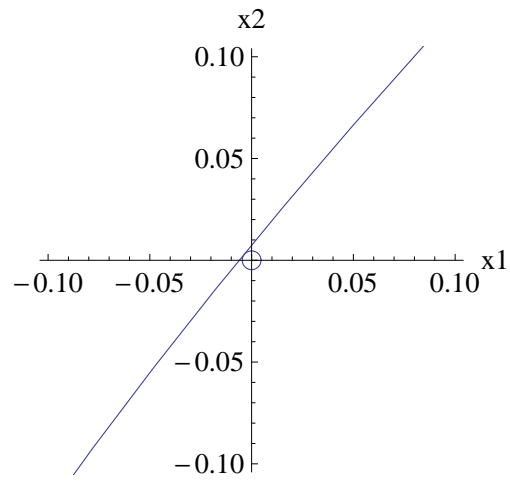


Рис. 2. Движение астероида с начальными данными $(x_0; y_0) = (1, 0, 0, -1.5991, 1, 0)$

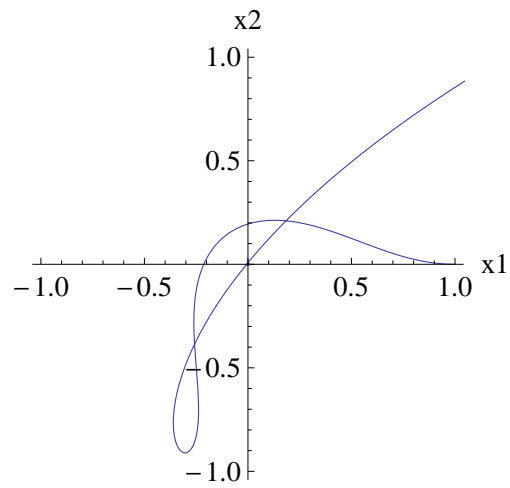


Рис. 3. Движение астероида с начальными данными $(x_0; y_0) = (1, 0, 0, -1.5991, 1, 0)$ (крупный масштаб)

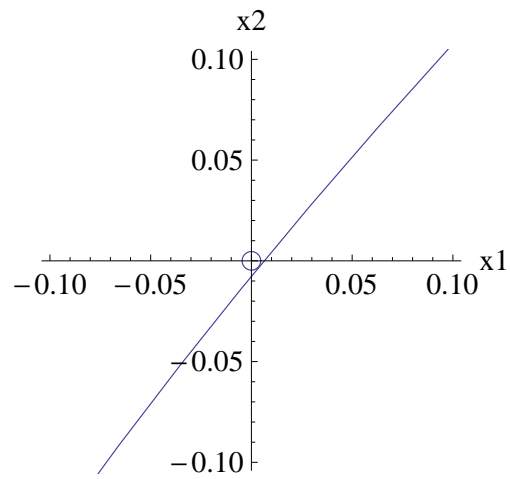


Рис. 4. Движение астероида с начальными данными $(x_0; y_0) = (1, 0, 0, -1.5955, 1, 0)$

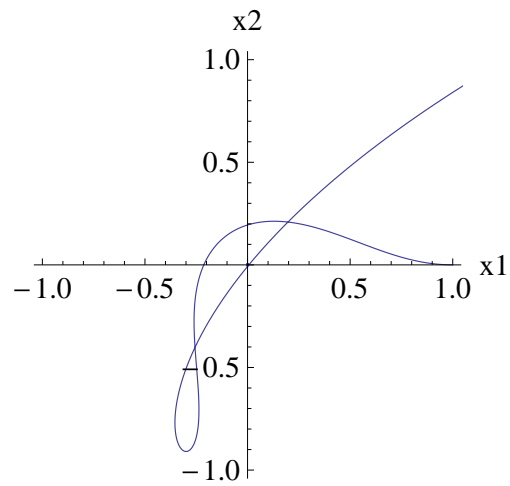


Рис. 5. Движение астероида с начальными данными $(x_0; y_0) = (1, 0, 0, -1.5955, 1, 0)$ (крупный масштаб)

Заключение. Проблема астероидной опасности непосредственно связана с существованием астероидов, которые представляют собой угрозу столкновения с поверхностью Земли, но для того чтобы говорить о реальной угрозе, необходимо, прежде всего, достаточно точно определить траекторию движения потенциально опасного небесного тела.

Таким образом, построенные в работе траектории могут дать представления о количестве астероидов и комет, пролетающих на близком расстоянии от Земли и проходящих через окрестность коллинеарной точки либрации L_1 . В то же время, проведенное исследование может являться дополнительным стимулирующим фактором для разработки проектов полета в эти области космического пространства, например, с целью исследования таких объектов или изменения их кинематических характеристик.

Литература

1. Маркеев А.П. Точки либрации в небесной механике и космодинамике. М.: Наука, 1978. 312 с.
2. Шмыров В.А. Стабилизация управляемого орбитального движения космического аппарата в окрестности коллинеарной точки либрации L_1 // Вестник СПбГУ. Сер. 10, 2005. Вып. 2. С. 193–199.
3. Шиманчук Д.В. Моделирование орбитального управляемого движения космического аппарата в окрестности коллинеарной точки либрации L_1 // Вестник СПбГУ. Сер. 10, 2010. Вып. 3. С. 86–92.

Большаков А.И.

Санкт-Петербургский государственный университет

Алгоритм выбора шага в методах плотной выдачи

Рекомендовано к публикации профессором Бабаджаняном Л.К.

Введение. Выбор длины шага важен в численных методах решения ОДУ [1–3]. Он позволяет контролировать локальную погрешность и избегать лишних вычислений. Предлагается новый способ, применимый к классу методов численного интегрирования ОДУ с механизмом плотной выдачи [1, 3, 4], основанный на экстраполяции погрешности. Рассматривается задача Коши

$$\dot{x} = f(x), \quad x(t_l) = x_l, \quad x \in R^n, \quad t \in R.$$

Как известно, приближение явным методом типа Рунге – Кутты $Rx(t, t_l, x_l)$ решения $x(t, t_l, x_l)$ этой задачи определяется формулами

$$k_1(t_l) = f(x_l), \quad k_i(t_l) = f\left(x_l + h \sum_{s=0}^{i-1} \beta_{is} k_s(t_l)\right), \quad i = 2, \dots, r;$$
$$Rx(t_l + \theta h, t_l, x_l) = x_l + h \sum_{i=1}^r \gamma_i(\theta) k_i(t_l),$$

где r – натуральное число, β_{is} – вещественные числа, а $\gamma_i(\theta)$ – вещественные функции, $\theta \in [0, 1]$. Величина шага h может быть как положительной, так и отрицательной. В первом случае приближение дается для $t > t_l$, во втором для $t < t_l$. Положим $x_{l+1} = Rx(t_{l+1}, t_l, x_l)$. Здесь x_{l+1} – приближенное значение решения, полученное методом Рунге – Кутты в точке t_{l+1} . Далее используется функция

$$\mathcal{E}(\tau; t_l, x_l; t_{l+1}, x_{l+1}) = \max_{i \in [1, n]} \frac{|(Rx(\tau, t_l, x_l) - Rx(\tau, t_{l+1}, x_{l+1}))_i|}{(|(x_l)_i| + |(x_{l+1})_i|)/2 + \delta},$$

при $\delta > 0$, $\tau \geq t_l$, которую назовем функцией погрешности.

Описание алгоритма. Он разбит на пять шагов. Предполагается, что известны приближенные значения решения в точках t_p , t_{p+1} , t_{p+2} и результатом работы алгоритма будет максимальная величина шага $h = t_{p+3} - t_{p+2}$, удовлетворяющая условиям на требуемую

точность. В результате первых трех шагов получается первое приближение h_I . Оставшиеся два шага предназначены для уточнения этой величины.

Шаг 1. Пусть известны величины

$$Rx(t_p, t_{p-1}, x_{p-1}), \quad Rx(t_{p+1}, t_p, x_p), \quad Rx(t_{p+2}, t_{p+1}, x_{p+1}).$$

Точками $\tau_1 = t_p, \tau_2, \dots, \tau_{M+2}$ и $\tau_{M+2}, \tau_{M+3}, \dots, \tau_{2M+3}$ разобьем промежутки $[t_p, t_{p+1}]$ и $[t_{p+1}, t_{p+2}]$ на $M+1$ частей каждый ($M \geq 0$). При плотной выдаче, позволяющей производить «дешевые вычисления» в промежуточных (между узловыми) точках τ_i , найдем величины

$$\begin{aligned} Rx(\tau_i, \tau_1, x_p), \quad Rx(\tau_i, \tau_{M+2}, x_{p+1}), \quad i \in [2, M+1]; \\ Rx(\tau_i, \tau_{M+2}, x_{p+1}), \quad Rx(\tau_i, \tau_{2M+3}, x_{p+2}), \quad i \in [M+3, 2M+2]. \end{aligned}$$

Шаг 2. По точкам τ_i и соответствующим значениям

$$\begin{aligned} \mathcal{E}(\tau_i; \tau_1, x_p; \tau_{M+2}, x_{p+1}), \quad i \in [1, M+2], \\ \mathcal{E}(\tau_i; \tau_{M+2}, x_{p+1}; \tau_{2M+3}, x_{p+2}), \quad i \in [M+3, 2M+3] \end{aligned}$$

методом наименьших квадратов приближаем функцию

$$E = \begin{cases} \mathcal{E}(\tau; \tau_1, x_p; \tau_{M+2}, x_{p+1}), & \tau \in [t_p, t_{p+1}], \\ \mathcal{E}(\tau; \tau_{M+2}, x_{p+1}; \tau_{2M+3}, x_{p+2}), & \tau \in [t_{p+1}, t_{p+2}] \end{cases}$$

параболой $a\tau^2 + b\tau + c$, $\tau \in [t_p, t_{p+2}]$.

Шаг 3. Пусть D – дискриминант уравнения $a\tau^2 + b\tau + c = \varepsilon$, ε – требуемая погрешность, а $t^* > t_{p+2}$ больший корень в случае $D > 0$. Тогда в качестве начального приближения h_I естественно положить

$$h_I = \begin{cases} t^* - t_{p+2}, & \text{если } D > 0, a > 0, t^* > t_{p+2}, \\ t_{p+2} - t_{p+1}, & \text{иначе.} \end{cases}$$

Шаг 4. Рассмотрим функцию

$$G(\tau) = \begin{cases} \mathcal{E}(\tau; t_p, x_p; t_{p+3}, x_{p+3}), & |\tau - t_p| \leq |\tau - t_{p+1}|, \\ \mathcal{E}(\tau; t_{p+1}, x_{p+1}; t_{p+3}, x_{p+3}), & \begin{cases} |\tau - t_p| > |\tau - t_{p+1}|, \\ |\tau - t_{p+1}| \leq |\tau - t_{p+2}|, \end{cases} \\ \mathcal{E}(\tau; t_{p+2}, x_{p+2}; t_{p+3}, x_{p+3}), & |\tau - t_{p+1}| > |\tau - t_{p+2}|, \end{cases}$$

т.е. $G(\tau)$ вычисляется как значение функции \mathcal{E} аргумента τ через значения в ближайшем узле (t_p, t_{p+1} или t_{p+2}) и в узле t_{p+3} .

Вычисляем $G(\tau_1), G(\tau_2), \dots$. Возможны два случая: а) существует τ_i такое, что $G(\tau_i) < \varepsilon$. Минимальное из таких τ_i обозначим τ_* и положим $h^0 = t_{p+3} - \tau_*$; б) такого τ_i не существует, тогда полагаем $h^0 = h_I$.

Шаг 5. Действуем по следующему стандартному алгоритму деления шага пополам:

1. $h := h^0$.
2. Если выполнено условие $\mathcal{E}(t_{p+2}; t_{p+2}, x_{p+2}; t_{p+2} + h, x_{p+3}) \leq \varepsilon$ или количество циклов превысило некоторое допустимое число N_{max} перейти к 4.
3. $h := h/2$, перейти к 2.
4. Выход.

Блок-схема. На рис. 1 графически представлен описанный выше алгоритм. Блоку «Максимизация шага LSQ» соответствует Шаг 4, а блоку «Вычисление шага методом FB (промах)» – Шаг 5. Стандартный алгоритм Шага 5 представлен на рис. 2. Он же используется для вычисления первых трех величин x_1, x_2, x_3 .



Рис. 1. Основной алгоритм

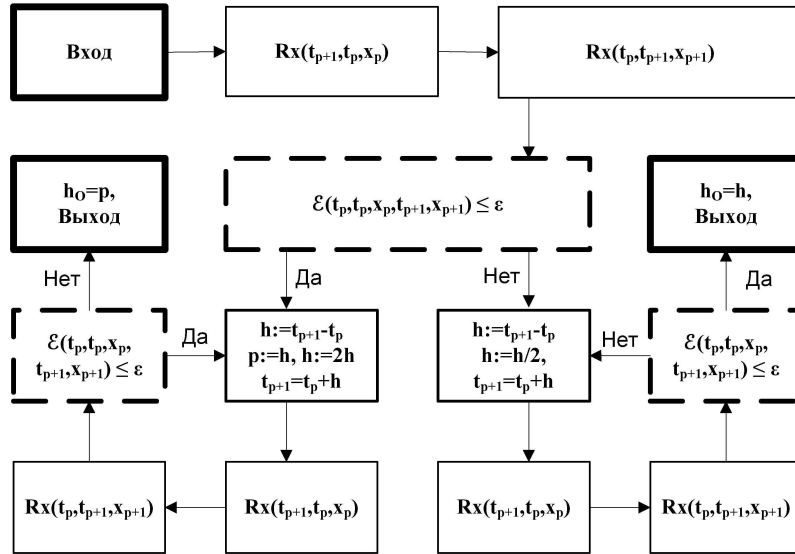


Рис. 2. Вычисление шага методом FB (промах)

Численный эксперимент. Для сравнения использовался метод Дорманда – Принса порядка 5(4) (DOPRI5 из [1]) и программа DP_LSQ, отличающаяся от DOPRI5 тем, что для выбора шага используется предложенный выше алгоритм, реализующая вышеописанный алгоритм для метода Дорманда – Принса с плотной выдчей порядка 4 [1, 5]. Вычисления проводились на промежутке $[0, \ln(\pi k/\alpha)]$, $k = 1, 2, \dots$ для системы уравнений

$$\begin{aligned} \dot{x}_1 &= \alpha x_2 x_3, & \dot{x}_2 &= -\alpha x_1 x_3, & \dot{x}_3 &= \alpha x_3, \\ x_1(0) &= \sin(1), & x_2(0) &= \cos(1), & x_3(0) &= 1, \end{aligned}$$

решением которой являются функции $x_1 = \sin e^{\alpha t}$, $x_2 = \cos e^{\alpha t}$, $x_3 = e^{\alpha t}$. Результаты приведены в таблице.

Таблица. Численные результаты

α , отрезок, погрешность, метод	Глобальная погрешность	Количество шагов	Относительное время счета
$\alpha = 2, t \in [0, \ln 10\pi/2], \varepsilon = 10^{-7}, \text{DOPRI5}$	$2 \cdot 10^{-6}$	176	1
$\alpha = 2, t \in [0, \ln 10\pi/2], \varepsilon = 10^{-7}, \text{DP_LSQ}$	$1.158 \cdot 10^{-6}$	173	0.95
$\alpha = 2, t \in [0, \ln 10\pi/2], \varepsilon = 10^{-10}, \text{DOPRI5}$	$2.4 \cdot 10^{-9}$	718	1
$\alpha = 2, t \in [0, \ln 10\pi/2], \varepsilon = 10^{-10}, \text{DP_LSQ}$	$3.3 \cdot 10^{-10}$	572	0.96
$\alpha = 2, t \in [0, \ln 10\pi/2], \varepsilon = 10^{-12}, \text{DOPRI5}$	$2.3 \cdot 10^{-11}$	1810	1
$\alpha = 2, t \in [0, \ln 10\pi/2], \varepsilon = 10^{-12}, \text{DP_LSQ}$	$2.8 \cdot 10^{-11}$	1243	0.87
$\alpha = 5, t \in [0, \ln 20\pi/5], \varepsilon = 10^{-10}, \text{DOPRI5}$	$7.6 \cdot 10^{-9}$	1460	1
$\alpha = 5, t \in [0, \ln 20\pi/5], \varepsilon = 10^{-10}, \text{DP_LSQ}$	$4.7 \cdot 10^{-10}$	1184	0.73
$\alpha = 5, t \in [0, \ln 200\pi/15], \varepsilon = 10^{-12}, \text{DOPRI5}$	$8 \cdot 10^{-10}$	37381	1
$\alpha = 5, t \in [0, \ln 200\pi/15], \varepsilon = 10^{-12}, \text{DP_LSQ}$	$2.2 \cdot 10^{-13}$	25972	0.81

Анализируя полученные результаты, можно прийти к выводу, что предлагаемый алгоритм работает и дает выигрыш во времени счета при сохранении порядка глобальной погрешности.

Литература

1. Хайрер Э., Нёрсетт С., Ваннер Г. Решение обыкновенных дифференциальных уравнений. Нежесткие задачи. М.: Мир, 1990. 512 с.
2. Babadzhanjanz L.K., Sarkissian D.R. Taylor series method for dynamic systems with control: Convergence and error estimates // Journal of Mathematical Sciences, 2004. Vol. 139. P. 7025–7046.
3. Бабаджанянц Л.К. Метод рядов Тейлора // Вестник СПбГУ. Сер. 10, 2010. Вып. 3. С. 13–29.
4. Shampine L.F. Some practical Runge-Kutta formulas // Mathematics of Computation, 1986. Vol. 46, № 173. P. 135–150.

Брэгман К.М.

Санкт-Петербургский государственный университет

Сведение дифференциальных уравнений к полиномиальной системе

Рекомендовано к публикации доцентом Пурьшиевой Ю.Ю.

Предисловие. Предлагается основанный на методе дополнительных переменных (МДП) алгоритм сведения к полиномиальной форме систем дифференциальных уравнений первого порядка, правые части которых можно записать при помощи четырех действий алгебры и любых допустимых суперпозиций функций конечного числа аргументов, принадлежащих пополняемому набору функций, называемому библиотекой. Алгоритм разработан на основе статьи [1] с учетом возможностей пакета «Mathematica» [2], что позволяет получать конечные результаты в аналитической форме. Алгоритм и пример его применения описываются во втором разделе, а в первом разделе рассматриваются необходимые предварительные сведения.

1. Введение в метод дополнительных переменных. Здесь кратко излагаются необходимые понятия и результаты из статьи [1].

1.1. Дифференциальные системы. Рассмотрим уравнения

$$\partial x_i / \partial t_j = f_i^j(x, \alpha), \quad i \in [1, m], j \in [1, s] \quad (1)$$

при $t = (t_1, \dots, t_s) \in R^s$, $x = (x_1, \dots, x_m) \in R^m$, $\alpha = (\alpha_1, \dots, \alpha_\omega) \in R^\omega$, $f_i^j \in R$, которые, в случае $s = 1$, запишутся как система ОДУ. Полиномиальной будем называть систему (1), в которой все функции f_i^j – полиномы по x_1, \dots, x_m . Говорят, что скалярная функция φ аргумента $x = (x_1, \dots, x_\sigma)$ удовлетворяет полиномиальной системе, если она является одной из компонент решения некоторой полиномиальной системы. Класс скалярных функций аргумента $x = (x_1, \dots, x_\sigma)$, удовлетворяющих полиномиальной системе, обозначают Σ_σ . Так как любую функцию аргумента $x = (x_1, \dots, x_{\sigma_1})$, принадлежащую классу Σ_{σ_1} , можно считать также и функцией аргумента $x = (x_1, \dots, x_{\sigma_2})$ класса Σ_{σ_2} при $\sigma_1 < \sigma_2$, то можно принять, что $\Sigma_1 \subset \Sigma_2 \subset \dots$. Большое число функций из справочников и компьютерных систем принадлежит Σ_1 (и тем более Σ_σ при $\sigma > 1$).

1.2. Метод дополнительных переменных (МДП). Этот метод сводит уравнения (1) к полиномиальной системе. Он состоит

в том, что находят функции $x_{m+1}(x_1, \dots, x_m), \dots, x_{m+k}(x_1, \dots, x_m)$ такие, что:

- правые части в (1) – полиномы по $x_1, \dots, x_m, x_{m+1}, \dots, x_{m+k}$;
- все производные $dx_{m+l}/dt_j = \sum_{i=1}^m f_i^j(x, \alpha) \partial x_{m+l} / \partial x_i$ дополнительных переменных x_{m+1}, \dots, x_{m+k} в силу уравнений (1) – полиномы по $x_1, \dots, x_m, x_{m+1}, \dots, x_{m+k}$.

Тогда x_1, \dots, x_{m+k} удовлетворяют полиномиальной системе.

Пример. Уравнение $\ddot{x} + k^2 \sin x = 0$ записываем как систему $\dot{x}_1 = x_2, \dot{x}_2 = -k^2 \sin x_1$, и, вводя переменные $x_3 = \sin x_1, x_4 = \cos x_1$, получаем $\dot{x}_1 = x_2, \dot{x}_2 = -k^2 x_3, \dot{x}_3 = x_2 x_4, \dot{x}_4 = -x_2 x_3$.

1.3. Классы функций и уравнений. Классом \mathcal{F}_m^σ , (σ, m – натуральные) называют множество скалярных функций аргумента $x = (x_1, \dots, x_m)$, которые можно получить из x_1, \dots, x_m при помощи конечного числа операций $+, -, \times, /$ и конечного числа функций $\varphi_1, \dots, \varphi_l \in \Sigma_\sigma$ и их суперпозиций. Если в число этих функций включить φ такую, что $\varphi(a) = 1/a$, то в определении \mathcal{F}_m^σ можно не использовать операцию деления, так как величина a/b равна $a \times \varphi(b)$. Символом $\mathcal{E}^s(\mathcal{F}_m^\sigma)$ называют класс уравнений (1), все правые части которых принадлежат \mathcal{F}_m^σ . К любой системе класса $\mathcal{E}^s(\mathcal{F}_m^\sigma)$ применим МДП. Предлагаемый далее в п. 2.1 для систем класса $\mathcal{E}^s(\mathcal{F}_m^\sigma)$ алгоритм МДП основан на понятии библиотеки.

1.4. Расширения и библиотеки. Если функции $\varphi_{1,1}, \dots, \varphi_{l,1}$ от $x = (x_1, \dots, x_\sigma)$ принадлежат Σ_σ , то каждая из них является компонентой решения полиномиальной системы. Если решение полиномиальной системы для $\varphi_{i,1}$ состоит из компонент $\varphi_{i,1}, \dots, \varphi_{i,n(i)}$, то множество $\mathcal{E}xt\{\varphi_{1,1}, \dots, \varphi_{l,1}\} = \cup_{i=1}^l \{\varphi_{i,1}, \dots, \varphi_{i,n(i)}\}$ называют расширением множества $\{\varphi_{1,1}, \dots, \varphi_{l,1}\}$. Так как одна и та же функция удовлетворяет различным полиномиальным системам, то расширение для $\{\varphi_{1,1}, \dots, \varphi_{l,1}\}$ не единственно.

Символом $D(\varphi)$, обозначим полиномиальную систему

$$\partial \varphi_i / \partial x_j = \Phi_i^j(\varphi, \alpha), \quad \varphi = (\varphi_1, \dots, \varphi_l), \quad i \in [1, l], \quad j \in [1, \sigma]$$

относительно компонент вектор-функции $\varphi = (\varphi_1, \dots, \varphi_l)$. Пару $L(\varphi) = (\varphi, D(\varphi))$ называют библиотекой функций $\varphi_1, \dots, \varphi_l$ класса Σ_σ , заданной уравнениями $D(\varphi)$, или просто библиотекой. В частности, библиотеку можно задать, добавив к ее уравнениям $D(\varphi)$ начальные условия. Если задана некоторая конкретная библиотека, а правые части системы (1) выражаются при помощи четырех

действий алгебры и суперпозиций некоторых из функций $\varphi_1, \dots, \varphi_l$ этой библиотеки, то, как мы увидим в п. 2., все необходимые дополнительные переменные метода МДП определяются последовательно как суперпозиции этих функций и полиномов по исходным неизвестным и введенным уже дополнительным переменным. Библиотеку можно задавать как базу данных, содержащую набор полиномиальных систем, а не задачи Коши для них, а начальные данные для дополнительных переменных можно определять по начальным данным исходной задачи Коши.

2. Алгоритм сведения к полиномиальной системе. Здесь будут рассмотрены алгоритм и пример.

2.1. Алгоритм сведения. Вначале опишем элементарное преобразование уравнений вида (1), составляющее основу алгоритма, а затем – схему алгоритма.

Элементарное преобразование системы класса $\mathcal{E}^s(\mathcal{F}_m^\sigma)$.

Рассмотрим полную систему вида (1) (при $s = 1$ это система ОДУ).

1. Ищем справа в (1) функцию вида $\varphi^k(P_1(x), \dots, P_\vartheta(x))$, где k и ϑ – натуральные, $\varphi = \varphi(p_1, \dots, p_\vartheta)$ – одна из функций библиотеки, а P_ν – полиномы по x_1, \dots, x_m .

2. Пусть $\varphi = \varphi_1, \dots, \varphi_\mu$ – все функции расширения функции φ и

$$\partial\varphi_i/\partial p_\nu = Q_i^\nu(\varphi_1, \dots, \varphi_\mu), \quad i \in [1, \mu], \nu \in [1, \vartheta],$$

где Q_i^ν – полиномы. Вводим новые (дополнительные) переменные

$$x_{m+1} = \varphi_1(P_1(x), \dots, P_\vartheta(x)), \dots, x_{m+\mu} = \varphi_\mu(P_1(x), \dots, P_\vartheta(x)).$$

3. Заменяя $\varphi_1(P_1, \dots, P_\vartheta), \dots, \varphi_\mu(P_1, \dots, P_\vartheta)$ в (1) на $x_{m+1}, \dots, x_{m+\mu}$ соответственно, получаем

$$\partial x_i/\partial t_j = g_i^j(x_1, \dots, x_m, x_{m+1}, \dots, x_{m+\mu}), \quad i \in [1, m], j \in [1, s].$$

4. Выписываем уравнения для введенных переменных

$$\partial x_{m+i}/\partial t_j = \sum_{\nu=1}^{\vartheta} Q_i^\nu(x_{m+1}, \dots, x_{m+\mu}) \times$$

$$\times \sum_{k=1}^m \partial P_\nu / \partial x_k \cdot g_k^j(x_1, \dots, x_m, x_{m+1}, \dots, x_{m+\mu}), \quad i \in [1, \mu], j \in [1, s].$$

Схема сведения. Шаг 1. Заносим систему и начальные условия в пакете «Mathematica» его средствами:

$$\partial x_i / \partial t_j = f_i^j(x_1, \dots, x_n), \quad x_i(t_0) = x_{i,0}, \quad i \in [1, n], j \in [1, s]. \quad (2)$$

Образует подбиблиотеку, состоящую из объединения расширений всех функций, участвующих в написании f_i^j в (2).

Шаг 2. Преобразуем систему шаг за шагом, пока не получим полиномиальную систему, применяя каждый раз элементарное преобразование к полученной системе. На первом шаге преобразование применяется к системе (2). В итоге для x_1, \dots, x_{n+M} получим полиномиальную систему

$$\partial x_i / \partial t_j = R_i^j(x_1, \dots, x_{n+M}), \quad i \in [1, n+M], j \in [1, s].$$

Шаг 3. Вычисляем начальные значения $x_{n+1,0}, \dots, x_{n+M,0}$.

2.2. Пример сведения. Рассматривается система ОДУ. Этот модельный, но не тривиальный пример иллюстрирует алгоритм сведения и предназначен для отладки соответствующей программы. Вначале естественно рассмотреть используемую библиотеку.

Библиотека. Она является объединением подбиблиотек, которые рассмотрим вместе с соответствующими дифференциальными уравнениями. Функции обозначаем $\varphi, \varphi_1, \dots$, а их аргументы p, p_1, \dots

Пб1. Состоит из функции одного аргумента $\varphi = \text{inv}(p) = p^{-1}$, которая удовлетворяет уравнению $d\varphi/dp = -\varphi^2$, а ее расширение состоит из нее одной.

Пб2. Состоит из двух функций одного аргумента: $\varphi_1 = \ln p$, $\varphi_2 = \text{inv}(p)$, удовлетворяющих системе $d\varphi_1/dp = \varphi_2$, $d\varphi_2/dp = -\varphi_2^2$, причем расширение функции φ_1 состоит из функций φ_1, φ_2 , а расширение функции φ_2 состоит из нее одной, т.е. совпадает с Пб1.

Пб3. Состоит из двух функций одного аргумента: $\varphi_1 = \sin p$, $\varphi_2 = \cos p$, удовлетворяющих системе $d\varphi_1/dp = \varphi_2$, $d\varphi_2/dp = -\varphi_1$, а расширение каждой из функций φ_1, φ_2 состоит из функций φ_1, φ_2 .

Пб4. Состоит из двух функций одного аргумента: $\varphi_1 = \text{ch}(p)$, $\varphi_2 = \text{sh}(p)$, удовлетворяющих системе $d\varphi_1/dp = \varphi_2$, $d\varphi_2/dp = \varphi_1$, а расширение каждой из функций φ_1, φ_2 состоит из функций φ_1, φ_2 .

Перейдем к самому примеру. **Шаг 1.** Заносим систему и начальные условия в пакете «Mathematica» его средствами:

$$dx_1/dt = x_3^3(\sin \cos(a \ln^2 x_2 + bx_3) + b \ln^4 x_2) + \sin^5(a \ln^2 x_2 + bx_3),$$

$$\begin{aligned}
dx_2/dt &= x_2^2 \cos \sin(a \ln^2 x_2 + bx_3) + \cos^4(a \ln^2 x_2 + bx_3), \\
dx_3/dt &= x_1(\operatorname{ch}^2(a \ln^2 x_2 + bx_3) + \sin(a \ln^2 x_2 + bx_3)) + \operatorname{sh}^5(a \ln^2 x_2 + bx_3), \\
x_i(t_0) &= x_{i,0}, \quad i = 1, 2, 3,
\end{aligned}$$

где $\alpha = (\alpha_1, \alpha_2) = (a, b) \in R$ – параметры.

Подбиблиотека состоит из объединения Пб1, ..., Пб4, т.е. из функций (записываем их как функции от p): $\ln p$, p^{-1} , $\sin p$, $\cos p$, $\operatorname{sh} p$, $\operatorname{ch} p$.

Шаг 2. Преобразуем уравнения шаг за шагом (шаг 2.1, шаг 2.2, ...), пока не получим полиномиальную систему, применяя на каждом шаге элементарное преобразование.

Шаг 2.1.

1. В качестве первой функции $\varphi(P)$ возьмем $\ln x_2$.
2. Функции расширения $\ln p$: $\varphi_1 = \ln p$, $\varphi_2 = p^{-1}$. Они удовлетворяют системе $d\varphi_1/dp = \varphi_2$, $d\varphi_2/dp = -\varphi_2^2$. Вводим дополнительные переменные $x_4 = \ln x_2$, $x_5 = x_2^{-1}$.
3. Заменяя $\ln x_2$, x_2^{-1} во всех их вхождениях в уравнения на x_4 , x_5 соответственно, получаем новую запись этих уравнений

$$\begin{aligned}
dx_1/dt &= x_3^3(\sin \cos(ax_4^2 + bx_3) + bx_4^4) + \sin^5(ax_4^2 + bx_3), \\
dx_2/dt &= x_2^2 \cos \sin(ax_4^2 + bx_3) + \cos^4(ax_4^2 + bx_3), \\
dx_3/dt &= x_1(\operatorname{ch}^2(ax_4^2 + bx_3) + \sin(ax_4^2 + bx_3)) + \operatorname{sh}^5(ax_4^2 + bx_3).
\end{aligned}$$

4. Уравнения для дополнительных переменных следующие

$$\begin{aligned}
dx_4/dt &= x_5(x_2^2 \cos \sin(ax_4^2 + bx_3) + \cos^4(ax_4^2 + bx_3)), \\
dx_5/dt &= -x_5^2(x_2^2 \cos \sin(ax_4^2 + bx_3) + \cos^4(ax_4^2 + bx_3)).
\end{aligned}$$

Итак, в результате шага 2.1 получена система

$$\begin{aligned}
dx_1/dt &= x_3^3(\sin \cos(ax_4^2 + bx_3) + bx_4^4) + \sin^5(ax_4^2 + bx_3), \\
dx_2/dt &= x_2^2 \cos \sin(ax_4^2 + bx_3) + \cos^4(ax_4^2 + bx_3), \\
dx_3/dt &= x_1(\operatorname{ch}^2(ax_4^2 + bx_3) + \sin(ax_4^2 + bx_3)) + \operatorname{sh}^5(ax_4^2 + bx_3), \\
dx_4/dt &= x_5(x_2^2 \cos \sin(ax_4^2 + bx_3) + \cos^4(ax_4^2 + bx_3)), \\
dx_5/dt &= -x_5^2(x_2^2 \cos \sin(ax_4^2 + bx_3) + \cos^4(ax_4^2 + bx_3)).
\end{aligned}$$

Шаг 2.2 – Шаг 2.5. Действуя аналогично, вводим переменные $x_6 = \sin(ax_4^2 + bx_3)$, $x_7 = \cos(ax_4^2 + bx_3)$, $x_8 = \operatorname{sh}(ax_4^2 + bx_3)$, $x_9 = \operatorname{ch}(ax_4^2 + bx_3)$, $x_{10} = \cos x_6$, $x_{11} = \sin x_6$, $x_{12} = \sin x_7$, $x_{13} = \cos x_7$. Окончательно, после шага 2.5 получаем полиномиальную систему

$$\begin{aligned} dx_1/dt &= x_3^3(x_{12} + bx_4^4) + x_6^5, \\ dx_2/dt &= x_2^2x_{10} + x_7^4, dx_3/dt = x_1(x_9^2 + x_6) + x_8^5, \\ dx_4/dt &= x_5(x_2^2x_{10} + x_7^4), dx_5/dt = -x_5^2(x_2^2x_{10} + x_7^4), \\ dx_6/dt &= x_7(2ax_4x_5(x_2^2x_{10} + x_7^4) + b(x_1(x_9^2 + x_6) + x_8^5)), \\ dx_7/dt &= -x_6(2ax_4x_5(x_2^2x_{10} + x_7^4) + b(x_1(x_9^2 + x_6) + x_8^5)), \\ dx_8/dt &= x_9(2ax_4x_5(x_2^2x_{10} + x_7^4) + b(x_1(x_9^2 + x_6) + x_8^5)), \\ dx_9/dt &= x_8(2ax_4x_5(x_2^2x_{10} + x_7^4) + b(x_1(x_9^2 + x_6) + x_8^5)), \\ dx_{10}/dt &= -x_{11}x_7(2ax_4x_5(x_2^2x_{10} + x_7^4) + b(x_1(x_9^2 + x_6) + x_8^5)), \\ dx_{11}/dt &= x_{10}x_7(2ax_4x_5(x_2^2x_{10} + x_7^4) + b(x_1(x_9^2 + x_6) + x_8^5)), \\ dx_{12}/dt &= -x_{13}x_6(2ax_4x_5(x_2^2x_{10} + x_7^4) + b(x_1(x_9^2 + x_6) + x_8^5)), \\ dx_{13}/dt &= x_{12}x_6(2ax_4x_5(x_2^2x_{10} + x_7^4) + b(x_1(x_9^2 + x_6) + x_8^5)). \end{aligned}$$

Исходные переменные и данные: $x_1, x_2, x_3; x_{1,0}, x_{2,0}, x_{3,0}$.

Дополнительные переменные и данные: $x_4 = \ln x_2$, $x_5 = x_2^{-1}$,
 $x_6 = \sin(ax_4^2 + bx_3)$, $x_7 = \cos(ax_4^2 + bx_3)$, $x_8 = \operatorname{sh}(ax_4^2 + bx_3)$,
 $x_9 = \operatorname{ch}(ax_4^2 + bx_3)$, $x_{10} = \cos x_6$, $x_{11} = \sin x_6$, $x_{12} = \sin x_7$,
 $x_{13} = \cos x_7$; $x_{4,0} = \ln x_{2,0}$, $x_{5,0} = x_{2,0}^{-1}$, $x_{6,0} = \sin(ax_{4,0}^2 + bx_{3,0})$,
 $x_{7,0} = \cos(ax_{4,0}^2 + bx_{3,0})$, $x_{8,0} = \operatorname{sh}(ax_{4,0}^2 + bx_{3,0})$, $x_{10,0} = \cos x_{6,0}$,
 $x_{9,0} = \operatorname{ch}(ax_{4,0}^2 + bx_{3,0})$, $x_{11,0} = \sin x_{6,0}$, $x_{12,0} = \sin x_{7,0}$, $x_{13,0} = \cos x_{7,0}$.

Литература

1. Бабаджанянц Л.К. Метод дополнительных переменных // Вестник СПбГУ. Сер.10, 2010. Вып. 1. С. 3–11.
2. Wolfram Mathematica 8 Documentation Center
<http://reference.wolfram.com/mathematica/guide/Mathematica.html>

Викулина Ю.И., Греков М.А.
Санкт-Петербургский государственный университет

Напряженно-деформированное состояние полуплоскости при учете поверхностного напряжения¹

1. Введение. Каждое материальное тело имеет границу, отделяющую его от других тел. В классических моделях механики сплошных сред эту границу называют поверхностью, обычно представляя ее как математическое понятие – двумерное многообразие.

Известно [1], что состояние приповерхностных слоев реальных тел существенно отличается от состояния внутренних слоев вследствие различных условий взаимодействия материальных частиц, а также в силу ряда других причин. Для наноразмерных структур наличие приповерхностных эффектов может быть причиной существенного отличия их физических свойств от свойств обычных материалов.

Как правило, влияние поверхностного напряжения на состояние идеально упругого тела не учитывается на макроуровне, так как оно практически незначительно в сравнении с влиянием других нагрузок. Моделью, учитывающей свойства поверхности в нанобъектах, служит теория упругости с поверхностными напряжениями, получившая бурное развитие в последние годы [2], [3]. Цель данной работы – исследовать эффект воздействия поверхностного напряжения на напряженное состояние границы упругой полуплоскости при периодическом нагружении.

2. Постановка задачи. Рассмотрим упругую среду, занимающую полупространство. Предполагаем, что среда находится в условиях плоской деформации под действием внешних сил. Таким образом, приходим к формулировке краевой задачи о деформации полуплоскости $\Omega = \{z : \text{Im } z < 0, \text{Re } z \in \mathbb{R}^1\}$, $z = x_1 + ix_2$, с границей Γ . Считаем, что упругие свойства границы отличны от упругих свойств полуплоскости.

Согласно закону Лапласа – Юнга [2] граничные условия имеют

¹Работа выполнена при финансовой поддержке РФФИ, проект № 11-01-00230, и СПбГУ, проекты № 9.0.165.2009, 9.37.129.2011

вид

$$\sigma_{22} - i\sigma_{12} = p(x_1) - i\frac{\partial\tau_{11}}{\partial x_1}, \quad z \in \Gamma, \quad (1)$$

где $p(x_1) = p_2(x_1) - ip_1(x_1)$ – известная комплекснозначная функция, удовлетворяющая условию Гельдера почти всюду при $x_1 \in \mathbb{R}^1$ и

$$\int_{-\lambda/2}^{\lambda/2} p(x_1)dx_1 = 0, \quad p(x_1 + \lambda) = p(x_1),$$

$\tau_{11}(x_1)$ – неизвестное поверхностное напряжение.

Считаем, что на бесконечности

$$\lim_{\text{Im } z \rightarrow -\infty} \sigma_{ij}(z) = \lim_{\text{Im } z \rightarrow -\infty} \omega(z) = 0, \quad (2)$$

где ω – угол поворота материальной частицы, σ_{ij} – компоненты напряжений в декартовой прямоугольной системе координат x_1, x_2 .

Допуская, что упругие свойства поверхности отличаются от аналогичных свойств основного материала, имеем следующее определяющее соотношение [4], связывающее в случае плоской деформации поверхностное напряжение τ_{11} с соответствующей деформацией ε_{11}^s :

$$\tau_{11} = (2\mu_s + \lambda_s)\varepsilon_{11}^s, \quad z = x_1, \quad (3)$$

где λ_s, μ_s – модули поверхностной упругости, аналогичные постоянным Ламе λ, μ для объемной изотропной упругости.

Из условия непрерывности перемещений $u_j, j = 1, 2$, при переходе от объема Ω к границе Γ

$$\lim_{z \rightarrow x_1} u_j(z) = u_j^s(x_1), \quad (4)$$

где u_j^s – перемещение точек границы Γ вдоль оси x_j , вытекает условие непрерывности деформации ε_{11}

$$\lim_{z \rightarrow x_1} \varepsilon_{11} = \varepsilon_{11}^s. \quad (5)$$

Соотношения (3)–(5) приводят к уравнению для определения поверхностного напряжения

$$\tau_{11} = (2\mu_s + \lambda_s)\varepsilon_{11}, \quad z = x_1. \quad (6)$$

Выражение для деформации ε_{11} найдем из решения краевой задачи (1), (2). Согласно [5], связь вектора напряжений $\sigma_n = \sigma_{nn} + i\sigma_{nt}$ на площадке с нормалью \mathbf{n} и вектора перемещений $u = u_1 + iu_2$ с комплексным потенциалом Гурса – Колосова – Мухелишвили Φ определяется соотношениями:

$$\sigma_n(z) = \Phi(z) + \overline{\Phi(z)} - \left(\Phi(\bar{z}) + \overline{\Phi(z)} - (z - \bar{z}) \overline{\Phi'(z)} \right) e^{-2i\alpha}, \quad (7)$$

$$-2\mu \frac{du}{dz} = -\varkappa \Phi(z) + \overline{\Phi(z)} - \left(\Phi(\bar{z}) + \overline{\Phi(z)} - (z - \bar{z}) \overline{\Phi'(z)} \right) e^{-2i\alpha}, \quad (8)$$

где α – угол наклона площадки к оси x_1 , $\varkappa = 3 - 4\nu$, ν – коэффициент Пуассона. Производная в (8) берется вдоль площадки, т.е. в направлении вектора \mathbf{t} , перпендикулярного орту \mathbf{n} и определяющего с ним правую систему координат n, t .

Решение краевой задачи (1), (2) для комплексного потенциала Φ имеет вид [5]

$$\Phi(z) = -\frac{1}{2\pi i} \int_{-\infty}^{\infty} \frac{p(t)}{t-z} dt + \frac{1}{2\pi i} \int_{-\infty}^{\infty} \frac{it^s(t)}{t-z} dt = I_p(z) + I_\tau(z), \quad (9)$$

где $t^s = d\tau_{11}/dx_1$.

Полагая $\alpha = \pi/2$ в (7) и $\alpha = 0$ в (8), на границе Γ получим

$$\sigma_{11}(x_1) + i\sigma_{12}(x_1) = \Phi^-(x_1) + 2\overline{\Phi^-(x_1)} + \Phi^+(x_1), \quad (10)$$

$$-2\mu \frac{du}{dx_1} = -\varkappa \Phi^-(x_1) - \Phi^+(x_1). \quad (11)$$

В равенствах (10), (11) $\Phi^\pm(x_1) = \lim_{x_2 \rightarrow \pm 0} \Phi(z)$. На основании соотношений (6) и (11) имеет место

$$\tau_{11} = (2\mu_s + \lambda_s) \frac{du_1}{dx_1} = D \operatorname{Re}(\varkappa \Phi^- + \Phi^+), \quad (12)$$

где $D = (2\mu_s + \lambda_s)/(2\mu)$.

После дифференцирования (12) приходим к уравнению относительно неизвестной функции t^s :

$$t^s(x_1) = D \operatorname{Re}(\varkappa \Phi'^-(x_1) + \Phi'^+(x_1)). \quad (13)$$

Уравнение (13) при использовании формул Сохоцкого – Племеля для интегралов в (9) принимает вид

$$\begin{aligned}
& t^s(x_1) - \frac{D(\varkappa + 1)}{2\pi} \int_{-\infty}^{+\infty} \frac{t^s(t)}{(t - x_1)^2} dt = \\
& = \frac{D(\varkappa - 1)}{2} p_2'(x_1) - \frac{D(\varkappa + 1)}{2\pi} \int_{-\infty}^{+\infty} \frac{p_1(t)}{(t - x_1)^2} dt.
\end{aligned} \tag{14}$$

Следует отметить, что однородное уравнение, соответствующее гиперсингулярному интегральному уравнению (14), имеет только нулевое решение. В противном случае, существовало бы отличное от константы поверхностное напряжение τ_{11} при отсутствии внешних усилий (при $p = 0$ и нулевых напряжениях на бесконечности), что для плоской поверхности невозможно. Если производная функции t^s удовлетворяет условию Гельдера, то уравнение (14) всегда имеет единственное решение при любой непрерывной правой части [6].

3. Численные результаты при синусоидальной нагрузке.

Пусть на границе Γ действует касательное усилие $p(x_1) = -i\sigma_0 \sin bx_1$, где $b = 2\pi/\lambda$. Функцию t^s ищем в виде ряда Фурье:

$$t^s(x_1) = \sum_{k=1}^{\infty} A_k \sin bx_1 + B_k \cos bx_1.$$

Для простоты вычислений сначала найдем выражения интегралов I_p и I_τ , а затем, воспользовавшись формулами Сохоцкого – Племяля и свойствами интегралов типа Коши [7], получим предельные значения функции Φ при стремлении к границе Γ . Итак

$$I_p(z) = \frac{\sigma_0}{2} \begin{cases} e^{ibz}, & \text{Im } z > 0, \\ e^{-ibz}, & \text{Im } z < 0, \end{cases} \tag{15}$$

$$I_\tau(z) = \frac{1}{2} \begin{cases} \sum_{k=1}^{\infty} C_k e^{ibkz}, & \text{Im } z > 0, \\ \sum_{k=1}^{\infty} \overline{C_k} e^{-ibkz}, & \text{Im } z < 0, \end{cases} \tag{16}$$

где $C_k = A_k + iB_k$.

Продифференцировав выражения (15), (16) и подставив результат в (13) с учетом (9), находим

$$t^s(x_1) = -\frac{2\pi\sigma_0 D(\varkappa + 1)}{2\lambda + 2\pi D(\varkappa + 1)} \sin bx_1. \tag{17}$$

Из (10) получаем выражения для кольцевых σ_{11} и касательных σ_{12} напряжений на Γ :

$$\sigma_{11} = \frac{2\sigma_0\lambda}{\lambda + \pi D(\varkappa + 1)} \cos bx_1, \quad \sigma_{12} = \left(\sigma_0 - \frac{\sigma_0 2\pi D(\varkappa + 1)}{2\lambda + 2\pi D(\varkappa + 1)} \right) \sin bx_1.$$

Интегрируя (17), вычислим поверхностные напряжения

$$\tau_{11}(x_1) = \int t^s(t) dt + C = \frac{\sigma_0 D \lambda (\varkappa + 1)}{2\lambda + 2\pi D(\varkappa + 1)} \cos bx_1 + C.$$

Константа C равна остаточному поверхностному напряжению, действующему на границе при отсутствии внешних сил.

При действии нормальных напряжений $p(z) = \sigma_0 \cos bx_1$ на границе Γ получаем решение, аналогичное предыдущему случаю:

$$t^s = -\frac{\pi\sigma_0 D(\varkappa - 1)}{\lambda + \pi D(\varkappa + 1)} \sin bx_1, \quad \tau_{11} = \frac{\sigma_0 D \lambda (1 - \varkappa)}{2\lambda + 2\pi D(\varkappa + 1)} \cos bx_1 + C,$$

$$\sigma_{11} = \frac{\sigma_0 \lambda + \pi\sigma_0 D(3 - \varkappa)}{\lambda + \pi D(\varkappa + 1)} \cos bx_1, \quad \sigma_{12} = -\frac{\pi\sigma_0 D(\varkappa - 1)}{\lambda + \pi D(\varkappa + 1)} \sin bx_1.$$

На рис. 1а, 2а приведены графики кольцевых напряжений на границе, полученные в случае первого вида нагружения, на рис. 1б, 2б – графики касательных напряжений в случае второго вида нагружения. В расчетах принято $\nu = 0,3$.

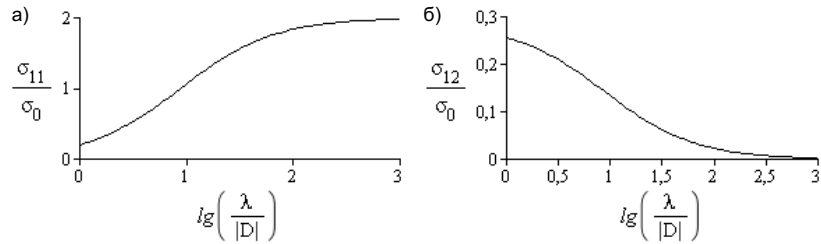


Рис. 1. Зависимость максимальных значений а) кольцевых и б) касательных напряжений от приведенной длины волны $\lambda/|D|$, $D > 0$

Заметим, что согласно работе [4] для алюминия при двух различных ориентациях кристаллов на поверхности $D = 0,113$ нм и

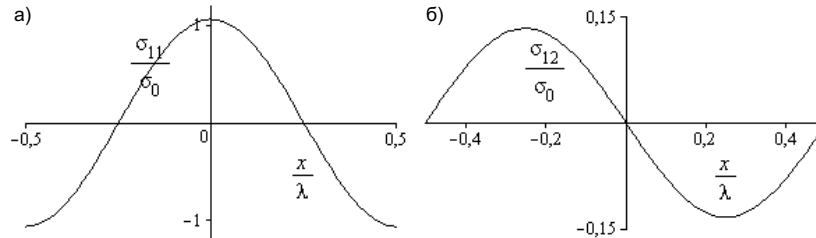


Рис. 2. Распределение а) кольцевых и б) касательных напряжений вдоль границы в пределах одного периода при $\lambda = 10|D|$, $D > 0$

$D = -0,166$ нм. Таким образом, значение длины волны $\lambda = 10|D|$ имеет порядок нанометра.

Литература

1. Подстригач Я.С., Повстенко Ю.З. Введение в механику поверхностных явлений в деформируемых твердых телах. Киев: Наукова думка, 1985. 200 с.
2. Gurtin M.E., Murdoch A.I. A continuum theory of elastic material surfaces // Arch. Ration. Mech. and Anal., 1975. Vol. 57, № 4. P. 291–323.
3. Альтенбах Х., Еремеев В.А., Морозов Н.Ф. Об уравнениях линейной теории оболочек при учете поверхностных напряжений // Изв. РАН. Серия: Механика твердого тела, 2010. № 3. С. 30–44.
4. Duan H.L., Wang J., Karimhaloo B.L. Theory of elasticity at the nanoscale // Advances in Applied Mechanics, 2008. № 42. P. 1–68.
5. Греков М.А. Сингулярная плоская задача теории упругости. СПб.: Изд-во СПбГУ, 2001. 192 с.
6. Линьков А.М. Комплексный метод граничных интегральных уравнений теории упругости. СПб.: Наука, 1999. 382 с.
7. Мусхелишвили Н.И. Некоторые основные задачи математической теории упругости. М.: Наука, 1966. 707 с.

Витвицкий С.И., Нестеров А.В., Нестеров С.В.
Кубанский государственный технологический университет

Анализ устойчивых решений уравнения Матье при моделировании периодически нестационарного объекта управления

Рекомендовано к публикации доцентом Терещенко И.В.

В настоящей работе приведены результаты моделирования параметрических колебаний ротора гироскопической системы (ГС). Технические подробности и математическую модель системы содержит работа [1]. Показано, что при установившейся угловой скорости ротора поведение ГС описывается однородным уравнением Матье

$$\ddot{z} + (l + 2q \cos 2\tau)z = 0. \quad (1)$$

Решением этого уравнения является функция $z(\tau)$, описывающая параметрические колебания ротора ГС. График этой функции служит удобным инструментом оценки устойчивости и качества изучаемого процесса (рис. 1).

Устойчивость ГС можно проверить также графически по диаграмме Айнса – Стретта [2], фрагмент которой показан на рисунке 2. Однако определить амплитуду и гармонический состав параметрических колебаний в этом случае невозможно. Поэтому синтез ГС проводят «вслепую», полагаясь на опыт и интуицию ее разработчика. В результате типового электромагнитного расчета ГС находят ее параметры, по которым затем вычисляют коэффициенты l и q уравнения Матье. Процедуру синтеза завершают проверкой устойчивости движения ротора ГС. Для этого используют аналитический критерий $ch(\mu\pi)$ [3] или диаграмму Айнса – Стретта. По уже известным координатам l и q на диаграмму наносят точку N_1 (рис. 2). Последняя может находиться в области устойчивости, неустойчивости или располагаться на границе названных областей b_1 .

Очевидно, что колебания ротора исследуемой ГС должны быть устойчивыми. В частности, такие колебания изображены на рис. 1. Вместе с тем, характеристическая точка N_1 должна лежать в первой области устойчивости, образованной осями координат и границей устойчивости b_1 (рис. 2). Любое положение характеристической

точки N_1 внутри этой области свидетельствует об устойчивости параметрических колебаний.

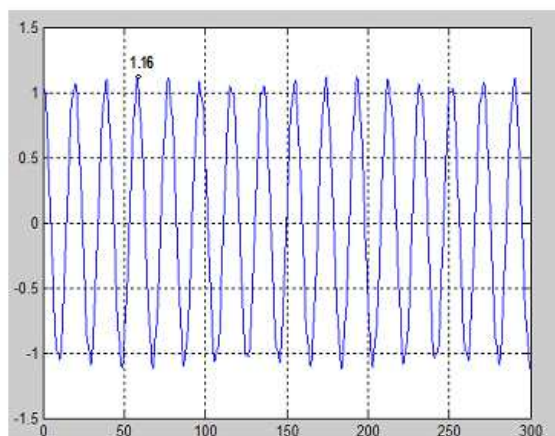


Рис. 1. График параметрических колебаний ротора ГС $z(\tau)$ при $l = 0,1$ и $q = 0,1$

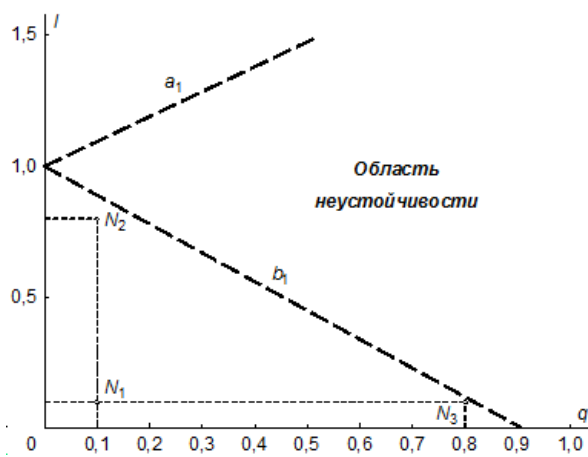


Рис. 2. Диаграмма Айнса – Стретта

Однако для эффективной эксплуатации рассматриваемой ГС одной положительной оценки устойчивости движения ее ротора может оказаться недостаточно. Может потребоваться также, чтобы амплитуда параметрических колебаний ротора не превышала некоторого допустимого значения $z_{\text{доп}}$. В связи с этим возникает задача поиска такого сочетания коэффициентов l и q (параметров ГС), которое обеспечит выполнение названного условия $z(\tau) \leq z_{\text{доп}}$. Поэтому с практической точки зрения возможное решение этой задачи целесообразно представить непосредственно на диаграмме Айнса – Стретта, изобразив на ней зависимости амплитуды параметрических колебаний от коэффициентов уравнения Матъе.

Искомые амплитуды определены в вычислительном эксперименте при решении задачи Коши для уравнения Матъе, приведенного к нормальной форме

$$\begin{cases} \dot{x}_1 = x_2, \\ \dot{x}_2 = -(l + 2q \cos 2\tau)x_1, \end{cases} \quad (2)$$

где $z = x_1$.

Численное интегрирование системы дифференциальных уравнений (2) произведено при начальных условиях

$$\begin{cases} x_1(0) = 1, \\ x_2(0) = 0 \end{cases}$$

с помощью функции *ode23tb* системы компьютерной математики MATLAB. Выбор этой функции обусловлен жесткостью уравнения Матъе и сделан согласно рекомендациям [4].

Диапазон изменения коэффициентов l и q соответствует размерам первой области устойчивости диаграммы Айнса – Стретта, т.е.

$$\begin{cases} 0 < l < 1, \\ 0 < q < 1. \end{cases}$$

Полученные решения $z(\tau) = x_1(\tau)$ отличаются друг от друга как в количественном (амплитудами A и частотами ω), так и в качественном отношении. Для сравнения на рис. 1, 3 и 4 показаны решения уравнения Матъе, коэффициенты которого соответственно равны

$$\begin{aligned} l = 0,1 \quad \text{и} \quad q = 0,1; \\ l = 0,8 \quad \text{и} \quad q = 0,1, \end{aligned} \quad (3)$$

$$l = 0,1 \quad \text{и} \quad q = 0,8. \quad (4)$$

На диаграмме Айнса – Стретта им соответствуют характеристические точки N_1 , N_2 и N_3 (рис. 2). Заметим, что последние две точки расположены в непосредственной близости от границы устойчивости.

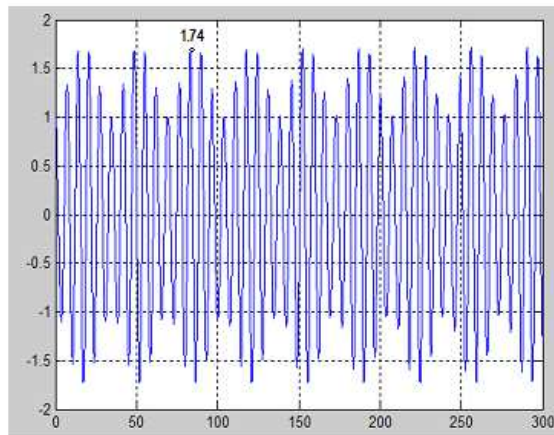


Рис. 3. График функции $z(\tau)$ при $l = 0,8$ и $q = 0,1$

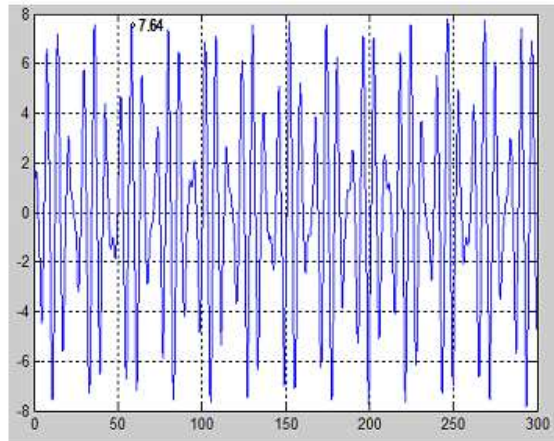


Рис. 4. График функции $z(\tau)$ при $l = 0,1$ и $q = 0,8$

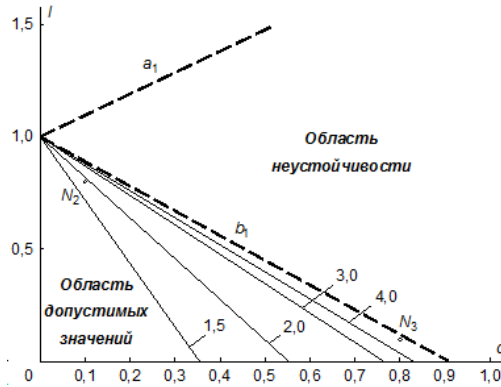


Рис. 5. Модифицированная диаграмма Айнса – Стретта с изоамплитудами

Анализ функций $z(\tau)$ свидетельствует о зависимости амплитуды параметрических колебаний A от коэффициентов уравнения Маттье, а также их соотношения. Наглядно эту зависимость выражают линии равной амплитуды $A = \text{const}$ (изоамплитуды). На рис. 5 семейство изоамплитуд совмещено с диаграммой Айнса – Стретта. На концах линий указаны соответствующие амплитуды параметрических колебаний.

Такая (модифицированная) диаграмма Айнса – Стретта может служить инструментом для выбора коэффициентов l и q (параметров ГС). Если при этом должно выполняться условие $z(\tau) < z_{\text{доп}}$, то допустимому значению $z_{\text{доп}}$ на модифицированной диаграмме соответствует изоамплитуда $A = \text{const}$. Последняя делит диаграмму на две части. Очевидно, что для выполнения условия $z(\tau) < z_{\text{доп}}$ характеристическая точка N должна быть левее линии равной амплитуды $A = z_{\text{доп}}$. Эту часть зоны устойчивости можно назвать областью допустимых значений.

Ранее уже был сделан вывод о том, что следует принимать такие значения коэффициентов l и q , при которых характеристическая точка располагается ближе к началу координат [5]. Модифицированная диаграмма Айнса – Стретта позволяет дать более точные рекомендации. Например, при проектировании сравниваются два варианта ГС. Их параметры (3), (4) таковы, что обе системы устойчивы. Соответствующие им характеристические точки N_2 и N_3 лежат в первой области устойчивости и равноудалены от начала координат (рис. 5), но в первом случае амплитуда параметрических колебаний ротора ГС $A_2 = 1,74$ (рис. 3), а во втором больше и равна $A_3 = 7,64$ (рис. 4).

Рассмотренный пример показывает преимущество первого варианта над вторым. На диаграмме Айнса – Стретта указанное превосходство выражается положением характеристической точки N_2 , которая находится ближе к оси ординат, чем точка N_3 .

Таким образом, сформулированная выше цель исследований достигнута. Для анализа устойчивых параметрических колебаний ротора ГС предложена модифицированная диаграмма Айнса – Стретта. Построенные в первой области устойчивости диаграммы изоамплитуды позволяют графически получить приближенную оценку амплитуды параметрических колебаний без интегрирования уравнения Матье. Кроме того, даны конкретные рекомендации по выбору коэффициентов уравнения Матье (параметров ГС), обеспечивающих ограничение амплитуды параметрических колебаний допустимым значением.

Литература

1. Добробаба Ю.П., Нестеров А.В., Нестеров С.В. Электропривод роторного зернистого фильтра и сирены. СПб.: Энергоатомиздат, 1991. 48 с.
2. Мак-Лахлан Н.В. Теория и приложения функций Матье. М.: Изд-во ИЛ, 1953. 475 с.
3. Кобринский А.Е. Механизмы с упругими связями. Динамика и устойчивость. М.: Наука, 1964. 392 с.
4. Канифольский Д.С., Нестеров А.В., Нестеров С.В. Цапов Г.М. О проблемах численного решения задачи Коши при моделировании линейных САУ в системе MATLAB // Процессы управления и устойчивость: Труды 41-й международной научной конференции аспирантов и студентов / Под ред. Н.В. Смирнова, Г.Ш. Тамасяна. СПб.: Издат. Дом С.-Петербур. гос. ун-та, 2010. С. 429–436.
5. Басси А.В., Нестеров А.В., Нестеров С.В. Исследование параметрических колебаний в электромеханических системах с переменным моментом инерции: Труды XVIII Международной Интернет-конференции молодых ученых и студентов по проблемам машиноведения (МИКМУС-2006). М.: Изд-во ИМАШ РАН, 2006. С. 166.

Дронь В.О.

Санкт-Петербургский государственный университет

О минимальных сплайнах третьего порядка на замкнутом промежутке¹

Рекомендовано к публикации профессором Демьяновичем Ю.К.

1. Введение. Минимальные сплайны третьего порядка ранее исследовались для бесконечной сетки на открытом интервале (см. [2]). Поскольку на практике обрабатываемые числовые потоки конечны, важно рассмотреть случай конечной сетки на отрезке вещественной оси. Для минимальных сплайнов второго порядка такое исследование проведено в работе [3].

Цель данной работы – провести подобное исследование для минимальных сплайнов третьего порядка. Здесь получены представления пространств минимальных сплайнов на замкнутом интервале в условиях, когда сплайновая сетка конечна.

2. О представлениях (X, A, φ) -сплайнов. Рассмотрим представления минимальных сплайнов третьего порядка на открытом и замкнутом интервалах.

2.1. Открытый интервал. На интервале $(\alpha, \beta) \in \mathbb{R}^1$ рассмотрим сетку

$$X : \dots < x_{-2} < x_{-1} < x_0 < x_1 < x_2 < \dots,$$

где $\alpha = \lim_{j \rightarrow -\infty} x_j$, $\beta = \lim_{j \rightarrow +\infty} x_j$ (случаи $\alpha = -\infty$, $\beta = +\infty$ не исключаются).

Каждому индексу $j \in \mathbb{Z}$ поставим в соответствие вектор-столбец $\mathbf{a}_j \in \mathbb{R}^4$ и введем квадратную матрицу четвертого порядка $A_j \stackrel{\text{def}}{=} (\mathbf{a}_{j-3}, \mathbf{a}_{j-2}, \mathbf{a}_{j-1}, \mathbf{a}_j)$.

Определение 1. Бесконечная цепочка векторов $\mathbf{A} \stackrel{\text{def}}{=} \{\mathbf{a}_j\}_{j \in \mathbb{Z}}$ называется *полной*, если $\det A_j \neq 0 \quad \forall j \in \mathbb{Z}$.

Множество всех бесконечных полных цепочек обозначим \mathbb{A}_∞ .

Пусть цепочка \mathbf{A} лежит в \mathbb{A}_∞ , а $\varphi(t)$ – четырехкомпонентная вектор-функция, которая задана на интервале (α, β) , и компоненты

¹Работа выполнена при финансовой поддержке РФФИ, гранты № 10-01-00245, 10-01-00297

которой линейно независимы на любом подынтервале (c, d) интервала (α, β) .

Координатными (X, \mathbf{A}, φ) -сплайнами третьего порядка называются функции $\omega_j(t)$, удовлетворяющие аппроксимационным соотношениям

$$\sum_j \mathbf{a}_j \omega_j(t) = \varphi(t) \quad \forall t \in (\alpha, \beta), \quad \text{supp } \omega_j \subset [x_j, x_{j+4}].$$

Этим соотношениям можно придать вид

$$\mathbf{a}_{k-3} \omega_{k-3}(t) + \mathbf{a}_{k-2} \omega_{k-2}(t) + \mathbf{a}_{k-1} \omega_{k-1}(t) + \mathbf{a}_k \omega_k(t) = \varphi(t) \quad (1)$$

$$\forall t \in (x_k, x_{k+1}),$$

где

$$\text{supp } \omega_j \subset [x_j, x_{j+4}]. \quad (2)$$

Теорема 1. Если $\mathbf{A} \in \mathbb{A}_\infty$, то из соотношений (1), (2), рассматриваемых при всех $k \in \mathbb{Z}$, однозначно определяются сплайны:

$$\begin{aligned} \omega_j(t) &= \frac{\det(\mathbf{a}_{j-3}, \mathbf{a}_{j-2}, \mathbf{a}_{j-1}, \varphi(t))}{\det(\mathbf{a}_{j-2}, \mathbf{a}_{j-2}, \mathbf{a}_{j-1}, \mathbf{a}_j)} & \text{при } t \in (x_j, x_{j+1}), \\ \omega_j(t) &= \frac{\det(\mathbf{a}_{j-2}, \mathbf{a}_{j-1}, \varphi(t), \mathbf{a}_{j+1})}{\det(\mathbf{a}_{j-2}, \mathbf{a}_{j-1}, \mathbf{a}_j, \mathbf{a}_{j+1})} & \text{при } t \in (x_{j+1}, x_{j+2}), \\ \omega_j(t) &= \frac{\det(\mathbf{a}_{j-1}, \varphi(t), \mathbf{a}_{j+1}, \mathbf{a}_{j+2})}{\det(\mathbf{a}_{j-1}, \mathbf{a}_j, \mathbf{a}_{j+1}, \mathbf{a}_{j+2})} & \text{при } t \in (x_{j+2}, x_{j+3}), \\ \omega_j(t) &= \frac{\det(\varphi(t), \mathbf{a}_{j+1}, \mathbf{a}_{j+2}, \mathbf{a}_{j+3})}{\det(\mathbf{a}_j, \mathbf{a}_{j+1}, \mathbf{a}_{j+2}, \mathbf{a}_{j+3})} & \text{при } t \in (x_{j+3}, x_{j+4}), \\ \omega_j(t) &= 0 & \text{при } t \notin [x_j, x_{j+4}], \end{aligned}$$

где $j \in \mathbb{Z}$ и $t \in (\alpha, \beta) \setminus X$.

Установленные в теореме представления функции ω_j однозначно выводятся из условий (1), (2) для указанных в теореме j и t .

Пусть $\mathbf{A} \in \mathbb{A}_\infty$. Линейная оболочка $\mathbb{S}(X, \mathbf{A}, \varphi)$ функций $\{\omega_j\}_{j \in \mathbb{Z}}$ называется пространством сплайнов третьего порядка на сетке X :

$$\mathbb{S}(X, \mathbf{A}, \varphi) \stackrel{\text{def}}{=} \left\{ u \mid u = \sum_j c_j \omega_j \quad \forall c_j \in \mathbb{R}^1 \right\}.$$

Нетрудно видеть, что для вектора

$$\mathbf{d}_k \stackrel{\text{def}}{=} \left(\det \begin{pmatrix} [\mathbf{a}_{k-3}]_1 & [\mathbf{a}_{k-2}]_1 & [\mathbf{a}_{k-1}]_1 \\ [\mathbf{a}_{k-3}]_2 & [\mathbf{a}_{k-2}]_2 & [\mathbf{a}_{k-1}]_2 \\ [\mathbf{a}_{k-3}]_3 & [\mathbf{a}_{k-2}]_3 & [\mathbf{a}_{k-1}]_3 \end{pmatrix}, -\det \begin{pmatrix} [\mathbf{a}_{k-3}]_0 & [\mathbf{a}_{k-2}]_0 & [\mathbf{a}_{k-1}]_0 \\ [\mathbf{a}_{k-3}]_2 & [\mathbf{a}_{k-2}]_2 & [\mathbf{a}_{k-1}]_2 \\ [\mathbf{a}_{k-3}]_3 & [\mathbf{a}_{k-2}]_3 & [\mathbf{a}_{k-1}]_3 \end{pmatrix}, \right. \\ \left. \det \begin{pmatrix} [\mathbf{a}_{k-3}]_0 & [\mathbf{a}_{k-2}]_0 & [\mathbf{a}_{k-1}]_0 \\ [\mathbf{a}_{k-3}]_1 & [\mathbf{a}_{k-2}]_1 & [\mathbf{a}_{k-1}]_1 \\ [\mathbf{a}_{k-3}]_3 & [\mathbf{a}_{k-2}]_3 & [\mathbf{a}_{k-1}]_3 \end{pmatrix}, -\det \begin{pmatrix} [\mathbf{a}_{k-3}]_0 & [\mathbf{a}_{k-2}]_0 & [\mathbf{a}_{k-1}]_0 \\ [\mathbf{a}_{k-3}]_1 & [\mathbf{a}_{k-2}]_1 & [\mathbf{a}_{k-1}]_1 \\ [\mathbf{a}_{k-3}]_2 & [\mathbf{a}_{k-2}]_2 & [\mathbf{a}_{k-1}]_2 \end{pmatrix} \right)^T$$

справедливы следующие соотношения

$$\mathbf{d}_j^T \mathbf{a}_{j-4} \neq 0, \quad \mathbf{d}_j^T \mathbf{a}_{j-3} = \mathbf{d}_j^T \mathbf{a}_{j-2} = \mathbf{d}_j^T \mathbf{a}_{j-1} = 0, \quad \mathbf{d}_j^T \mathbf{a}_j \neq 0. \quad (3)$$

Первое и последнее соотношение являются следствием полноты цепочки; оставшиеся можно установить, заметив, что $\mathbf{d}_k^T \mathbf{x} = \det(\mathbf{a}_{k-3}, \mathbf{a}_{k-2}, \mathbf{a}_{k-1}, \mathbf{x})$.

Теорема 2. Для функций ω_j , $j = k-3, k-2, k-1, k$, на промежутке (x_k, x_{k+1}) справедливы представления

$$\omega_{k-3}(t) = \frac{\mathbf{d}_{k+1}^T \varphi(t)}{\mathbf{d}_{k+1}^T \mathbf{a}_{k-3}}, \quad (4)$$

$$\omega_{k-2}(t) = \frac{\mathbf{d}_{k+2}^T \varphi(t)}{\mathbf{d}_{k+2}^T \mathbf{a}_{k-2}} - \frac{\mathbf{d}_{k+2}^T \mathbf{a}_{k-3}}{\mathbf{d}_{k+2}^T \mathbf{a}_{k-2}} \frac{\mathbf{d}_{k+1}^T \varphi(t)}{\mathbf{d}_{k+1}^T \mathbf{a}_{k-3}}, \quad (5)$$

$$\omega_{k-1}(t) = \frac{\mathbf{d}_{k-1}^T \varphi(t)}{\mathbf{d}_{k-1}^T \mathbf{a}_{k-1}} - \frac{\mathbf{d}_{k-1}^T \mathbf{a}_k}{\mathbf{d}_{k-1}^T \mathbf{a}_{k-1}} \frac{\mathbf{d}_k^T \varphi(t)}{\mathbf{d}_k^T \mathbf{a}_k}, \quad (6)$$

$$\omega_k(t) = \frac{\mathbf{d}_k^T \varphi(t)}{\mathbf{d}_k^T \mathbf{a}_k}, \quad (7)$$

причем каждое из них зависит лишь от векторов \mathbf{a}_{k-3} , \mathbf{a}_{k-2} , \mathbf{a}_{k-1} , \mathbf{a}_k .

Доказательство. Доказываемые формулы являются одним из вариантов представлений $\omega_j(t)$, получаемых путем последовательного умножения (1) на \mathbf{d}_s^T , где $s \in \{k-3, k-2, k-1, k\}$ или $s \in \{k+1, k+2, k+3, k+4\}$. Зависимость доказываемых формул лишь от перечисленных в теореме векторов очевидна, ибо при фиксированном $t \in (x_k, x_{k+1})$ они представляют собой решение системы линейных уравнений (1).

Теорема доказана.

Теорема 3. Для функции ω_j справедливы представления

$$\begin{aligned}\omega_j(t) &= \frac{\mathbf{d}_j^T \varphi(t)}{\mathbf{d}_j^T \mathbf{a}_j} && \text{при } t \in (x_j, x_{j+1}), \\ \omega_j(t) &= \frac{\mathbf{d}_j^T \varphi(t)}{\mathbf{d}_j^T \mathbf{a}_j} - \frac{\mathbf{d}_j^T \mathbf{a}_{j+1}}{\mathbf{d}_j^T \mathbf{a}_j} \frac{\mathbf{d}_{j+1}^T \varphi(t)}{\mathbf{d}_{j+1}^T \mathbf{a}_{j+1}} && \text{при } t \in (x_{j+1}, x_{j+2}), \\ \omega_j(t) &= \frac{\mathbf{d}_{j+4}^T \varphi(t)}{\mathbf{d}_{j+4}^T \mathbf{a}_j} - \frac{\mathbf{d}_{j+4}^T \mathbf{a}_{j-1}}{\mathbf{d}_{j+4}^T \mathbf{a}_j} \frac{\mathbf{d}_{j+3}^T \varphi(t)}{\mathbf{d}_{j+3}^T \mathbf{a}_{j-1}} && \text{при } t \in (x_{j+2}, x_{j+3}), \\ \omega_j(t) &= \frac{\mathbf{d}_{j+4}^T \varphi(t)}{\mathbf{d}_{j+4}^T \mathbf{a}_j} && \text{при } t \in (x_{j+3}, x_{j+4}).\end{aligned}$$

Координатная функция $\omega_j(t)$ определяется узлами $x_j, x_{j+1}, x_{j+2}, x_{j+3}, x_{j+4}$ и векторами $\mathbf{a}_{j-3}, \mathbf{a}_{j-2}, \mathbf{a}_{j-1}, \mathbf{a}_j, \mathbf{a}_{j+1}, \mathbf{a}_{j+2}, \mathbf{a}_{j+3}$.

Доказательство. Представления функции ω_j , устанавливаемые в теореме, вытекают из формул (4)–(7), где следует взять $k = j, j+1, j+2, j+3$. Зависимость от перечисленных узлов очевидна, а зависимость от указанных в теореме векторов получается из предыдущей теоремы при $k = j, j+1, j+2, j+3$.

Теорема доказана.

Теорема 4. Пусть вектор-функция $\varphi(t)$ s -кратно кусочно дифференцируема на интервале (α, β) и ее s -ая производная может быть продолжена на интервал (α, β) непрерывным образом. Для того чтобы s -ую производную функции $\omega_j(t)$ можно было продолжить на (α, β) непрерывным образом, необходимо и достаточно, чтобы

$$\det \left(\mathbf{a}_{k-3}, \mathbf{a}_{k-2}, \mathbf{a}_{k-1}, \varphi^{(s)}(x_k) \right) = 0 \quad \forall k \in \mathbb{Z}.$$

Доказательство эквивалентности последнего соотношения и s -кратной дифференцируемости $\omega_j(t)$ аналогично доказательству, приведенному в [3].

Замечание 1. Если $\varphi(t) = (1, t, t^2, t^3)^T$, то ω_j – полиномиальный сплайн третьей степени с носителем $\text{supp } \omega_j = [x_j, x_{j+4}]$.

2.2. Закрытый промежуток. Из бесконечной сетки X выделим конечную сетку X_N ,

$$X_N : x_0 < x_1 < \dots < x_{N-1} < x_N,$$

а из полной бесконечной цепочки \mathbf{A} выделим конечную цепочку векторов \mathbf{A}_N ,

$$\mathbf{A}_N = \{\mathbf{a}_{-3}, \mathbf{a}_{-2}, \mathbf{a}_{-1}, \dots, \mathbf{a}_{N-1}\}.$$

Введем обозначения

$$a \stackrel{\text{def}}{=} x_0, \quad b \stackrel{\text{def}}{=} x_N, \quad J_s \stackrel{\text{def}}{=} \{-3, -2, -1, 0, \dots, s\}, \quad s \in \mathbb{N}.$$

Сузим все функции пространства $\mathbb{S}(X, \mathbf{A}, \varphi)$ на множество $[a, b] \setminus X$. Совокупность этих сужений представляет собой конечномерное линейное пространство

$$\mathbb{S}_{[a,b]}(X_N, \mathbf{A}_N, \varphi) \stackrel{\text{def}}{=} \left\{ u \mid u = \sum_{j \in J_{N-1}} c_j \omega_j \quad \forall c_j \in \mathbb{R}^1, j \in J_{N-1} \right\}.$$

Для измеримого (по Лебегу) множества $\mathcal{M} \subset \mathbb{R}^1$ обозначим $\text{mes}(\mathcal{M})$ его лебегову меру.

Определение 2. Пусть система $\{f_j\}$ состоит из функций $f_j(t)$, заданных почти везде на промежутке (α, β) , и $[a, b] \subset (\alpha, \beta)$. Система функций $\{f_j \mid \text{mes}(\text{supp} f_j \cap (a, b)) > 0\}$ называется *сужением* системы $\{f_j\}$ на промежуток $[a, b]$.

Теорема 5. Пусть цепочка векторов $\{\mathbf{a}_j\}_{j \in \mathbb{Z}}$ полная, а компоненты вектор-функции $\varphi(t)$ представляют собой линейно независимую систему функций на любом интервале $(c, d) \subset (\alpha, \beta)$. Тогда сужение системы функций $\{\omega_j\}_{j \in \mathbb{Z}}$ на любой интервал $(c, d) \subset (\alpha, \beta)$ является линейно независимой системой на упомянутом интервале.

Доказательство. Из полноты цепочки $\{\mathbf{a}_j\}_{j \in \mathbb{Z}}$ следует однозначная разрешимость аппроксимационных соотношений (1), (2), а из линейной независимости компонент правой части – линейная независимость системы функций $\{\omega_j\}_{j \in \mathbb{Z}}$.

Теорема доказана.

Для дальнейших построений кроме векторов цепочки \mathbf{A} потребуется привлечь еще векторы $\mathbf{a}_{-6}, \mathbf{a}_{-5}, \mathbf{a}_{-4}, \mathbf{a}_N, \mathbf{a}_{N+1}, \mathbf{a}_{N+2}$. Заметим, что благодаря полноте бесконечной цепочки \mathbf{A} конечная цепочка векторов $\mathbf{A}' \stackrel{\text{def}}{=} \{\mathbf{a}_{-6}, \mathbf{a}_{-5}, \dots, \mathbf{a}_{N+2}\}$ обладает свойством

$$\det(\mathbf{a}_{j-3}, \mathbf{a}_{j-2}, \mathbf{a}_{j-1}, \mathbf{a}_j) \neq 0 \quad \forall j \in \{-3, -2, \dots, N+2\}.$$

Построим цепочку $\{\mathbf{d}_j\}$, полагая $\mathbf{d}_j^T \mathbf{x} \stackrel{\text{def}}{=} \det(\mathbf{a}_{j-3}, \mathbf{a}_{j-2}, \mathbf{a}_{j-1}, \mathbf{x})$, где $j = -3, -2, \dots, N, N+1, N+2$.

Теорема 6. В условиях теоремы 5 сужения функций ω_j образуют линейно независимую систему на $[a, b]$. Для них справедливы заключения теорем 2 и 3. Кроме того,

$$\dim \mathbb{S}_{[a,b]}(X_N, \mathbf{A}_N, \varphi) = N + 3.$$

Доказательство теоремы следует из определения сужения системы функций и из определения пространства $\mathbb{S}_{[a,b]}(X_N, \mathbf{A}_N, \varphi)$ применением теоремы 5.

Теорема 7. Функция $u_N(t) \stackrel{\text{def}}{=} \sum_{j=-3}^{N-1} c_j \omega_j(t)$, $t \in [a, b]$, является следом функции $u \stackrel{\text{def}}{=} \sum_j c_j \omega_j$ на отрезке $[a, b]$, лежит в пространстве $\mathbb{S}_{[a,b]}(X_N, \mathbf{A}_N, \varphi)$ и полностью определяется набором узлов $\{x_j\}_{j \in \{0, \dots, N\}}$, набором векторов $\{\mathbf{a}_j\}_{j \in J_{N-1}}$ и набором коэффициентов $\{c_j\}_{j \in J_{N-1}}$.

Утверждение теоремы очевидным образом вытекает из теоремы 2.

Полученные представления пространств минимальных сплайнов третьего порядка на конечной сетке позволяют проводить дальнейшее исследование данных функций на замкнутом промежутке. В частности, можно получить представления формул декомпозиции и реконструкции для соответствующих сплайн-вейвлетных разложений.

Литература

1. Завьялов Ю.С., Квасов Б.И., Мирошниченко В.Л. Методы сплайн-функций. М.: Наука, 1980. 352 с.
2. Демьянович Ю.К. Всплесковые разложения на неравномерной сетке // Труды СПбМОб, 2007. Т. 13. С. 27–51
3. Демьянович Ю.К., Иванцова О.Н. Гладкость пространств сплайнов третьего порядка // Сб. Математические модели. Теория и приложения, 2006. Вып. 7. СПб. С. 58–64.

Иванов А.Н.

Санкт-Петербургский государственный университет

Символьные вычисления в моделировании динамики пучков заряженных частиц

Рекомендовано к публикации профессором Андриановым С.Н.

Введение. Моделирование динамики пучков заряженных частиц является востребованной и актуальной задачей. Существует несколько подходов к исследованию и анализу динамики пучка. В работе использован один из возможных методов интегрирования дифференциальных уравнений – матричный формализм [1]. Особенностью данного подхода является то, что он позволяет в удобной (компактной) форме записывать решение нелинейного уравнения движения в символьном виде. Кроме того, указанный метод может использоваться в технологии параллельного программирования и успешно применяться для исследования длительной эволюции пучка заряженных частиц [2].

В качестве среды для проведения вычислительного эксперимента выбран пакет *Mathima* [3]. Данная система компьютерной алгебры является бесплатной и, в то же время, предоставляет широкие возможности для проведения как символьных, так и численных вычислений с последующей визуализацией полученных результатов.

Основные положения. Матричный формализм позволяет записать решение динамической системы в аналитическом виде как сумму бесконечного ряда [1]. В общем случае уравнение движения может быть записано в виде

$$\frac{d\mathbf{X}}{ds} = \mathbf{F}(\mathbf{X}, s). \quad (1)$$

При условии измеримости по s функции $\mathbf{F}(\mathbf{X}, s)$ и $\mathbf{F}(\mathbf{X}, s) \in C^\infty(x)$, уравнение (1) можно представить в виде:

$$\frac{d\mathbf{X}}{ds} = \sum_{k=0}^{\infty} \mathbb{P}^{1k}(s) \mathbf{X}^{[k]}, \quad (2)$$

где $\mathbf{X}^{[k]} = \underbrace{\mathbf{X} \otimes \dots \otimes \mathbf{X}}_{k \text{ раз}}$ – кронекерова степень вектора, а матрицы \mathbb{P}^{1k} вычисляются по формулам

$$\{\mathbb{P}^{1k}(t)\}_{ij} = \frac{1}{k_1! \dots k_n!} \frac{\partial^k \mathbf{F}_i(\mathbf{X}, t)}{\partial x_1^{k_1} \dots \partial x_n^{k_n}} \Big|_{x_1=\dots=x_n=0}.$$

Замечание. Здесь и далее будем различать кронекерово произведение векторов и обычное тензорное. При вычислении кронекерова произведения при наличии совпадающих компонент будем переходить к вектору меньшей размерности. Например, для $\mathbf{X} = (x_1, x_2)^T$ вместо обычного тензорного произведения векторов $\mathbf{X} \otimes \mathbf{X} = (x_1^2, x_1x_2, x_1x_2, x_2^2)^T$ получим вектор $\mathbf{X}^{[2]} = (x_1^2, x_1x_2, x_2^2)^T$.

Более подробную информация по вычислению матриц \mathbb{P}^{1k} и кронекерова произведения можно найти в [1]. Там же описан метод последовательных приближений, руководствуясь которым, решение уравнения (2) можно записать в виде

$$\mathbf{X}(s) = \sum_{\mathbf{k}=0}^{\infty} \mathbb{R}^{1k}(s|s_0) \mathbf{X}_0^{[k]}, \quad (3)$$

где $\mathbf{X}_0 = \mathbf{X}(s_0)$ – начальное условие в задаче Коши для (1).

Постановка задачи и решение. В рамках данной работы рассматривается движение частиц в циклических ускорителях, основанных на электростатических элементах. В качестве примера выводится решение уравнения движения частиц в цилиндрическом конденсаторе в виде, соответствующем формуле (3).

В стационарном случае, когда частица движется в окрестности траектории с постоянным радиусом R_{eq} , уравнение её движения имеет вид [4]:

$$\begin{aligned} x'' + \frac{2}{R_{eq}^2} x - \frac{1}{R_{eq}^3} x^2 &= 0, \\ y'' &= 0, \end{aligned} \quad (4)$$

где $R_{eq}^2 = M^2 \ln \frac{R_2}{R_1} / (2emU_0)$, $M = mvr$ – некоторый инвариант, R_1 и R_2 – радиусы внутреннего и внешнего электродов соответственно, $2U_0$ – разность потенциалов между электродами.

Легко заметить, что система уравнений (4) представляет собой два независимых уравнения второго порядка. Поэтому рассмотрим движение частицы в пространстве (x, x') :

$$\frac{d}{ds} \begin{pmatrix} x \\ x' \end{pmatrix} = \begin{pmatrix} x' \\ -2/R_{eq}^2 x + 1/R_{eq}^3 x^2 \end{pmatrix}. \quad (5)$$

В соответствии с формулой (2) уравнение (5) может быть представлено в виде

$$\frac{d}{ds} \begin{pmatrix} x \\ x' \end{pmatrix} = \sum_{k=0}^{\infty} \mathbb{P}^{1k}(s) \begin{pmatrix} x \\ x' \end{pmatrix}^{[k]}, \quad (6)$$

где $\mathbb{P}^{11} = \begin{pmatrix} 0 & 1 \\ -2/R_{eq}^2 & 0 \end{pmatrix}$, $\mathbb{P}^{12} = \begin{pmatrix} 0 & 0 & 0 \\ 1/R_{eq}^3 & 0 & 0 \end{pmatrix}$, $\mathbb{P}^{1k} = \mathbb{O}$, $k = \overline{3, \infty}$.

Руководствуясь алгоритмом, описанным в работе [1], построим решение уравнения (6), используя систему компьютерной алгебры Махима [4].

Пусть \mathbb{R}^{11} – матрицант линейризованного уравнения $\frac{d\mathbf{X}}{dt} = \mathbb{P}^{11}\mathbf{X}$, решение которого может быть найдено при помощи следующего кода:

```
(%i1) eq1: 'diff(x(s),s)=y(s)$
(%i2) eq2: 'diff(y(s),s)=-b^2*x(s)$
(%i3) desolve([eq1,eq2], [x(s),y(s)]);
(%o3) [x(s)=(y(0)*sin(b*s))/b+x(0)*cos(b*s),
       y(s)=y(0)*cos(b*s)-x(0)*b*sin(b*s)]
```

Таким образом, при $b = \sqrt{2}/R_{eq}$ имеем

$$\mathbb{R}^{11}(s) = \begin{pmatrix} \cos(b(s-s_0)) & \frac{1}{b} \sin(b(s-s_0)) \\ -b \sin(b(s-s_0)) & \cos(b(s-s_0)) \end{pmatrix}.$$

Для построения матрицы $\mathbb{R}^{12}(s|0) = \int_0^s \mathbb{R}^{11}(s|\tau) \mathbb{P}^{12}(\tau) \mathbb{R}^{22}(\tau|0) d\tau$

требуется дополнительно вычислить матрицу $\mathbb{R}^{22} = (\mathbb{R}^{11})^{[2]}$.

Замечание. При вычислении кронекеровой степени квадратной $(n \times n)$ -матрицы A также происходит уменьшение размерности результата в соответствии с тождеством $(AX)^{[k]} = A^{[k]}X^{[k]}$, где X – вектор-столбец размерности n . Отметим, что при использовании

символьных операций, вычисление кронекеровой степени матрицы сводится к вычислению коэффициентов произведения полиномов соответствующих степеней [1].

Отметим, что операции возведения полинома в степень и подсчёта коэффициентов при подобных слагаемых реализуемы в кодах Maxima и представимы в виде набора простых команд:

```
(%i4) R11: matrix([cos(b*(s-s_0)), 1/b*sin(b*(s-s_0))],
                  [-b*sin(b*(s-s_0)), cos(b*(s-s_0))]);
(%i5) expres1: expand((R11[1,1]*x+R11[1,2]*y)^2)$
(%i6) expres2: expand((R11[1,1]*x+R11[1,2]*y)*
                    (R11[2,1]*x+R11[2,2]*y))$
(%i7) expres3: expand((R11[2,1]*x+R11[2,2]*y)^2);
(%i8) R22: matrix([coeff(expres1,x^2),ev(coeff(expres1,x),
                                             y=1),coeff(expres1,y^2)],
                  [coeff(expres2,x^2),ev(coeff(expres2,x),
                                             y=1),coeff(expres2,y^2)],
                  [coeff(expres3,x^2),ev(coeff(expres3,x),
                                             y=1),coeff(expres3,y^2)]);
(%o7)
```

$$\begin{aligned} & \cos(b(s-s_0))^2 y^2 - 2b \cos(b(s-s_0)) \sin(b(s-s_0)) xy + \\ & + b^2 \sin(b(s-s_0))^2 x^2 \end{aligned}$$

(%o3)

Продолжая вычисления в соответствии с алгоритмом, описанным в [1], получим окончательный результат:

$$\mathbf{X}(s) = \sum_{k=1}^{\infty} \mathbb{R}^{1k}(s|0) \mathbf{X}_0^{[k]},$$

где при $b = \sqrt{2}/R_{eq}$ имеем

$$\begin{aligned} \mathbb{R}^{11}(s|0) &= \begin{pmatrix} \cos(bs) & \frac{1}{b} \sin(bs) \\ -b \sin(bs) & \cos(bs) \end{pmatrix}, \\ \mathbb{R}^{12}(s|0) &= \begin{pmatrix} R_{11}^{12} & R_{12}^{12} & R_{13}^{12} \\ R_{21}^{12} & R_{22}^{12} & R_{23}^{12} \end{pmatrix}, \\ \mathbb{R}^{1k}(s|0) &= \mathbb{O}, k = \overline{3, \infty}, \end{aligned}$$

$$\begin{aligned}
R_{11}^{12} &= \frac{b^2}{2^{3/2}} \left(-\frac{(\cos(2bs) - 3)}{6b} - \frac{\cos(bs)}{3b} \right), \\
R_{12}^{12} &= \frac{b}{2^{1/2}} \left(\frac{\sin(bs)}{3b} - \frac{\sin(2bs)}{6b} \right), \\
R_{13}^{12} &= \frac{b^2}{2^{3/2}} \left(\frac{(\cos(2bs) + 3)}{6b} - \frac{2\cos(bs)}{3b} \right), \\
R_{21}^{12} &= \frac{b^3}{2^{3/2}} \left(\frac{\sin(2bs)}{3b} + \frac{\sin(bs)}{3b} \right), \\
R_{22}^{12} &= \frac{b^2}{2^{1/2}} \left(\frac{\cos(bs)}{3b} - \frac{\cos(2bs)}{3b} \right), \\
R_{23}^{12} &= \frac{b^2}{2^{3/2}} \left(\frac{2\sin(bs)}{3b} - \frac{\sin(2bs)}{3b} \right).
\end{aligned}$$

Таким образом, решение системы (6) с точностью до нелинейных членов второго порядка может быть записано как

$$\mathbf{X}(s) = \mathbb{R}^{11}(s) (x_0, x'_0)^T + \mathbb{R}^{12}(s) (x_0^2, x_0 x'_0, x_0'^2)^T. \quad (7)$$

Заключение. Существуют различные способы интегрирования систем дифференциальных уравнений [5]. Для примера, сравним два метода решения системы (6). На рис. 1 приведены графики решения ($x_0 = 0, 2$, $x'_0 = 0$), полученные методом матричного формализма (7) и численным методом Рунге – Кутты четвертого порядка.

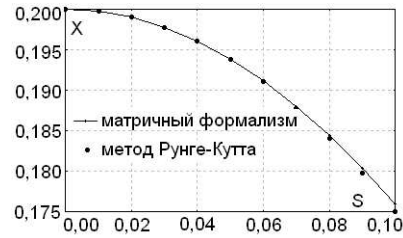


Рис. 1. Графики решений

Взаимная ошибка (расхождение результата) двух приведенных методов имеет порядок 10^{-3} . Отметим, что в отличие от итеративных методов интегрирования, матричный формализм позволяет записать решение в символьном виде. Данное свойство может

быть полезно для проведения аналитических исследований построенных решений. Кроме того, такой алгоритм построения решения может быть легко распараллелен. В то время как большинство численных методов требуют для этих целей специальных модификаций.

Развитие систем компьютерной алгебры упрощает работу по вычислению матриц большой размерности и предоставляет возможности их автоматического построения и верификации конечного результата.

Литература

1. Андрианов С.Н. Динамическое моделирование систем управления пучками частиц. СПб.: Изд-во СПбГУ, 2002. 376 с.
2. Andrianov S.N. Role of parallel and distributed computing in beam physics // Nuclear Instruments and Methods, 2004. Vol. 519. P. 37–41.
3. Maxima, a Computer Algebra System. <http://maxima.sourceforge.net/ru/>
4. Senichev Y., Moeller S.P. Beam Dynamics in Electrostatic Rings. <http://accelconf.web.cern.ch/AccelConf/PAPERS/MOP1B04.pdf>
5. Рошаль А.С. Моделирование заряженных пучков. М.: Атомиздат, 1979. 224 с.

Иванов В.А.

Санкт-Петербургский государственный университет

Двухкомпонентная плоскость с криволинейной трещиной в окрестности межфазной линии¹

Рекомендовано к публикации профессором Мальковым В.М.

Введение. Рассматривается плоская задача теории упругости (плоская деформация и плоское напряженное состояние) для бесконечной упругой плоскости, содержащей криволинейную трещину в форме дуги окружности вблизи линии раздела полуплоскостей. На бесконечности заданы внешние усилия и угол поворота. При решении задачи используются результаты работ [1–4].

Постановка задачи. Рассматривается плоскость S , состоящая из двух полуплоскостей S_1 и S_2 , выполненных из разных материалов. В декартовых координатах (x_1, x_2) прямая $x_2 = 0$ является линией раздела полуплоскостей. Предполагаем, что криволинейная трещина расположена целиком в нижней полуплоскости S_1 .

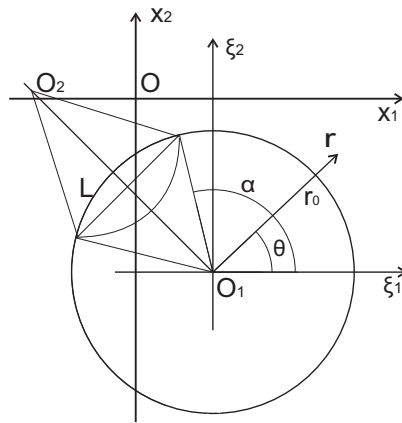


Рис. 1. Двухкомпонентная плоскость с криволинейной трещиной

Введем другую систему декартовых координат (ξ_1, ξ_2) с началом координат в центре окружности O_1 , на которой расположена трещина (рис. 1), радиус окружности обозначим $r = r_0$.

Для комплексных переменных $z = x_1 + ix_2$ и $\zeta = \xi_1 + i\xi_2$ имеем зависимость $z = z_0 + \zeta$, где z_0 – положение центра окружности. В полярных координатах (r, θ) трещине соответствует промежуток $\alpha \leq \theta \leq \beta$. На рис. 1 показаны два вида трещин: выпуклая вверх и выпуклая вниз. Радиусы окружностей трещин одинаковы и равны r_0 .

¹Работа выполнена при финансовой поддержке РФФИ, проект № 09-01-00656

Граничные условия задачи. На бесконечности при $|z| \rightarrow \infty$ заданы напряжения и углы поворота (свои для каждой полуплоскости $S_k, k = 1, 2$)

$$\sigma_{ij} \rightarrow \sigma_{ij}^{k\infty}, \quad \omega \rightarrow \omega^{k\infty}, \quad |z| \rightarrow \infty,$$

причем

$$\sigma_{12}^{1\infty} = \sigma_{12}^{2\infty} = \sigma_{12}^{\infty}, \quad \sigma_{22}^{1\infty} = \sigma_{22}^{2\infty} = \sigma_{22}^{\infty}.$$

На берегах криволинейной трещины L задана внешняя нагрузка

$$(\sigma_{rr} - i\sigma_{r\theta})^+ = p^*(z_c), \quad (\sigma_{rr} - i\sigma_{r\theta})^- = q^*(z_c),$$

где $\sigma_{rr}, \sigma_{r\theta}$ – компоненты тензора напряжений в базисе нормали и касательной к контуру трещины, $z_c = z_0 + \zeta_c = z_0 + r_0 e^{i\theta}$ – значение координаты z на трещине. Внешняя нагрузка в общем случае различна на двух берегах трещины и может быть не самоуравновешенной. Функции $p^*(z_c), q^*(z_c)$ непрерывны и удовлетворяют условию Гёльдера.

На линии сопряжения полуплоскостей $x_2 = 0$ имеем условия непрерывности напряжений и производных перемещений

$$(\sigma_{22} - i\sigma_{21})^+(x_1) = (\sigma_{22} - i\sigma_{21})^-(x_1),$$

$$(u'_1 + iu'_2)^+(x_1) = (u'_1 + iu'_2)^-(x_1),$$

здесь штрих означает производную по переменной x_1 .

Решение краевой задачи строится в виде суммы решений двух задач: первой является задача для двухкомпонентной плоскости без трещины со скачками напряжений и перемещений на линии сопряжения полуплоскостей, второй – задача для однородной плоскости с криволинейной трещиной [1], [4].

Граничные условия первой задачи на линии раздела

$$[(\sigma_{22} - i\sigma_{21})_1]^+ - [(\sigma_{22} - i\sigma_{21})_1]^- = \Delta\sigma(x_1),$$

$$[(u'_1 + iu'_2)_1]^+ - [(u'_1 + iu'_2)_1]^- = \Delta u'(x_1).$$

Граничные условия на трещине во второй задаче

$$[(\sigma_{rr} - i\sigma_{r\theta})_2]^+ = p(z_c), \quad [(\sigma_{rr} - i\sigma_{r\theta})_2]^- = q(z_c).$$

Функции скачков напряжений $\Delta\sigma(x_1)$ и производных перемещений $\Delta u'(x_1)$ первой задачи и функции внешней нагрузки на берегах трещины второй задачи $p(z_c)$ и $q(z_c)$ являются искомыми.

В первой задаче напряжения и производные перемещения выражаются через комплексные потенциалы по формулам Колосова – Мусхелишвили [2]

$$(\sigma_{22} - i\sigma_{21})_1 = \Phi_k(z) + \overline{\Phi_k}(z) + z\overline{\Phi'_k}(z) + \overline{\Psi_k}(z),$$

$$(\sigma_{11} + i\sigma_{12})_1 = \Phi_k(z) + \overline{\Phi_k}(z) - z\overline{\Phi'_k}(z) - \overline{\Psi_k}(z),$$

$$2\mu_k(u'_1 + iu'_2)_1 = \kappa_k\Phi_k(z) - \overline{\Phi_k}(z) - z\overline{\Phi'_k}(z) - \overline{\Psi_k}(z).$$

Комплексные потенциалы можно записать через две комплексные функции $h(z)$, $r(z)$

$$\Phi_2(z) = \frac{r(z) + \mu_2 h(z)}{\mu_2 + \mu_1 \kappa_2}, \quad \Omega_1(z) = \frac{r(z) - \mu_1 \kappa_2 h(z)}{\mu_2 + \mu_1 \kappa_2}, \quad z \in S_2,$$

$$\Phi_1(z) = \frac{r(z) + \mu_1 h(z)}{\mu_1 + \mu_2 \kappa_1}, \quad \Omega_2(z) = \frac{r(z) - \mu_2 \kappa_1 h(z)}{\mu_1 + \mu_2 \kappa_1}, \quad z \in S_1,$$

Функция $\Omega(z)$ выражается через функции $\Phi(z)$ и $\Psi(z)$ [2].

Сами комплексные функции $h(z)$, $r(z)$ находятся из решения граничных задач Римана – Гильберта для первой задачи

$$h(z) = \frac{1}{2\pi i} \int_{-\infty}^{\infty} \frac{\Delta\sigma(x_1)}{x_1 - z} dx_1 + h(\infty),$$

$$r(z) = \frac{\mu_1 \mu_2}{\pi i} \int_{-\infty}^{\infty} \frac{\Delta u'(x_1)}{x_1 - z} dx_1 + r(\infty).$$

Постоянные $h(\infty)$, $r(\infty)$ находятся из граничных условий на бесконечности

$$h(\infty) = \frac{1}{4}(\sigma_{11}^{1\infty} + \sigma_{11}^{2\infty} + \sigma_{22}^{\infty}) - (\sigma_{22}^{\infty} - i\sigma_{21}^{\infty}),$$

$$r(\infty) = \frac{1}{4}\mu_2 \kappa_1 (\sigma_{11}^{1\infty} + \sigma_{22}^{\infty}) - \frac{1}{4}\mu_1 (\sigma_{11}^{2\infty} + \sigma_{22}^{\infty}) + \mu_1 (\sigma_{22}^{\infty} - i\sigma_{21}^{\infty}).$$

Во второй задаче используются полярные координаты (r, θ) , функции Колосова – Мусхелишвили имеют вид

$$s_{rr} + is_{r\theta} = \Phi(\zeta) + \overline{\Phi}(\zeta) - \overline{\zeta\Phi'(\zeta)} - \frac{\overline{\zeta}}{\zeta}\overline{\Psi(\zeta)},$$

$$s_{\theta\theta} - is_{\theta r} = \Phi(\zeta) + \overline{\Phi}(\zeta) + \overline{\zeta\Phi'(\zeta)} + \frac{\overline{\zeta}}{\zeta}\overline{\Psi(\zeta)},$$

$$2\mu(v_r + iv_\theta) = [\kappa\varphi(\zeta) - \zeta\overline{\varphi'(\zeta)} - \overline{\psi(\zeta)}]e^{-i\theta}.$$

Если внешняя нагрузка на берегах трещины одинакова $p^*(z_c) = q^*(z_c)$, то $p(z_c) = q(z_c)$ и формулы для комплексных потенциалов запишутся так

$$\Phi(\zeta) = \Omega(\zeta) = \frac{X(\zeta)}{4\pi i} \int_L \frac{(p+q)(t)dt}{X^+(t)(t-\zeta)}, \quad (1)$$

где $X(\zeta) = \frac{1}{\sqrt{(\zeta-a)(\zeta-b)}}$, $a = r_0 e^{i\alpha}$, $b = r_0 e^{i\beta}$.

Вывод интегрального уравнения. Путем преобразования граничных условий на линии раздела и на трещине получим интегральное уравнение Фредгольма второго рода относительно одной искомой функции $(p+q)(\theta)$ на промежутке $\alpha \leq \theta \leq \beta$ [1]

$$a_1[\Phi(\zeta_1) + \overline{\Phi}(\zeta_1) + \overline{\zeta\Phi'(\zeta_1)} + \zeta\overline{\Phi'(\zeta_1)} + \Psi(\zeta_1) + \overline{\Psi}(\zeta_1)] +$$

$$+ [a_1(\Phi(\zeta_1) + \overline{\zeta\Phi'(\zeta_1)} + \Psi(\zeta_1)) + a_1(z_c - \overline{z_c})(2\Phi'(\zeta_1) + \overline{\zeta\Phi''(\zeta_1)} +$$

$$+ \Psi'(\zeta_1)) - a_2\Phi(\zeta_1)]e^{2i\theta} = -0,5[(p+q) - (p^* + q^*)](z_c) -$$

$$- \Phi_1(\infty) - \overline{\Phi_1(\infty)} - [\overline{\Phi_1(\infty)} - \Omega_1(\infty)]e^{-2i\theta}. \quad (2)$$

Постоянные $a_1 = \frac{\mu_2 - \mu_1}{\mu_1 + \mu_2\kappa_1}$, $a_2 = \frac{\mu_1\kappa_2 - \mu_2\kappa_1}{\mu_2 + \mu_1\kappa_2}$.

Потенциал $\Psi(\zeta)$ вычисляется по формуле

$$\Psi(\zeta) = \frac{r_0^2}{\zeta^2}\Phi(\zeta) - \frac{r_0^2}{\zeta}\Phi'(\zeta) - \frac{r_0^2}{\zeta^2}\overline{\Omega}\left(\frac{r_0^2}{\zeta}\right).$$

Решение интегрального уравнения. Уравнение (2) имеет единственное решение, которое непрерывно на рассматриваемом промежутке. Разобьем на части отрезок $[\alpha, \beta]$ и на каждом частичном промежутке $[\theta_{k-1}, \theta_k]$, $k = \overline{1, n}$, искомую нагрузку представим

$$(p+q)(\theta) = c_k, \quad c_k = \text{const.}$$

Используя такое представление, можно записать выражение (1) в следующем виде

$$\Phi(\zeta) = \frac{1}{4\sqrt{(\zeta-a)(\zeta-b)}} \sum_{k=1}^n c_k \int_{t_{k-1}}^{t_k} \frac{\sqrt{(t-a)(t-b)}}{t-\zeta} dt. \quad (3)$$

Результаты работы. Для вычисления коэффициентов c_k с помощью уравнения (2) предполагается использовать метод коллокаций. После нахождения коэффициентов находятся функции $\Phi(z)$, $\Psi(z)$ и интересующие нас характеристики.

Практическое значение представляют формулы для нахождения напряжений на линии раздела полуплоскостей [1]

$$\begin{aligned} (\sigma_{22} - i\sigma_{21})_1(x_1) = & (\sigma_{22}^\infty - i\sigma_{21}^\infty) + \frac{\mu_2(1 + \kappa_1)}{\mu_2 + \mu_1\kappa_2} \Phi(\zeta_0) + \\ & + \frac{\mu_2(1 + \kappa_1)}{\mu_1 + \mu_2\kappa_1} \left(\overline{\Phi(\zeta_0)} + \zeta_0 \overline{\Phi'(\zeta_0)} + \overline{\Psi(\zeta_0)} \right). \end{aligned} \quad (4)$$

Выполнены расчёты для частного случая: прямолинейной трещины вблизи линии раздела полуплоскостей при наличии постоянного давления на плоскости и отсутствии напряжений на бесконечности.

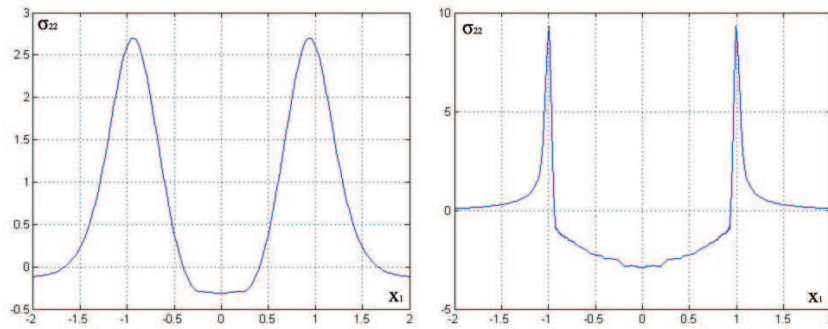


Рис. 2. Нормальные напряжения на линии раздела при $p = q = \text{const}$, $d = 0,5$ (слева), $d = 0,05$ (справа)

Используя метод коллокации, по формуле (4) построены графики функций напряжения на линии раздела полуплоскостей при условиях: $\mu_2/\mu_1 = 3$; $\kappa_1 = \kappa_2 = 1,8$; $\nu = 0,3$; $d = 0,5$; $d = 0,05$, где d – расстояние от трещины до линии раздела (рис. 2, 3).

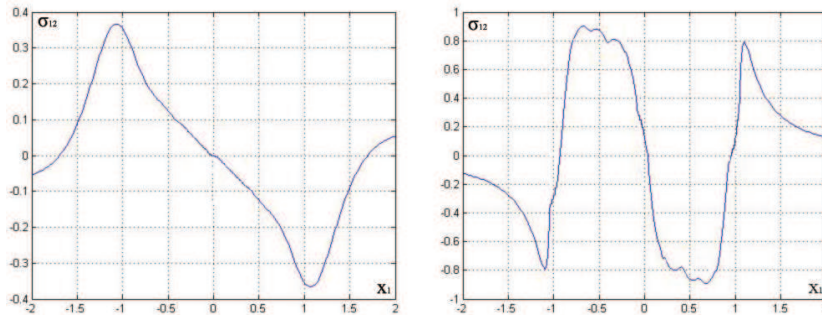


Рис. 3. Касательные напряжения на линии раздела при $p = q = \text{const}$, $d = 0,5$ (слева), $d = 0,05$ (справа)

Заключение. В работе рассмотрена краевая задача теории упругости для бесконечной упругой плоскости, содержащей криволинейную трещину в окрестности линии разделения материалов. Основное внимание уделено решению интегрального уравнения Фредгольма. Рассмотрены разные варианты параметров трещины и материалов. Разработана программа в среде MATLAB, для построения графиков напряжений на линии раздела полуплоскостей в зависимости от различных параметров задачи. Построены графики напряжений для прямолинейной трещины, позволяющие оценить влияние отношения модулей сдвига материалов на напряжения.

Литература

1. Мальков В.М., Малькова Ю.В. Трещина в форме дуги окружности, расположенная вблизи поверхности раздела материалов // Вестник СПбГУ. Сер. 1, 2010. Вып. 1. С. 152–166.
2. Мусхелишвили Н.И. Некоторые основные задачи математической теории упругости. М.: Наука, 1966. 708 с.
3. Малькова Ю.В. Некоторые задачи для двухкомпонентной плоскости с криволинейными трещинами. СПб.: Изд-во СПбГУ, 2008. 160 с.
4. Греков М.А. Сингулярная плоская задача теории упругости. СПб.: СПбГУ. 2001. 192 с.

Климаков А.А.

Санкт-Петербургский государственный университет

Моделирование электростатической линзы

Рекомендовано к публикации профессором Егоровым Н.В.

Введение. Развитие теории источников электронов и ионов тесно связано с общим развитием фундаментальных и прикладных наук: математического моделирования, физики поверхности, ускорителей, нанoeлектроники и т. д. Как известно, источники электронов на основе полевой электронной эмиссии по своим основным характеристикам превосходят широко применяемые на практике источники на основе термоэлектронной эмиссии [1].

Основные отличия полевых катодов от наиболее широко применяющихся сегодня термокатодов заключаются в следующем: более узкий энергетический спектр полевых электронов, малые размеры эмитирующей области, безынерционность, компактность, экономичность (отсутствие расхода энергии на принудительный нагрев). Некоторые из вышеперечисленных характеристик не являются свойством полевой эмиссии как физического феномена, а обусловлены техническими приемами, применяемыми для создания условий, необходимых для ее возникновения. А именно, необходимые значения напряженности поля в большинстве практических случаев возможно получить, лишь придав эмиттеру форму острия [2].

Явление полевой электронной эмиссии может использоваться для создания широкого круга приборов и устройств. В первую очередь, это источники электронов, применяемые в электронных микроскопах, плоских дисплеях на основе полевой эмиссии, системах диагностики поверхности, высокочастотных радио-передающих системах, приборах микро- и нанoeлектроники [3].

При проектировании различных устройств возникает необходимость расчета электростатического поля проводников различной конфигурации. В данной работе рассматривается электростатическая задача для электронной пушки на основе полевого острия [4].

В аналитических моделях полевых катодов в большинстве своем используются системы с границами, параллельными координатным осям. Однако, для практических моделей интересны системы с коугольными элементами.

Постановка задачи. Цель работы – проверка возможности моделирования распределения электростатического поля электронно-оптической системы со сложной геометрией аналитическими методами (методом парных уравнений).

Для этого требовалось найти распределение потенциала методами парных уравнений и конечных элементов. Метод конечных элементов был взят как основной, поскольку корректность распределения потенциала согласно этому методу была проверена ранее [5]. В итоге будет проведено сравнение результатов методов.

Физическая модель. Для моделирования потенциала используем уравнение Лапласа, для сокращения объема вычислений – осевую симметрию системы. Расчеты будем проводить в цилиндрической системе координат.

Для нахождения потенциала методом парных уравнений была взята геометрия системы учета без острия (рис. 1).

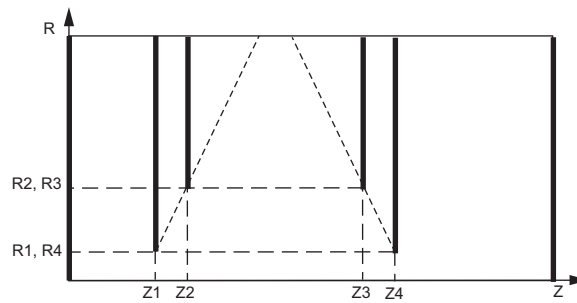


Рис. 1. Схематическое изображение электронно-оптической системы для МПУ

Метод конечных элементов. Для моделирования потенциала используем уравнение Лапласа (Пуассона, $f := 0$), решение которого численно аппроксимируем

$$-\Delta u = f, \quad u \in \Omega \in R^3 \quad (1)$$

с граничными условиями:

1. Дирихле $u = u_D, u \in \Gamma_D$.
2. Неймана $\vec{n}(\nabla u) = g, u \in \Gamma_N$,

где $f \in L_2(\Omega)$, $u_D \in H^1(\Omega)$, $g \in L_2(\Gamma_N)$. Слабая формулировка может быть получена путем умножения уравнения (1) на тестовую функцию $w \in L_2(\Omega) = \{w \in L_2 | w = 0, w \in \Gamma_D\}$ и интегрирования по Ω :

$$-\int_{\Omega} w u dv = \int_{\Omega} w f dv.$$

Интегрирование по частям дает

$$\int_{\Omega} \nabla w \nabla u dv - \int_{\Gamma_N} w \frac{\partial u}{\partial n} da = \int_{\Omega} w f dv.$$

Подставив граничные условия, получим

$$\int_{\Omega} \nabla w \nabla u dv = \int_{\Omega} w f dv + \int_{\Gamma_N} w g da.$$

Включив граничные условия и заменив однородным решением $v \in L_2$ таким, что $v = u - u_D$, $v = 0$ ($v \in \Gamma_D$), для нахождения слабой формулировки уравнения Пуассона, необходимо определить такую функцию $v \in L_2$, чтобы выполнялось:

$$\int_{\Omega} (\nabla v + \frac{\partial v}{\partial \rho}) \nabla w dv = \int_{\Gamma_N} g w da + \int_{\Omega} \nabla u_D \nabla w dv.$$

Метод парных уравнений. Для решения задачи нахождения потенциала методом парных уравнений рассмотрим систему

$$\begin{aligned} \Delta V(r, z) &= 0, \\ V(r, \pm Z_i) \Big|_{r \geq R_i} &= \pm U_i, V(r, z) = -V(r, -z). \end{aligned}$$

Решение представимо в следующем виде

$$\begin{aligned} V(r, z) &= U_i + \frac{U_{i+1} - U_i}{Z_{i+1} - Z_i} (z - Z_i) + \\ &+ \int_0^{\infty} \left[\frac{\text{sh } \lambda(Z_{i+1} - z)}{\text{sh } \lambda(Z_{i+1} - Z_i)} F_i(\lambda) + \frac{\text{sh } \lambda(z - Z_i)}{\text{sh } \lambda(Z_{i+1} - Z_i)} F_{i+1}(\lambda) \right] J_0(\lambda r) d\lambda, \\ &Z_i < z < Z_{i+1}, \end{aligned}$$

$$F_i(\lambda) = \int_0^{R_i} \phi_i(t) \sin \lambda t dt.$$

После преобразований приходим к системе интегральных уравнений Фредгольма второго рода, решив которую, получим распределение потенциала:

$$\begin{aligned} & -\frac{1}{\pi} \int_0^{R_{i-1}} K_{i-1}(x, t) \phi_{i-1}(t) dt + \\ & + 2\phi_i(x) + \frac{1}{\pi} \int_0^{R_i} K_{ii}(x, t) \phi_i(t) dt - \\ & - \frac{1}{\pi} \int_0^{R_{i+1}} K_{i+1i}(x, t) \phi_{i+1}(t) dt = \Phi_i(x), \end{aligned}$$

$$\Phi_i(x) = \frac{2}{\pi} \int_0^x \frac{y V_i(y)}{\sqrt{x^2 - y^2}} dy,$$

$$K_{i-1}(x, t) = -2 \int_0^\infty \frac{1}{\operatorname{sh} \lambda(Z_i - Z_{i-1})} \sin \lambda x \sin \lambda t d\lambda,$$

$$\begin{aligned} K_{ii}(x, t) = & -2 \int_0^\infty \left[\frac{\exp(-\lambda(Z_{i+1} - Z_i))}{\operatorname{sh} \lambda(Z_{i+1} - Z_i)} + \right. \\ & \left. + \frac{\exp(-\lambda(Z_i - Z_{i-1}))}{\operatorname{sh} \lambda(Z_i - Z_{i-1})} \right] \sin \lambda x \sin \lambda t d\lambda. \end{aligned}$$

Результаты. Найдено распределение потенциала методами парных уравнений и конечных элементов, произведено сравнение результатов методов.

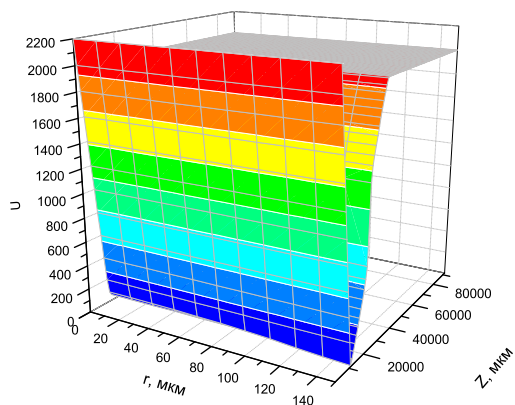


Рис. 2. Распределение потенциала при моделировании фокусирующих линз четырьмя диафрагмами

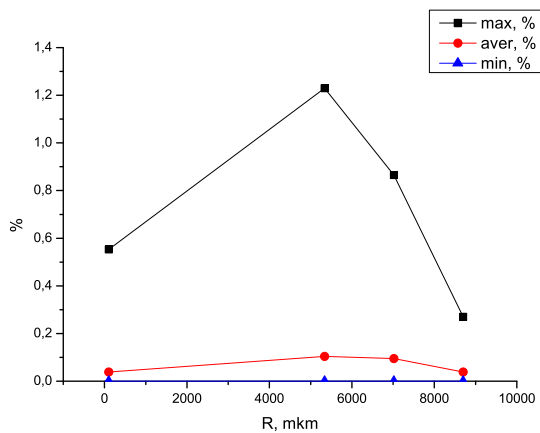


Рис. 3. Максимальная, минимальная и средняя разности между распределениями потенциала при изменении положения второй и третьей диафрагмы

На рис. 2 представлено распределение потенциала при моделировании фокусирующих линз четырьмя диафрагмами по методу парных уравнений.

На рис. 3 указана максимальная, минимальная и средняя разности между распределениями потенциала при изменении положения второй и третьей диафрагмы.

Согласно полученным результатам моделирование электростатического потенциала выбранной системы можно производить МПУ с допустимой погрешностью.

Литература

1. Ильин В.П. Алгоритмы и методы расчета электронно-оптических систем. Новосибирск: ИТМ, 1983. 215 с.
2. Самарский А.А., Гулин А.В. Численные методы математической физики. М.: Научный мир, 2003. 310 с.
3. Уфлянд Я.С. Метод парных уравнений в задачах математической физике. Л.: Наука, 1977. 285 с.
4. Соболев С.Л. Уравнения математической физики. М.: Наука, 1966. 165 с.
5. Шушкевич Г.Ч., Электростатическое поле тонкой незамкнутой сферической оболочки и тора // Журнал технической физики. 1998. Т. 68, № 7. С. 177–179.
6. Chevailler J.P. // J. Electrostatics. 1981. Vol. 9, № 2. P. 307–313.
7. Edgcombe C. J. Experimental, computational study of field emission characteristics // Philosophical Magazine. 2002. Part B, Vol. 82(9): 987–1007. P. 135–137.

Кодрик В.В., Кутузов С.А.
Санкт-Петербургский государственный университет

Моделирование внешнего поля Земли с помощью гравитирующих стержней

Введение. Моделирование внешнего гравитационного поля Земли важно как в познавательном аспекте, так и для вычисления траекторий естественных и искусственных тел [1]. Простейшую модель центрального поля шара построил Ньютон. Учет несферичности Земли привел к использованию разложения потенциала по сферическим функциям (см., например, [1]). Ряд содержит зональные, секториальные и тессеральные гармоники. Значения параметров разложения приближенно определяются по наблюдениям Луны, ИСЗ и по гравиметрическим и геодезическим измерениям. Известный набор значений получил название «Стандартная Земля» [1].

Хорошее приближение дает модель двух неподвижных центров [1]. Обобщением является построение системы из многих материальных точек (см., например, [3]). Кондратьев [4] разработал теорию эквигравитирующих стержней, моделирующих поля пространственных тел.

Для нашей задачи мы предлагаем модель из трех гравитирующих стержней, жестко связанных с Землей. Она позволяет хорошо учесть как сплюснутость Земли, так и ее ротационную асимметрию. Потенциал выражается в конечном виде через элементарные функции.

1. Общее выражение потенциалов стержней. Асимптотика. Введем геоцентрическую декартову систему координат. Ось z направим вдоль оси вращения Земли, ось x – в точку гринвичского меридиана на экваторе, ось y дополнит систему до правой. На каждой из осей расположим по бесконечно тонкому стержню с линейной плотностью $\mu_u(u)$ на симметричном относительно начала отрезке $[-a_u, +a_u]$. Полагаем, что концы отрезков находятся внутри Земли, а распределение масс симметрично относительно начала. Здесь и далее u может заменяться на одно из обозначений x, y, z , указывая, к которому стержню относится формула.

Применяя формулу для гравитирующего отрезка из [5], находим

выражения для трех потенциалов:

$$\Phi_u(x, y, z) = G \int_{-a_u}^{a_u} \frac{\mu_u(u')}{\sqrt{U_u}} du', \quad U_u = R_u^2 + (u' - u)^2, \quad (1)$$

где G – гравитационная постоянная и

$$R_z^2 = x^2 + y^2, \quad R_x^2 = y^2 + z^2, \quad R_y^2 = x^2 + z^2. \quad (2)$$

Поле u -стержня представляет собой тело вращения с осью u . Тотальное поле с потенциалом

$$\Phi(x, y, z) = \Phi_z(R_z, z) + \Phi_x(R_x, x) + \Phi_y(R_y, y), \quad (3)$$

не обладает ротационной симметрией. Для Земли главный вклад дает первое слагаемое, моделирующее сплюснутость вдоль оси вращения. Два других слагаемых описывают ротационную асимметрию и служат небольшими поправками.

Введем также сферические координаты (r, φ, λ) рассматриваемой точки (x, y, z) :

$$x = r \cos \varphi \cos \lambda, \quad y = r \cos \varphi \sin \lambda, \quad z = r \sin \varphi, \quad r^2 = x^2 + y^2 + z^2,$$

так что r есть геоцентрическое расстояние, φ – широта, λ – долгота точки. Рассматривая внешнее поле Земли, следует считать, что расстояние больше ее полярного радиуса.

Разложим потенциалы (1) по степеням r^{-1} . Для этого разделим числитель и знаменатель подынтегрального выражения на r . Раскрывая скобки под корнем, замечаем, что получается производящая функция полиномов Лежандра, не зависящих от переменных интегрирования. Получаем формулу того же вида, который приводит к стандартному разложению [1]:

$$\Phi_u(x, y, z) = G \sum_{n=0}^{\infty} \frac{v_n^{(u)}}{r^{n+1}} P_n \left(\frac{u}{r} \right), \quad (4)$$

$$v_n^{(u)} = \int_{-a_u}^{a_u} \mu_u(s) s^n ds, \quad n = 0, 1, \dots$$

Здесь $\nu_n^{(u)}$ – начальные моменты распределений масс, а $\nu_0^{(u)} = m_u$ – масса u -стержня, которую считаем положительной. Аргументы полиномов Лежандра очевидным образом выражаются через широту и долготу. Первый член ряда описывает асимптотику потенциалов. На бесконечности стержень притягивает как материальная точка. Задавая распределение масс $\mu_z(z | \mathbf{p})$ выражениями с N свободными параметрами, $\mathbf{p} = (p_1, \dots, p_N)$, можно в принципе получить $N - 1$ зональных гармоник в стандартном разложении потенциала.

С помощью теоремы сложения для полиномов Лежандра формулы (4) могут быть преобразованы к стандартному виду, содержащему присоединенные функции Лежандра.

2. Потенциал полиномиального распределения масс. Зададим линейную плотность u -стержня $\mu_u(u)$ в виде полинома

$$\mu_u(u) = \sum_{k=0}^{K_u} p_k^{(u)} u^k. \quad (5)$$

Задавая $\mu_u(u)$ и вычисляя моменты распределения $\nu_n^{(u)}$ по формуле (4), получаем систему линейных алгебраических уравнений

$$\sum_{k=0}^{K_u} A_{ik}^{(u)} p_k^{(u)} = \nu_i^{(u)}, \quad i = \overline{0, K_u}, \quad A_{ik}^{(u)} = \frac{a_u^{i+k+1} - (-a_u)^{i+k+1}}{i+k+1}. \quad (6)$$

Для четного $K_u = 2n$ и $a_u = 1$ система уравнений принимает вид

$$2 \begin{pmatrix} 1 & 0 & \frac{1}{3} & 0 & \frac{1}{5} & \cdots & \frac{1}{n+1} \\ 0 & \frac{1}{3} & 0 & \frac{1}{5} & 0 & \cdots & \frac{1}{n+2} \\ \frac{1}{3} & 0 & \frac{1}{5} & 0 & \frac{1}{7} & \cdots & \frac{1}{n+3} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{1}{n+1} & 0 & \frac{1}{n+3} & 0 & \frac{1}{n+5} & \cdots & \frac{1}{2n+1} \end{pmatrix} \begin{pmatrix} p_0^{(u)} \\ p_1^{(u)} \\ p_2^{(u)} \\ \vdots \\ p_n^{(u)} \end{pmatrix} = \begin{pmatrix} \nu_0^{(u)} \\ 0 \\ \nu_2^{(u)} \\ \vdots \\ \nu_{2n}^{(u)} \end{pmatrix}. \quad (7)$$

Разрешая систему (7), получим искомые коэффициенты полинома $p_0^{(u)}, \dots, p_n^{(u)}$. Заметим, что матрица системы для больших K_u является плохо обусловленной. Тогда даже малому изменению моментов будут соответствовать большие изменения коэффициентов. Эмпирические значения моментов приведены в [1].

Для нахождения потенциала подставляем полином (5) в интеграл (1). Это дает выражение, в котором интегралы берутся подстановками Эйлера:

$$\Phi_u(x, y, z) = G \sum_{k=0}^{K_u} p_k \chi_k^{(u)}(x, y, z), \quad \chi_k^{(u)}(x, y, z) = \int_{-a_u}^{a_u} \frac{u'^k}{\sqrt{U_u}} du'. \quad (8)$$

3. Потенциал «зубчатого» распределения масс. Построим кусочно-линейные распределения масс, графики которых имеют пилообразный вид. Расположим $2l_u + 1$ зубцов симметрично относительно начала вдоль u -стержня. Зубцы образуют равнобедренные треугольники с основаниями $2h_k^{(u)}$ и высотой $1/h_k^{(u)}$, где k – номер зубца. Зубцу приписываем массу $m_k^{(u)}$. Так как площадь треугольника всегда равна единице, то при $h_k^{(u)} \rightarrow 0$ получаем дельта-функцию, которая отображает точечную массу величиной $m_k^{(u)}$. Положение материальной точки совпадает с координатой вершины треугольника u_k . Функция зубца имеет вид

$$f_k^{(u)}(u) = \left(h_k^{(u)}\right)^{-2} \begin{cases} h_k^{(u)} - |u - u_k|, & u \in [u_k - h_k^{(u)}, u_k + h_k^{(u)}], \\ 0, & u \notin [u_k - h_k^{(u)}, u_k + h_k^{(u)}]. \end{cases} \quad (9)$$

Считаем, что зубцы не пересекаются друг с другом, а крайние зубцы не выходят за интервал $[-a_u, a_u]$.

Распределения масс и потенциалы стержней получают вид

$$\mu_u(u) = \sum_{k=-l}^l m_k^{(u)} f_k^{(u)}(u), \quad \sum_{k=-l}^l m_k^{(u)} = m_u, \quad (10)$$

$$\Phi_u(x, y, z) = G \sum_{k=-l}^l m_k^{(u)} \psi_k^{(u)}(x, y, z).$$

Удельные потенциалы зубцов $\psi_k^{(u)}$ получаются по формуле (1) с заменой пределов интегрирования на $b_k^{(u)} = u_k - h_k^{(u)}$ и $c_k^{(u)} = u_k + h_k^{(u)}$. Интегрирование дает выражение в элементарных функциях. Величины $a_u, m_u, m_k^{(u)}, u_k, h_k^{(u)}$ являются параметрами модели. Зубчатое распределение охватывает широкий класс моделей и включает в качестве частных случаев одну, две и более материальных точек на

координатных осях. Увеличивая ширину зубцов, можно проследить эффект «размазывания» точек по отрезкам.

4. Эквипотенциали для различных распределений масс. Эквипотенциали, т.е. поверхности уровня потенциала, представляют собой средство наглядного описания поля. Напряженности поля, являющиеся градиентами потенциала, направлены по нормальям к эквипотенциалам. Здесь демонстрируется несколько эквипотен-

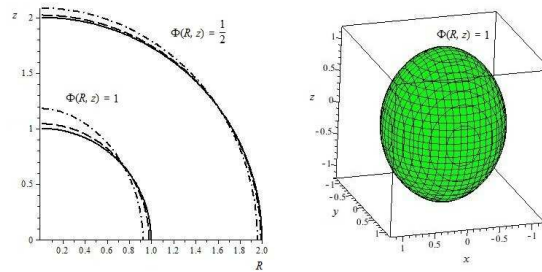


Рис. 1. Эквипотенциали z -стержня

циалей для простых случаев. Для наглядности значимость x - и y -стержней сильно преувеличена. Рассмотрим z -стержень с единичной массой. Распределение линейной плотности на стержне моделируется одним зубцом. На рис. 1 слева изображены сечения эквипотенциалей плоскостью, проходящей через ось z . Значениям

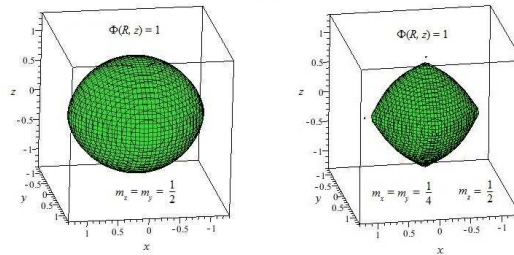


Рис. 2. Эквипотенциали x - и y -стержней (слева) и тотального потенциала (справа)

$h_0^{(z)} = 1$, $h_0^{(z)} = \frac{1}{2}$, $h_0^{(z)} = \frac{1}{16}$ соответствуют штрих-пунктирная, пунк-

тирная и сплошная линии. Видно, что при уменьшении $h_0^{(z)}$ различие между эквипотенциальными поверхностями существенно уменьшается. На рис. 2 слева изображены эквипотенциали стержней, которые используются для моделирования ротационной асимметрии потенциала. Справа изображена эквипотенциаль системы трех стержней с длиной, равной двум. Распределение масс каждого стержня моделируется одним зубцом.

Заключение. Предлагаемая модель внешнего гравитационного поля Земли является довольно гибкой благодаря наличию большого числа параметров. Хотя количество получаемых зональных гармоник может быть сколь угодно большим, часть секториальных и тессеральных гармоник не может быть получена в принципе. Их можно получить вращением экваториальных стержней относительно друг друга. Другой подход состоит в том, чтобы, не требуя точного совпадения, строить средне-квадратическую аппроксимацию. Тогда значения параметров модели получатся в результате минимизации функционала согласия моделей.

Модели такого рода можно строить и для других планет Солнечной системы. Эти модели интересны и для теоретических исследований хаотических орбит пробных материальных точек.

Литература

1. Аксенов Е.П. Теория движения искусственных спутников Земли. М.: Наука, 1977. 360 с.
2. Субботин М.Ф. Курс небесной механики. Т. 3. Л.–М.: Гостехиздат, 1949. 280 с.
3. Антонов В.А., Тимошкова Е.И., Холшевников К.В. Введение в теорию ньютоновского потенциала. М.: Наука, 1988. 272 с.
4. Кондратьев Б.П. Теория потенциала. Новые методы и задачи с решениями. М.: Мир, 2007. 512 с.
5. Антонов В.А., Никифоров И.И., Холшевников К.В. Элементы теории гравитационного потенциала и некоторые случаи его явного выражения. СПб.: Изд-во СПбГУ, 2008. 208 с.

Котов П.А.

Воронежская государственная технологическая академия

О динамических погрешностях диссипативных динамических систем

Рассматриваются варианты динамических моделей диссипативных систем представимых вещественными дифференциальными уравнениями второго порядка с постоянными элементами коэффициентной системы, скалярным возмущающим воздействием и предлагаются конструктивные основы устойчивости динамических систем.

1. Предварительные соображения. Динамическими погрешностями первого рода называют разность между решениями уравнения

$$m\ddot{x} + B\dot{x} + Sx \pm F_{mp} = f_1(t)$$

и уравнения

$$Sx \pm F_{mp} = f_1(t),$$

где под $f_1(t)$ подразумевается сила, подлежащая измерению в данном приборе [1] (для упрощений силой кулоновского трения F_{mp} пренебрегают). При таком рассмотрении динамическая погрешность первого рода представляет собой разность между решением уравнения

$$m\ddot{x} + B\dot{x} + Sx = f(t) \tag{1}$$

для случая, когда $\frac{S}{m} > \left(\frac{B}{2m}\right)^2$, формируемого в виде

$$x = \frac{\exp -ht}{\sqrt{k^2 - h^2}} \left\{ \sin \sqrt{k^2 - h^2}t \int \exp ht \frac{f(t)}{m} \cos \sqrt{k^2 - h^2}tdt \right\} - \frac{\exp -ht}{\sqrt{k^2 - h^2}} \left\{ \cos \sqrt{k^2 - h^2}t \int \exp ht \frac{f(t)}{m} \sin \sqrt{k^2 - h^2}tdt \right\},$$

где $k^2 = \frac{S}{m}$, $h = \frac{B}{2m}$ и выражением так называемых «статических» отклонений точки приведения всех сил в механизме, которое равно

$\frac{f(t)}{S}$. Эта разность разбивается на два слагаемых, а именно: на разность между положениями, определяемыми «вынужденным» движением и «статическим» отклонением точки, и на отклонение вследствие собственных колебаний системы.

В случае гармонических функций

$$f_1(t) = A \sin(\omega t + \varepsilon),$$

отклонение x_B точки вследствие вынужденного движения будет таким:

$$x_B = \frac{A}{S} \lambda \sin(\omega t + \varepsilon - \delta),$$

где λ – коэффициент динамической восприимчивости системы; δ – сдвиг фаз между вынужденными и возмущающими колебаниями; Отношение амплитуды вынужденных колебаний к амплитуде статических отклонений называется коэффициентом динамической восприимчивости системы и обозначается λ . Амплитуда статических отклонений равна AS^{-1} [1].

Итак, пусть $f(t) = A \sin(\omega t + \varepsilon)$; A – амплитуда колебаний возмущающей силы; ε – какая-то постоянная безразмерная угловая величина.

Тогда решение уравнения (1) будет

$$x = \frac{A}{m\sqrt{(k^2 - \omega^2) + 4\omega^2 h^2}} \sin(\omega t + \varepsilon - \delta) + \exp -ht \left(C_1 \sin \sqrt{k^2 - h^2}t + C_2 \cos \sqrt{k^2 - h^2}t \right),$$

где δ – постоянная величина, равная

$$\delta = \arctg \frac{2\omega h}{k^2 - \omega^2}.$$

Статическое отклонение x_{cm} от точки в данном случае равно:

$$x_{cm} = \frac{A}{S} \sin(\omega t + \varepsilon).$$

Отсюда разность для этих двух отклонений записывается так [1]:

$$\begin{aligned} x_B - x_{cm} &= \frac{A}{S} [\lambda \sin(\omega t + \varepsilon - \delta) - \sin(\omega t + \varepsilon)] = \\ &= \frac{A}{S} (1 + \lambda^2 - 2\lambda \cos \delta)^{\frac{1}{2}} \sin(\omega t + \varepsilon - \mu), \end{aligned}$$

где $\mu = \operatorname{arctg} \frac{\lambda \sin \delta}{\lambda \cos \delta - 1}$.

В моменты, когда $\omega t + \varepsilon - \mu = (2n + 1)\frac{\pi}{2}$, где n – целое число, т.е. когда

$$t = \frac{(2n + 1)\frac{\pi}{2} - \varepsilon + \mu}{\omega},$$

разность между величинами x_B и x_{cm} становится наибольшей и достигает величины

$$x_B - x_{cm} = \frac{A}{S} \sqrt{(1 - \lambda \cos \delta)^2 + \lambda^2 \sin^2 \delta}.$$

Кроме величины $x_B - x_{cm}$, собственные колебания, пока они еще не затухли, дают также какие-то добавочные отклонения, которые обычно трудно учесть. Эти погрешности алгебраически складываются с погрешностью, получающейся вследствие наличия разности между отклонениями x_B и x_{cm} . При этом механизм, имея собственные колебания, не сразу устанавливается, даже когда $f(t)$ есть постоянная величина. Поэтому для погашения собственных колебаний требуется успокоение. Отмечается, что при увеличении степени успокоения собственные колебания затухают быстрее. Собственные колебания системы, как известно, могут быть до тех пор, пока $(\frac{B}{2m})^2 < \frac{S}{m}$. Если же величина $(\frac{B}{2m})^2 \geq \frac{S}{m}$, то имеет место апериодическое собственное движение. При этом система, выведенная из равновесия, в случае когда функция $f(t)$ есть постоянная величина, асимптотически возвращается к равновесию, не переходя через него. Во многих приборах (например, в компасах) наиболее принятым полагать успокоение, когда $(\frac{B}{2m})^2 = 0,5 \frac{S}{m}$ ($h = \frac{B}{2m} \approx 0,707k$).

Актуальны вопросы существования и единственности решения исходных уравнений.

2. Особенности динамических систем при наличии трения. Пусть динамическая система при влиянии кулоновского трения при вынужденных колебаниях представляется уравнением [1]

$$m\ddot{x} + B\dot{x} + Sx - F_{mp} = A \cos \omega t, \quad (2)$$

где ω – круговая частота вибрации, x – происходящее вследствие вибрации отклонение от некоторого среднего положения точки приведения всех сил в механизме (приборе).

Решение уравнения (2) имеет вид

$$x = \frac{A}{S} \lambda \cos(\omega t - \delta) + \frac{F_{mp}}{S} + \exp -h(t - t_1) \times \\ \times \left[C_3 \sin \sqrt{k^2 - h^2} (t - t_1) + C_4 \cos \sqrt{k^2 - h^2} (t - t_1) \right].$$

Откуда

$$\dot{x} = -\frac{A}{S} \lambda \omega \sin(\omega t - \delta) + \\ + \exp -h(t - t_1) \left\{ -h \left[C_3 \sin \sqrt{k^2 - h^2} (t - t_1) + C_4 \cos \sqrt{k^2 - h^2} (t - t_1) \right] \right. \\ \left. + \sqrt{k^2 - h^2} C_3 \cos \sqrt{k^2 - h^2} (t - t_1) - \sqrt{k^2 - h^2} C_4 \sin \sqrt{k^2 - h^2} (t - t_1) \right\}.$$

Далее формируются три равенства:

$$C_3 n \sin \frac{\sqrt{k^2 - h^2}}{\omega} \pi + C_4 \left(1 + n \cos \frac{\sqrt{k^2 - h^2}}{\omega} \pi \right) = -2 \frac{F_{mp}}{S}, \\ -\frac{A}{S} \lambda \omega \sin(\omega t_1 - \delta) - h C_4 + \sqrt{k^2 - h^2} C_3 = 0, \\ \frac{A}{S} \lambda \omega \sin(\omega t_1 - \delta) - h n \left[C_3 \sin \frac{\sqrt{k^2 - h^2}}{\omega} \pi + C_4 \cos \frac{\sqrt{k^2 - h^2}}{\omega} \pi \right] + \\ + \sqrt{k^2 - h^2} n \left[C_3 \cos \frac{\sqrt{k^2 - h^2}}{\omega} \pi - C_4 \sin \frac{\sqrt{k^2 - h^2}}{\omega} \pi \right] = 0, \\ n = \exp -\frac{h}{\omega} \pi.$$

Из этих равенств определяются C_3, C_4, t_1 . Из второго равенства определяется сдвиг фаз $\omega t_1 - \delta$. Далее учитывается, что

$$x_0 = \frac{A}{S} \lambda \cos(\omega t_1 - \delta) + \frac{F_{mp}}{S} + C_4.$$

Заслуживают внимания исследования устойчивости рассматриваемого варианта системы при движении по инерции с удельной массой, записываемого детерминированным уравнением:

$$\frac{d^2x}{dt^2} + B \frac{dx}{dt} + k_0 x - F = 0, \quad F > 0, \quad (3)$$

$$x(t_0) = x_0; \quad \left. \frac{dx}{dt} \right|_{t_0} = x_1, \quad t_0 \neq \infty \quad (4)$$

и вариантом решения:

$$C_1 \exp r_1 t + C_2 \exp r_2 t + k_0^{-1} F, \quad t \geq 0.$$

Возможные соотношения диагностирования функциональных постоянных интегрирования искомого действительного решения детерминированного уравнения предлагаются таким вариантом разработанной схемы вещественных соотношений при $t_0 < \infty$:

$$x_0 - C_1 \exp r_1 t_0 - k_0^{-1} F = C_2 \exp r_2 t_0,$$

$$x_1 + C_1 r_1 \exp r_1 t_0 + x_0 r_2 - C_1 r_2 \exp r_1 t_0 = k_0^{-1} F r_2,$$

$$\frac{x_1 + x_0 r_2 - k_0^{-1} F r_2}{r_2 - r_1} \exp^{-1} r_1 t_0 \neq 0, \quad r_1 t_0 \neq -\infty,$$

$$C_1 \exp r_1 t_0 = \frac{x_1 + x_0 r_2 - k_0^{-1} F r_2}{r_2 - r_1}, \quad x_0 \neq \frac{x_1 + x_0 r_2 - k_0^{-1} F r_2}{r_2 - r_1} + k_0^{-1} F,$$

$$\frac{-x_0 r_1 - x_1 + k_0^{-1} F r_1}{r_2 - r_1} \exp^{-1} r_2 t_0 = C_2.$$

Другие возможные соотношения диагностирования функциональных постоянных интегрирования искомого действительного решения

детерминированного уравнения предлагаются иным вариантом разработанной схемы вещественных соотношений при $t_0 < \infty$:

$$\begin{aligned}
x_0 - C_2 \exp r_2 t_0 - k_0^{-1} F &= C_1 \exp r_1 t_0, \\
x_1 + x_0 r_1 - C_2 r_1 \exp r_2 t_0 + C_2 r_2 \exp r_2 t_0 &= k_0^{-1} F r_1, \\
\frac{x_1 + x_0 r_1 - k_0^{-1} F r_1}{r_1 - r_2} \exp^{-1} r_2 t_0 &\neq 0, \\
\frac{x_1 + x_0 r_1 - k_0^{-1} F r_1}{r_1 - r_2} - C_2 \exp r_2 t_0 &= 0, \quad r_2 t_0 \neq -\infty, \\
x_0 &\neq \frac{x_1 + x_0 r_1 - k_0^{-1} F r_1}{r_1 - r_2} + k_0^{-1} F, \\
\frac{-x_0 r_2 - x_1 + k_0^{-1} F r_2}{r_1 - r_2} &= C_1 \exp r_1 t_0, \quad r_1 t_0 \neq -\infty.
\end{aligned}$$

Развернутый вариант действительного решения исследуемого вещественного детерминированного уравнения разработан так:

$$\begin{aligned}
&\frac{-x_0 r_2 - x_1 + k_0^{-1} F r_2}{r_1 - r_2} \exp r_1 t - t_0 + \\
&+ \frac{x_1 + x_0 r_1 - k_0^{-1} F r_1}{r_1 - r_2} \exp r_2 t - t_0 + k_0^{-1} F.
\end{aligned}$$

Устойчивый вариант [2] динамической диссипативной системы (3), (4) определяется непрерывным решением однородного уравнения при $t \rightarrow \infty$, $N < +\infty$.

Заслуживают внимания вопросы существования и единственности решения рассматриваемого в настоящей работе вещественного уравнения второго порядка с периодическим возмущающим воздействием. Предлагается разработка такого положения.

Лемма. *Для существования единственного измеримого решения детерминированного уравнения второго порядка*

$$\frac{d^2 x}{dt^2} + p \frac{dx}{dt} + qx = f(\omega t), \quad p > 0, \quad q > 0, \quad t \in R_+$$

$$x(t_0) = x_0, \left. \frac{dx}{dt} \right|_{t_0} = x_1$$

с непрерывным периодическим возмущением в правой части и фиксированными детерминированными скалярными начальными значениями координат и скорости требуется выполнение условия:

$$f(\omega t) \in C_{L^2[t_0, t]}, \quad \omega \neq \omega_0,$$

где ω_0 – круговая частота свободных колебаний системы.

Актуальными представляются вопросы формирования директивного управляющего сигнала системы из рассматриваемого класса динамических систем со стационарным фазовым пространством и действительными отрицательными корнями характеристического полинома, обоснования условий реализации оптимальных процессов в системах высокого порядка [3].

Литература

1. Тихменев С.С. Элементы точных приборов. Руководство по расчету и конструированию. М., Оборонгиз, 1956, 356 с.
2. Котов П.А. Математические вопросы динамики стабильных элементарных частиц с фиксированным точечным зарядом // Моделирование систем и процессов. Воронеж, 2010. Вып. 1-2. С. 29.
3. Репников А.В. Колебания в оптимальных системах автоматического регулирования. М., Машиностроение, 1968. 239 с.

Курушина С.Е., Желнов Ю.В, Максимов В.В.
Самарский государственный аэрокосмический университет

**Фазовый переход
«беспорядок – порядок – беспорядок»
в стохастических системах реакция-диффузия**

Системы типа реакция-диффузия имеют широкую область приложений в активных физических, химических, экологических и экономических системах. При этом большинство работ посвящено исследованию свойств этих систем в автономном случае. Однако в реальной ситуации та или иная система, как правило, находится под действием внешних случайных полей.

Учет влияния флуктуирующего окружения на двухкомпонентную распределенную систему реакция-диффузия приводит к стохастическим уравнениям вида

$$\begin{aligned} \partial x_1 / \partial t &= P_1(x_1, x_2, \chi_1, \dots, \chi_n) + \\ &+ f_1(\vec{r}, t) P_2(x_1, x_2, \chi_{l+1}, \dots, \chi_n) + D_1 \nabla^2 x_1, \\ \partial x_2 / \partial t &= Q_1(x_1, x_2, \chi_1, \dots, \chi_m) + \\ &+ f_2(\vec{r}, t) Q_2(x_1, x_2, \chi_{k+1}, \dots, \chi_m) + D_2 \nabla^2 x_2, \end{aligned} \quad (1)$$

где x_1, x_2 – функции состояния системы, χ_1, \dots, χ_n – управляющие параметры системы, D_1, D_2 – коэффициенты диффузии, $f_1(\vec{r}, \tau)$, $f_2(\vec{r}, \tau)$ – стационарные однородные изотропные гауссовы поля с нулевыми средними и корреляционным тензором

$$\langle f_i(\vec{r}', t') f_j(\vec{r}, t) \rangle = F_i(|\vec{r}' - \vec{r}|) \exp(-k_i |t' - t|) \delta_{ij}, \quad i, j = 1, 2.$$

В [1] для системы (1) были получены стохастические уравнения, описывающие эволюцию амплитуд незатухающих мод $\xi_{\vec{k}u}$ (параметров порядка). После усреднения этих уравнений по ансамблю реализаций получена система, описывающая изменение во времени моментов параметров порядка. Она была использована для исследования

эволюции системы (1) в малой окрестности точки бифуркации Тьюринга, в результате которой рождается новое статистически стационарное состояние системы. Сделан вывод о том, что мультипликативные флуктуации параметров приводят к смещению стационарного состояния каждой моды, к изменению собственных чисел задачи, и как следствие, к расширению области неустойчивости, смещению точки бифуркации Тьюринга и изменению скорости образования пространственных структур. При этом интенсивность флуктуаций и характерные пространственный и временной масштабы внешнего случайного поля становятся бифуркационными параметрами.

Однако полученные уравнения не позволяют исследовать поведение системы (1) вдали от точки бифуркации. В этом случае система (1) может быть исследована численно, либо может быть получено и исследовано уравнение Колмогорова – Фоккера – Планка для параметров порядка. В данной работе решается последняя задача.

Для вывода уравнения Колмогорова – Фоккера – Планка в уравнениях для параметров порядка, полученных в [1], ограничимся слагаемыми, имеющими третий порядок малости. Тогда эти уравнения примут вид

$$\begin{aligned}
& \frac{\partial \xi_{\vec{k}_u}}{\partial t} = \lambda_1(\vec{k}_u) \xi_{\vec{k}_u} + \\
& + \sum_{\vec{k}'_u, \vec{k}''_u} \sigma_{11}^{(1)}(\vec{k}_u, \vec{k}'_u, \vec{k}''_u) \xi_{\vec{k}'_u} \xi_{\vec{k}''_u} \delta(\vec{k}'_u + \vec{k}''_u - \vec{k}_u) + \\
& + \sum_{\vec{k}'_u, \vec{k}''_u, \vec{k}'''_u} w(\vec{k}_u, \vec{k}'_u, \vec{k}''_u, \vec{k}'''_u, \vec{k}_s) \xi_{\vec{k}'_u} \xi_{\vec{k}''_u} \xi_{\vec{k}'''_u} \delta(\vec{k}'_u + \vec{k}''_u + \vec{k}'''_u - \vec{k}_u) + \\
& + \sum_{\varphi=1}^2 O_{\varphi}^{*(1)}(\vec{k}_u) P_{\varphi}^{(0)} f_{\varphi, \vec{k}_u}(t) + \\
& + \sum_{\varphi, \vec{k}'_u} \eta_{\varphi}(\vec{k}_u, \vec{k}'_u) \xi_{\vec{k}'_u} f_{\varphi, \vec{k}_u - \vec{k}'_u}(t) = F_{\vec{k}_u}(t).
\end{aligned} \tag{2}$$

Функции

$$\lambda_1(\vec{k}_u), \quad \sigma_{11}^{(1)}(\vec{k}_u, \vec{k}'_u, \vec{k}''_u), \quad w(\vec{k}_u, \vec{k}'_u, \vec{k}''_u, \vec{k}'''_u, \vec{k}_s),$$

$$O_\varphi^{*(1)}(\vec{k}_u)P_\varphi^{(0)}, \quad \eta_\varphi(\vec{k}_u, \vec{k}'_u)$$

в виду громоздкости здесь не представлены и описаны в [1]. Процессы

$$f_{\varepsilon, \vec{k}}(t) = \int f_\varepsilon(\vec{r}, t) e^{-i\vec{k}\vec{r}} d\vec{r}$$

– компоненты случайного векторного поля $\vec{f}(t)$, имеющие нулевые средние, ε и \vec{k} – индексные аргументы этого поля. Корреляционные функции для компонент поля $\vec{f}(t)$ будут иметь вид

$$\langle f_{j, \vec{k}}(t) f_{l, \vec{k}'}(\tau) \rangle = F_j^{kor}(|\vec{k}|) \delta(\vec{k} - \vec{k}') \exp(-k_t|t - \tau|) \delta_{jl}.$$

Уравнение Колмогорова – Фоккера – Планка для системы (2) может быть представлено в общем виде [2]

$$\begin{aligned} \partial\omega/\partial t = & - \sum_{\vec{k}_u} \frac{\partial}{\partial \xi_{\vec{k}_u}} \left\{ \left(\langle F_{\vec{k}_u} \rangle + \right. \right. \\ & \left. \left. + \sum_{\vec{q}} \int_{-\infty}^0 K \left[\frac{\partial F_{\vec{k}_u}(t)}{\partial \xi_{\vec{q}}} , F_{\vec{q}}(\tau) \right] d\tau \right) \omega \right\} + \\ & + \sum_{\vec{k}_u, \vec{q}} \frac{\partial^2}{\partial \xi_{\vec{k}_u} \partial \xi_{\vec{q}}} \left\{ \left(\int_{-\infty}^0 K \left[F_{\vec{k}_u}(t), F_{\vec{q}}(\tau) \right] d\tau \right) \omega \right\}, \end{aligned} \quad (3)$$

где использовано обозначение

$$K[F_{\vec{k}_u}(t), F_{\vec{q}}(\tau)] = \langle F_{\vec{k}_u}(t) F_{\vec{q}}(\tau) \rangle - \langle F_{\vec{k}_u}(t) \rangle \langle F_{\vec{q}}(\tau) \rangle.$$

С учетом явного вида слагаемых системы (2) уравнение (3) для многомерной плотности вероятности сводится к виду

$$\begin{aligned}
\partial\omega(\{\xi_{\vec{k}_u}\})/\partial t = & -\sum_{\vec{k}_u} \frac{\partial}{\partial \xi_{\vec{k}_u}} \left\{ \left[\lambda_1(\vec{k}_u) \xi_{\vec{k}_u} + \right. \right. \\
& + \sum_{\vec{k}'_u, \vec{k}''_u} \sigma_{11}^{(1)}(\vec{k}_u, \vec{k}'_u, \vec{k}''_u) \xi_{\vec{k}'_u} \xi_{\vec{k}''_u} \delta(\vec{k}'_u + \vec{k}''_u - \vec{k}_u) + \\
& + \sum_{\vec{k}'_u, \vec{k}''_u, \vec{k}'''_u} w(\vec{k}_u, \vec{k}'_u, \vec{k}''_u, \vec{k}'''_u, \vec{k}_s) \times \\
& \times \xi_{\vec{k}'_u} \xi_{\vec{k}''_u} \xi_{\vec{k}'''_u} \delta(\vec{k}'_u + \vec{k}''_u + \vec{k}'''_u - \vec{k}_u) + \\
& + \sum_{\vec{q}, \varphi} 0,5 \left(\eta_\varphi(\vec{k}_u, \vec{q}) O_\varphi^{*(1)}(\vec{q}) P_\varphi^{(0)} F_\varphi^{kor}(|\vec{k}_u - \vec{q}|) \delta(\vec{k}_u - 2\vec{q}) + \right. \\
& \left. + \eta_\varphi(\vec{k}_u, \vec{q}) \eta_\varphi(\vec{q}, 2\vec{q} - \vec{k}_u) \xi_{2\vec{q} - \vec{k}_u} F_\varphi^{kor}(|\vec{k}_u - \vec{q}|) \right] \omega \Big\} + \quad (4) \\
& + \sum_{\vec{k}_u, \vec{q}} \frac{\partial^2}{\partial \xi_{\vec{k}_u} \partial \xi_{\vec{q}}} \left\{ \left[\sum_{\varphi=1}^2 0,5 \left(O_\varphi^{*(1)}(\vec{k}_u) P_\varphi^{(0)} \right)^2 F_\varphi^{kor}(|\vec{k}_u|) + \right. \right. \\
& + \sum_{\varphi} 0,5 \eta_\varphi(\vec{q}, \vec{q} - \vec{k}_u) O_\varphi^{*(1)}(\vec{k}_u) P_\varphi^{(0)} \xi_{\vec{q} - \vec{k}_u} F_\varphi^{kor}(|\vec{k}_u|) + \\
& + \sum_{\varphi} \eta_\varphi(\vec{k}_u, \vec{k}_u - \vec{q}) O_\varphi^{*(1)}(\vec{q}) P_\varphi^{(0)} \xi_{\vec{k}_u - \vec{q}} F_\varphi^{kor}(|\vec{q}|) + \\
& + \sum_{\varphi, \vec{q}'} 0,5 \eta_\varphi(\vec{k}_u, \vec{k}_u - \vec{q} + \vec{q}') \eta_\varphi(\vec{q}, \vec{q}') \times \\
& \left. \left. \times \xi_{\vec{k}_u - \vec{q} + \vec{q}'} \xi_{\vec{q}'} F_\varphi^{kor}(|\vec{q} - \vec{q}'|) \right] \omega \right\}.
\end{aligned}$$

Уравнение (4) существенно упрощается в одномерном случае, если условия неустойчивости выполняются только для одной моды с волновым вектором \vec{k}_c и амплитудой $\xi_{\vec{k}_c}$

$$\begin{aligned}
\partial\omega/\partial t = & -\frac{\partial}{\partial\xi_{kc}} \left(\left[\lambda_1(k_c) + \sum_{\varphi} 0,5 \eta_{\varphi}^2(k_c, k_c) F_{\varphi}^{kor}(0) \right] \xi_{kc} + \right. \\
& + \left. \left(2w(k_c, k_c, k_c, k_c, 0) + w(k_c, k_c, k_c, k_c, 2k_c) \right) \xi_{kc}^3 \right] \omega) + \\
& + \frac{\partial^2}{\partial\xi_{kc}^2} \left\{ \left[\sum_{\varphi=1}^2 0,5 \left(O_{\varphi}^{*(1)}(k_c) P_{\varphi}^{(0)} \right)^2 F_{\varphi}^{kor}(|k_c|) + \right. \right. \\
& \left. \left. + 1,5 \sum_{\varphi} \eta^2(k_c, k_c) \xi_{kc}^2 F_{\varphi}^{kor}(0) \right] \omega \right\}. \tag{5}
\end{aligned}$$

Стационарное решение уравнения (5) имеет вид

$$\omega = N (d\xi_{kc}^2 + c)^{\frac{ad-cb}{2d^2}} \exp(b\xi_{kc}^2/2d),$$

где

$$b = 2w(\vec{k}_c, \vec{k}_c, \vec{k}_c, \vec{k}_c, 0) + w(\vec{k}_c, \vec{k}_c, \vec{k}_c, \vec{k}_c, 2\vec{k}_c),$$

$$a = \lambda_1(\vec{k}_c) - 2,5 \eta^2(\vec{k}_c, \vec{k}_c) F_{\varphi}^{kor}(0),$$

$$d = 1,5 \sum_{\varphi=1}^2 \eta^2(\vec{k}_c, \vec{k}_c) F_{\varphi}^{kor}(0),$$

$$c = \sum_{\varphi=1}^2 0,5 \left(O_{\varphi}^{*(1)}(\vec{k}_c) P_{\varphi}^{(0)} \right)^2 F_{\varphi}^{kor}(0),$$

$$N = 1 / \int_{-\infty}^{\infty} (d\xi_{kc}^2 + c)^{\frac{ad-cb}{2d^2}} \exp(b\xi_{kc}^2/2d) d\xi_{kc}$$

– нормировочная постоянная.

Уравнения (4)–(6) были применены для исследования модифицированной системы Гирера – Майнхардта, в которой учтены флуктуации скорости образования автокаталитической переменной и ее постоянной распада [3]:

$$\begin{aligned}
\partial a / \partial t = & \rho c (1 + f_2(\vec{r}, t)) / \nu k + a^2 / h - \mu (1 + f_1(\vec{r}, t)) a / \nu + \Delta a, \\
\partial h / \partial t = & a^2 - h + D \Delta h. \tag{6}
\end{aligned}$$

Здесь $D = D_h/D_a$, ρ – скорость образования автокаталитической переменной a , μ – ее постоянная распада, D_a – коэффициент диффузии активатора; ν – постоянная скорости распада демпфирующей переменной h , D_h – коэффициент диффузии ингибитора.

На рис. 1 представлен вид стационарного распределения плотности вероятности для одной неустойчивой моды системы (7). Параметры модели выбраны так, что система находится вдали от точки бифуркации Тьюринга, в закритической области.

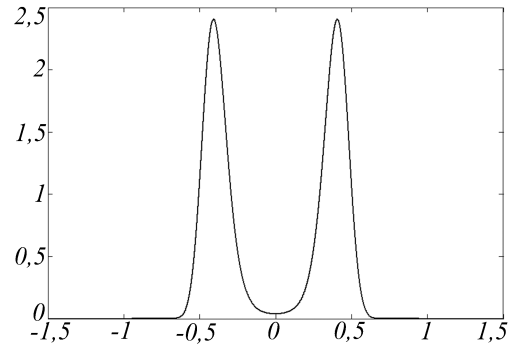


Рис. 1. Расщепление максимума стационарной плотности вероятности для одной неустойчивой моды системы (7). Параметры модели: $\rho = 0,1$; $\mu = 1$; $D = 400$; $k_{f1} = 1$; $k_{f2} = 1$; $\theta_1 = 0,01$; $\theta_2 = 0,01$

В вероятностной трактовке эффект возникновения нового статистически стационарного состояния проявляется в расщеплении максимума плотности вероятности на два симметричных, что продемонстрировано на рис. 1. При этом математическое ожидание параметра порядка становится отличным от нуля, а дисперсия возрастает до макроскопических значений. При малой интенсивности мультипликативных флуктуаций полученный результат хорошо согласуется с результатами работ [1, 3]. При больших шумах (рис. 2) наблюдается обратный эффект – слияние максимумов плотности вероятности, т.е. система возвращается в состояние с сильно нерегулярным поведением.

Моделирование эволюции системы (7) показало хорошее соответствие результатов численного и аналитического исследований.

Таким образом, на основании полученного в работе уравнения

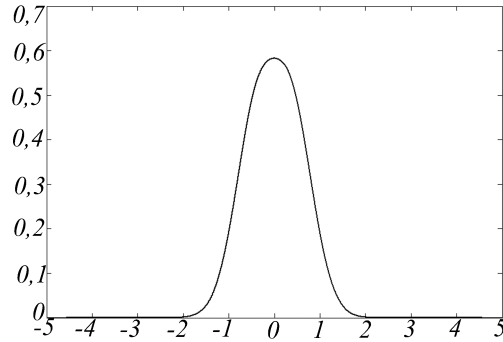


Рис. 2. Слияние максимумов стационарной плотности вероятности для одной неустойчивой моды системы (7) при больших шумах. Параметры модели: $\rho = 0, 1$; $\mu = 1$; $D = 400$; $k_{f1} = 1$; $k_{f2} = 1$; $\theta_1 = 0,8$; $\theta_2 = 0,01$

Колмогорова – Фоккера – Планка для параметров порядка двух-компонентной системы реакция-диффузия показано, что в системах такого типа, и частности в системе Гирера – Майнхардта, возникает шумоиндуцированный фазовый переход «беспорядок – порядок – беспорядок» (“reentrant phase transition”).

Литература

1. Курушина С.Е. Анализ поведения пространственных диссипативных структур в системах реакция-диффузия в поле мультипликативных флуктуаций в окрестности точки бифуркации // Компьютерная оптика, 2010. Т. 34. Вып. 1. С. 343–352.
2. Стратонович Р.Л. Случайные процессы в динамических системах / Под ред. Ю.Л. Климонтовича, Ю.М. Романовского. М. – Ижевск: ИКИ, 2009. 592 с.
3. Курушина С.Е., Иванов А.А. Диссипативные структуры в системе «реакция-диффузия» в поле мультипликативных флуктуаций // Известия ВУЗов. Прикладная нелинейная динамика, 2010. № 3. С. 85–103.

Михеев В.С., Михеев С.Е.

Санкт-Петербургский государственный университет

Интерполяция по линейчатой таблице

При интерполяции полиномом функции двух переменных (x, y) большое значение имеет расположение узлов. В регулярном случае, т.е., когда узлы расположены в виде прямолинейной решетки с $k + 1$ координатами по x и $m + 1$ координатами по y (всего $(k + 1)(m + 1)$ узлов), как известно, существует интерполятор (полином, решающий интерполяционную задачу) степени не более k по x и не более m по y . Известны и оценки погрешности для него. В произвольном случае расположения узлов можно, не взирая на одну из координат узлов (например на y), построить интерполятор только по одной переменной. Но степень его в общем случае окажется лишь на единицу меньше числа узлов, что может быть чересчур велико.

Однако существуют задачи с частичной упорядоченностью узлов, которая позволяет существенно снизить максимум степеней интерполятора по каждой переменной. Здесь будет рассмотрена упорядоченность в виде линейчатой таблицы.

Пусть множество узлов состоит из подмножеств с одинаковыми первыми координатами. Множества вторых координат от подмножества к подмножеству могут различаться. (Если они совпадают, то это задача интерполяции по прямоугольной таблице.) Пусть z — значение интерполяционной функции в узлах.

Назовем линейчатой таблицей таблицу в j -ой строке которой m_j элементов.

Таблица 1. Исходные данные

y_0	(x_{00}, z_{00})	(x_{01}, z_{01})	\dots	(x_{0m_0}, z_{0m_0})
\dots	\dots	\dots	\dots	\dots
y_k	(x_{k0}, z_{k0})	(x_{k1}, z_{k1})	\dots	(x_{km_k}, z_{km_k})

Пусть $m := \max(m_0, \dots, m_k)$. Опишем алгоритм построения интерполятора невысоких степеней по x и y . Интерпретируем i -ую

строку таблицы как самостоятельную интерполяционную задачу по m_i узлам с переменной x . Представим ее решение в таком виде

$$p(x, j) = \sum_{i=0}^{m_j} a_{ij} x^i.$$

В этом месте Б.П. Демидович [1] для прямоугольной таблицы предлагал вычислять для заданного x все значения $p(x, j) := P(x, y_j)$, $j = \overline{1, k}$, а затем решать еще одну интерполяционную задачу по y с этими значениями. Здесь мы поступим несколько иначе, что приведет к построению интерполятора от двух переменных и к возможности оценки его погрешности.

Дополним множество коэффициентов $A_j := \{a_{ij}\}_{i=1}^{m_j}$, когда $m_j < m$, нулями, т. е., $a_{ij} := 0$, $i = \overline{m_j + 1, m}$. Тогда можно ввести $m + 1$ интерполяционных задач по переменной y .

Таблица 2. Интерполяция коэффициентов

y	y_0	\dots	y_k
$a_j(y)$	a_{0j}	\dots	a_{kj}

Они имеют решениями полиномы $a_j(y)$ степени не более k . Очевидно, что полином

$$P(x, y) = \sum_{j=0}^m a_j(y) x^j \quad (1)$$

удовлетворяет линейчатой таблице. Таким образом справедлива

Теорема 1. *Существует интерполятор степени по x не более m и по y не более k для задачи интерполяции по таблице 1.*

Этот алгоритм является незначительной модификацией известного алгоритма интерполяции по прямоугольной таблице [2]. Однако в этом случае оценка погрешности интерполятора (1) заметно сложнее.

Теорема 2. *Построенный интерполятор (1) – единственный, который решает задачу интерполяции по таблице 1 со степенью по x не более m и по y не более k .*

Доказательство. Пусть есть еще один интерполятор $Q(x, y)$ со степенью по x не более m и по y не более k . Представим его в аналогичном виде

$$Q(x, y) = \sum_{j=0}^m b_j(y)x^j,$$

где $b_j(y)$ – полином степени не более k . Тогда, полагая $\check{a}_j := b_j - a_j$, видим, что полином

$$P(y, x) - Q(x, y) =: \check{P}(x, y) = \sum_{j=0}^m \check{a}_j(y)x^j$$

удовлетворяет таблице 1 с нулевыми значениями $\{z_i\}$. Это означает, что $\check{P}(x, y_i) = c(y_i) \prod_{s=0}^{m+1} (x - x_s)$, $i = \overline{0, k}$. Поскольку степень по x полинома \check{P} не более m , последние равенства влекут $c(y_i) = 0$, $i = \overline{0, k}$, т.е. $c(y) = d \prod_0^k (y - y_i)$. Поскольку полином c имеет степень не более k , последнее равенство влечет $c(y) = 0$. Значит

$$P(x, y) \equiv Q(x, y).$$

Теорема 2 доказана.

Как известно, остаточный член для полинома $P_j(x)$, когда f имеет производную по y порядка $m + 1$, принимает вид

$$r_j(x) \equiv f(x, y_j) - P_j(x) = \frac{f^{(m_j+1)}(\xi_j, y_j)}{(m_j + 1)!} \omega_j(x), \quad (2)$$

где $\omega_j(x) = (x - x_{j0}) \dots (x - x_{jm_j})$ и $\xi_j \in \text{convex}(x_{j0}, \dots, x_{jm_j})$. Здесь и далее функция $\text{convex}(\cdot)$ сопоставляет своим аргументам их выпуклую оболочку.

Выберем $m + 1$ узел x_0, \dots, x_m для переменной x . В частности, это могут быть значения первой переменной в узлах x_{j0}, \dots, x_{jm_j} для некоторого j , когда $m_j = m$. Или же можно назначить сетку равномерно покрывающих выпуклую оболочку значений первых переменных во всех узлах. Мы же для удобства оценки погрешности возьмем Чебышевскую систему для этой выпуклой оболочки.

Рассмотрим вспомогательную задачу интерполирования по прямоугольной таблице с узлами $u_{ij} = (x_i, y_j)$, $i = \overline{0, m}$, $j = \overline{0, k}$, со значениями $F_{ij} = P(x_i, y_j)$.

Очевидно, что сам полином P и будет нужным решением вспомогательной задачи с ограничениями на степень по x не более m и на степень по y не более k . Эту же вспомогательную задачу можно трактовать как задачу интерполяции исходной функции f по значениям, полученным с погрешностями, в узлах $\{u_{ij}\}$.

Таким образом, общую погрешность интерполятора P можно представить в виде методической погрешности интерполяции по прямоугольной таблице и неустраняемой погрешности от начальных данных.

Положим $r_{ij} := r_j(x_i)$, $i = \overline{0, m}$, $j = \overline{0, k}$. Тогда неустраняемая погрешность приобретает вид

$$\Delta(x, y) = \sum_{j=0}^k \sum_{i=0}^m \frac{\omega(x)\Omega(y)r_{ij}}{(x-x_i)\omega'(x_i)(y-y_j)\Omega'(y_j)},$$

где $\omega(x) = (x-x_0)\dots(x-x_m)$, $\Omega(y) = (y-y_0)\dots(y-y_k)$.

Оценим Δ . Пусть x и y находятся внутри диапазонов по x и по y соответственно. Пусть диапазон узлов x_0, \dots, x_m имеет длину l_x и

$$\left| \frac{\partial f^{(m+1)}(\xi, y_j)}{\partial x^{m+1}} \right| \leq M_x, \quad j = \overline{0, k}, \quad \forall \xi \in \text{convex}(x_0, \dots, x_m).$$

Тогда

$$|r_j(x)| \leq \frac{M_x}{(m+1)!} l_x^m, \quad j = \overline{0, k}, \quad \forall \xi \in \text{convex}(x_0, \dots, x_m).$$

Отсюда

$$|r_{ij}| \leq \frac{M_x}{(m+1)!} l_x^m, \quad i = \overline{0, m}, \quad j = \overline{0, k},$$

и

$$|\Delta(x, y)| \leq \frac{M_x l_x^m}{(m+1)!} \sum_{j=0}^k \sum_{i=0}^m \left| \frac{\omega(x)\Omega(y)}{(x-x_i)\omega'(x_i)(y-y_j)\Omega'(y_j)} \right|. \quad (3)$$

Для получения методической погрешности проведем интерполяцию с помощью полинома Ньютона

$$f(x, y) = \sum_0^m \omega_i(x) f(x_0, \dots, x_i, y) + \omega_{k+1}(x) f(x_0, \dots, x_k, x, y). \quad (4)$$

Здесь обобщенная разность $(x - x_0) \dots (x - x_{i-1})$ обозначена через $\omega_i(x)$, причем $\omega_0(x) = 1$. Разделенные разности под знаком суммы тоже разложим на полином и некоторую добавку с помощью полинома Ньютона:

$$f(x_0, \dots, x_i, y) = \sum_{j=0}^k \hat{\omega}_j(y) f(x_0, \dots, x_i, y_0, \dots, y_j) + \hat{\omega}_{k+1}(y) f(x_0, \dots, x_i, y_0, \dots, y_k, y), \quad (5)$$

где $\hat{\omega}_j(y) = (y - y_0) \dots (y - y_{j-1})$, $\hat{\omega}_0(x) = 1$. Подставим (5) в (4):

$$f(x, y) = \sum_{i=0}^m \omega_i(x) \sum_{j=0}^k \hat{\omega}_j(y) f(x_0, \dots, x_i, y_0, \dots, y_j) + R_{k,m}(x),$$

где

$$R_{k,m}(x, y) = \sum_{i=0}^m \omega_i(x) \hat{\omega}_{k+1}(y) f(x_0, \dots, x_i, y_0, \dots, y_k, y) + \omega_{m+1}(x) f(x_0, \dots, x_m, x, y). \quad (6)$$

Заметим, что

$$\begin{aligned} \sum_{i=0}^m \omega_i(x) f(x_0, \dots, x_i, y_0, \dots, y_k, y) &= \\ = f(x, y_0, \dots, y_k, y) - \omega_{m+1}(x) f(x_0, \dots, x_i, x, y_0, \dots, y_k, y). \end{aligned} \quad (7)$$

Используя (7) в (6), получаем симметричное по x и y представление остаточного члена

$$R_{k,m}(x, y) = \omega_{m+1}(x) f(x_0, \dots, x_m, y) + \hat{\omega}_{k+1}(y) f(x, y_0, \dots, y_k, y) - \omega_{m+1}(x) \hat{\omega}_{k+1}(y) f(x_0, \dots, x_m, x, y_0, \dots, y_k, y). \quad (8)$$

Как известно, разделенные разности и производные функции φ связаны соотношением

$$\varphi(v_0, \dots, v_i) = \frac{\varphi^{(i)}(s)}{i!}, \quad s \in \text{convex}(v_0, \dots, v_i).$$

Поэтому, считая, что f дифференцируема по аргументам достаточное число раз, имеем

$$R_{k,m}(x,y) = \frac{\partial^{m+1} f(s_x, y)}{\partial x^{m+1}} \frac{\omega_{m+1}(x)}{(m+1)!} + \frac{\hat{\omega}_{k+1}(y)}{(k+1)!} \frac{\partial^{k+1} f(x, s_y)}{\partial y^{k+1}} - \frac{\omega_{m+1}(x) \hat{\omega}_{k+1}(y)}{(k+1)!(m+1)!} \frac{\partial^{m+k+2} f(\xi, \delta)}{\partial x^{m+1} \partial y^{k+1}}, \quad (9)$$

где $s_x, \xi \in \text{convex}(x_0, \dots, x_m)$ и $s_y, \delta \in \text{convex}(y_0, \dots, y_k)$.

Оценкой полной погрешности можно назначить сумму правых частей неравенств (3) и (9).

Литература

1. Демидович Б. П., Марон И. А. Основы вычислительной математики. М.: Наука, 1966. 664 с.
2. Кирич Н. Е. Интерполяционные методы оценивания на решениях функциональных включений // Известия РАН. Теория и системы управления, 1995. Т. 3. С. 206–209.

Подзывалов Е.А.

Санкт-Петербургский государственный университет

Метод глобально-оптимального поиска структуры канала согласования

Рекомендовано к публикации профессором Андриановым С.Н.

1. Введение. Проблема выбора оптимальной структуры *канала согласования* возникает в процессе разработки структуры комплексов современных циклических установок. Дело в том, что в подобных сложных системах необходимо «перевести» пучок из одной области пространства в другую при одновременном обеспечении дополнительных ограничений. Иными словами, основной функцией *канала согласования* является транспортировка некоторого входного множества частиц в заданное выходное.

Учитывая сложность подобных систем и строгие требования к характеристикам пучка, необходимо осуществлять отбор вариантов оптимальной структуры канала согласования из найденного семейства допустимых решений, обеспечивающих необходимое согласование с нечетко сформулированными критериями оптимальности, что является спецификой именно данного класса задач. Степень надежности подобных установок, требуемые функциональные характеристики и возможность выбора оптимальных значений параметров «заставляют» исследователя, так или иначе, решать задачу *глобальной оптимизации*.

Методы глобальной оптимизации находят свое применение в теории ускорителей сравнительно недавно [1, 2]. В статье [2], например, рассматривается задача поиска оптимальных параметров линеаризованной системы согласования, минимизирующей потерю частиц в канале. Включение в задачу дополнительных функционалов или ограничений на управляющие параметры требует от исследователя проведения тщательного анализа их взаимного влияния.

Рассмотренный в работе метод многоцелевой оптимизации, основанный на идеологии *эволюционных алгоритмов* [3], относится к категории многопараметрических методов поиска. Основным достоинством представленного подхода является возможность поиска глобального и локальных экстремумов для построенного функционала.

Разработанный на основе идеологии *генетических алгоритмов*

программный продукт позволяет осуществлять подобного рода гибкий анализ семейства допустимых решений.

2. Постановка задачи согласования. В данной работе под каналом согласования будем понимать систему управления пучками частиц, состоящую из управляющих элементов (диполи, квадруполь) вместе с дополнительными ограничениями на характеристики пучка в канале.

Динамика частиц в электромагнитных полях в окрестности некоторой выделенной (опорной) траектории s описывается уравнениями Ньютона – Лоренца:

$$\frac{d\mathbf{x}}{ds} = \mathbf{F}(\mathbf{x}, s), \quad (1)$$

где $\mathbf{F}(\mathbf{x}, s) = \mathbf{F}(\mathbf{x}, \mathbf{E}, \mathbf{B}, s)$ сила Ньютона – Лоренца, \mathbf{E} – вектор электрической напряженности, \mathbf{B} – вектор магнитной индукции поля. Уравнения движения (1) в общей форме могут быть записаны в следующем виде (см., например, [4])

$$\frac{d\mathbf{x}}{ds} = \sum_{k=0}^{\infty} \mathbf{P}^{1k}(\mathbf{E}, \mathbf{B}, s) \mathbf{x}^{[k]}, \quad (2)$$

где $\mathbf{x}^{[k]} = \mathbf{x} \otimes \dots \otimes \mathbf{x}$ (где символ « \otimes » обозначает кронекерово произведение); \mathbf{P}^{1k} – матрицы с элементами, равными k -й производной компонент вектор-функции $\mathbf{F}(\mathbf{x}, s)$; $\mathbf{x} = \{x, x', y, y'\}^*$ – вектор поперечных фазовых координат для двумерного случая (символ « $*$ » обозначает транспонирование) и $x' = dx/ds$.

Решения уравнений (2) могут быть найдены в виде ряда (3), сходящегося на конечном интервале $[s, s_0]$, с использованием идеологии матричного формализма [4]:

$$\mathbf{x}(s) = \sum_{k=0}^{\infty} \mathbf{R}^{1k}(\mathbf{E}, \mathbf{B}; s|s_0) \mathbf{x}_0^{[k]}, \quad (3)$$

где $\mathbf{x}_0 = \mathbf{x}(s_0)$ – «начальное значение» фазового вектора, \mathbf{R}^{1k} – матричные пропагаторы системы (3).

Для описания основных ограничений на характеристики пучка будем использовать следующий общий вид функционала [5]

$$J[\mathbf{u}] = \int_{s_0}^{s_T} \int_{\mathfrak{M}(s)} g_1(\mathbf{u}, \mathbf{x}, \tau) d\mathbf{x} d\tau + \int_{\mathfrak{M}_T} g_2(\mathbf{u}, \mathbf{x}, T) d\mathbf{x}, \quad (4)$$

где функция g_1 описывает распределение функциональных критериев внутри системы (если это необходимо), g_2 – конечные требования к пучку, \mathbf{u} – вектор управляющих параметров.

3. Глобально-оптимальный подход. В силу ограниченности вычислительных ресурсов и в зависимости от типов введенных ограничений на параметры модели и целевых функционалов применяют различные методы оптимизации (см., например, [1, 3]). Основная идея методов глобальной оптимизации заключается в нахождении некоторого множества элементов x^* из \mathcal{X} , удовлетворяющих вводимым ограничениям на множестве функционалов $\mathcal{F} = \{f_1, \dots, f_n\}$.

Обычно заранее определить зависимость между этими целевыми функциями сложно, поэтому многие исследователи применяют вероятностные методы для поиска оптимальных решений задач. Простейший метод задания механизма поиска оптимальных решений основывается на вычислении взвешенной суммы всех функций $f_i \in \mathcal{F}$ в виде $g(x) = \sum_{i=1}^n \omega_i f_i$. Однако взвешенная сумма применяется только при оптимизации функций, по крайней мере, одинаковых степеней. Поэтому в процессе решения задачи выбор целевых функций и ограничений возлагается на разработчика модели.

В работе [6] предприняты попытки поиска оптимальной структуры оптических систем в линейном приближении в соответствии со стандартной стратегией исследования подобных систем. Для простых систем задача разрешается с помощью детерминированных методов градиентного спуска. В случае сложных систем нет гарантии нахождения глобального оптимума, что приводит к необходимости разработки методов поиска множества экстремумов [7], с целью выбора оптимального из них с точки зрения специфики задачи.

Подобная концепция глобальной оптимизации относится к решению задачи поиска множества оптимальных структур для канала согласования, в которую также будут включены целевые функции, отражающие вводимые исследователем ограничения. В данной статье осуществляется синтез оптимального управления с целью его реализации на реальных установках.

4. Практическая реализация. В качестве примера поиска оптимальной структуры согласующего канала рассмотрим простую систему управления пучками частиц, состоящую из четырех квадрупольных линз. В качестве управляющих параметров выступают приведенные градиенты линз k_i , $i = \overline{1,4}$ (при соответствующей вариации величин k_i по знаку, что соответствует последовательности

фокусировки-дефокусировки), их длины L_i , расстояния между линзами λ_i , начальный и конечный свободные промежутки a , g соответственно.

Рассмотрим некоторые основные характеристики пучка, наблюдаемые в системе, с помощью матрицы огибающих

$$\mathbf{S}(s) = \mathbf{R}(s|s_0)\mathbf{S}_0\mathbf{R}^*(s|s_0). \quad (5)$$

Используя матричный формализм [4] и полагая, что \mathbf{M}_{obj} является матрицей фокусирующей части системы, запишем общий матрицант в следующей форме:

$$\mathbf{R} = \mathbf{R}_g\mathbf{M}_{\text{obj}}\mathbf{R}_a, \quad (6)$$

где матрицы \mathbf{R}_g и \mathbf{R}_a будут использованы для расчета условия фокусировки пучка *из точки в точку*: $r_{12} = r_{34} = 0$, где r_{12} соответствует фокусировке по координате x , r_{34} – по координате y . Выпишем данную зависимость для параметров a и g :

$$g = -\frac{am_{11} + m_{12}}{am_{21} + m_{22}}, \quad \frac{am_{11} + m_{12}}{am_{21} + m_{22}} = \frac{am_{33} + m_{34}}{am_{43} + m_{44}}, \quad (7)$$

где m_{ij} – элементы матрицы \mathbf{M}_{obj} . Заметим, что элементы r_{11} и r_{33} со значением $r_{11} > 1$, $r_{33} > 1$ соответствуют увеличению (в случае $r_{11} < 1$, $r_{33} < 1$ – уменьшению) размера пучка после прохождения канала согласования [4].

Таким образом, функциональные ограничения на пучок на мишени (например, (7)) являются основными для реализации канала согласования. Однако нужно понимать, что требуется выполнение других важных ограничений на характеристики пучка на протяжении всего согласующего канала. Подставив (6) в (5), найдем матрицу $\mathbf{S}(s)$, позволяющую отслеживать огибающие пучка в канале согласования. Выбор матрицы огибающей для описания пучка во многих задачах обеспечивает необходимую достаточность. В частности, диагональные элементы представляют либо наибольшие, либо усредненные значения по соответствующим координатам.

Вводя пороговые значения огибающей $\mathbf{S}_{\text{opt}}(s)$ на всем участке канала согласования, допустимое отклонение матрицы огибающей $\mathbf{S}(s)$ от порогового значения можно записать в некоторой подходящей функциональной метрике:

$$\|\mathbf{S} - \mathbf{S}_{\text{opt}}\| \leq \varepsilon.$$

В качестве демонстрации работоспособности метода решается задача поиска значений градиентов линз k_i с точки зрения минимизации пучка на мишени. Одним из функциональных критериев часто выступает число $F = R_{11}^2 + R_{33}^2$, характеризующее коэффициент сжатия ($F \leq 1$) или растяжения (для $F > 1$) размеров пучка на мишени.

На рис. 1 отражены области, соответствующие сжатию пучка при $F \leq 1$.

Выбор того или иного решения осуществляется с учетом его устойчивости к отклонениям других управляющих параметров (так называемая «проблема допусков»).

Для поиска оптимального решения при минимуме функционала F следует воспользоваться данными на рис. 2, 3, на которых отражен характер поведения функционала при различных значениях параметров k_i .

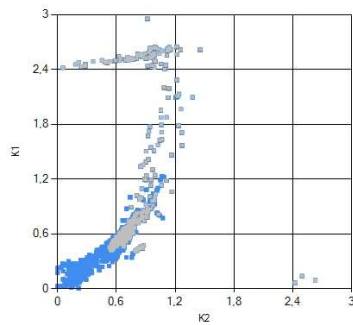


Рис. 1. Градиенты линз, удовлетворяющие условию $F \leq 1$

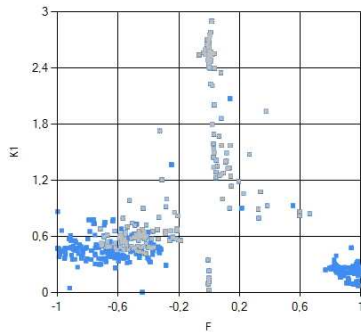


Рис. 2. Градиент k_1 для $F \leq 1$

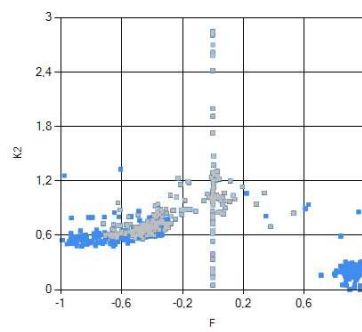


Рис. 3. Градиент k_2 для $F \leq 1$

Как показано на рис. 2, 3 минимальным значениям функционала для градиента k_1 ($F \leq 0, 1$) соответствуют интервалы $[1, 2; \dots; 1, 8]$ и $[2, 4; \dots; 2, 8]$, для $k_2 - [0, 6; \dots; 1, 2]$. Данный процесс можно продолжать, сужая допустимые значения функционала F . Предложенный алгоритм позволяет использовать методы и технологии высокопроизводительных вычислений, а также применять подобную идеоло-

гию для расчета оптимальных структур различного класса сложных систем управления пучками частиц.

Окончательный выбор оптимального решения возлагается на разработчика системы, на основании его опыта и с учетом введенных им целевых функций.

Заключение. Рассмотренный метод глобально-оптимального поиска структуры канала согласования предоставляет возможность опытному исследователю осуществлять выбор семейства оптимальных структур для системы управления пучками частиц.

Подобный подход может быть использован в задачах синтеза сложных физических структур. Планируется применение реализованного программного комплекса для моделирования и предварительных расчетов проекта NICA (в частности, выбор структуры канала согласования между существующим кольцом «Нуклотрона» и проектируемым кольцом разгона – «Бустером»).

Литература

1. Yang L., Robin D. Global optimization of the magnetic lattice using genetic algorithms // EPAC'08, Genoa, Italy, 2008. P. 3050–3052.
2. Gao W., Wang L. Low emittance lattice optimization using multiobjective genetic algorithm // IPAC'10, Kyoto, Japan, 2010. P. 4515–4517.
3. Weise T. Global Optimization Algorithms: Theory and Application, 2nd Edition. <http://www.it-weise.de/projects/book.pdf>
4. Андрианов С.Н. Динамическое моделирование систем управления пучками частиц. СПб.: Изд-во СПбГУ, 2004. 368 с.
5. Ovsyannikov D.A., Merkurjev S.V. On beam dynamics optimization // EPAC'04, Lucerne, Switzerland, 2004. P. 2152–2154.
6. Autin B., Royer Ph. Optimum optical systems // EPAC'00, Vienna, Austria, 2000. P. 1396–1398.
7. Schuster E., Allen C.K. Optimized beam matching using extremum seeking // PAC'05, Knoxville, Tennessee, USA, 2005. P. 4269–4271.

Прокопенко А.В.

Санкт-Петербургский государственный университет

Траектории звёзд в модели галактики, допускающей три интеграла движения

Рекомендовано к публикации профессором Кутузовым С.А.

Введение. В данной работе используется подход к построению модели галактики с заданием гравитационного потенциала. Рассматривается стационарная модель с ротационной симметрией, что обеспечивает существование двух интегралов движения: интеграла энергии и интеграла площадей. В [1] Кузмин, накладывая дополнительное ограничение на потенциал, ввёл третий квадратичный по скоростям интеграл движения, независимый от первых двух.

Им же рассмотрены семейства моделей с различными параметрами и предложен метод вычисления траекторий, однако конкретные модели подробно не исследовались, и приложение метода не было продемонстрировано. Здесь выбирается одна из моделей с параметром $n = 3$, и рассчитываются траектории звёзд, для чего решается проблема начальных значений.

1. Единицы измерения. Все фигурирующие в данной работе переменные и функции являются безразмерными. Переход к размерным величинам осуществляется умножением на соответствующие размерные величины, которые служат единицами длины, потенциала и скорости [2]:

$$u_R = \frac{z_0}{\beta}, \quad u_\varphi = \frac{G\beta}{z_0}, \quad u_v = \sqrt{u_\varphi}, \quad \beta \equiv \sqrt{1 - \gamma^2}, \quad (1)$$

где G – гравитационная постоянная, z_0 – масштабный параметр, $0 \leq \gamma < 1$ – структурный параметр модели. В случае $\gamma = 0$ модель обращается в диск, а когда $\gamma \rightarrow 1$, система становится сферической.

2. Модель Кузмина $n = 3$. Безразмерный потенциал для такой модели имеет вид

$$\varphi(\zeta_1, \zeta_2) = \frac{1}{\zeta_1 + \zeta_2}, \quad 1 \leq \zeta_1 < \infty, \quad \gamma \leq \zeta_2 < 1,$$

где ζ_1, ζ_2 – разновидность безразмерных эллиптических координат.

Используя соотношения, приведённые в работе [1], находим выражение для потенциала в цилиндрических координатах R, z :

$$\varphi(R, z) = \frac{1}{\sqrt{R^2 + (Q(R, z) + 1)^2}}, \quad Q \equiv \sqrt{\gamma^2 + \frac{z^2 + \gamma^2 R^2}{\beta^2}}. \quad (2)$$

3. Интегралы движения. В работе [1] приведены выражения для трёх интегралов движения в цилиндрических координатах (θ – азимутальная цилиндрическая координата):

$$\begin{cases} E = \varphi(R, z) - (v_m^2 + v_\theta^2)/2, & v_m^2 = v_R^2 + v_z^2, \\ I = Rv_\theta, \end{cases} \quad (3)$$

$$K = (Rv_z - zv_R)^2 + z^2 v_\theta^2 + \beta^2 (v_z^2 - 2\varphi^*),$$

где E – интеграл энергии (полная энергия, взятая с обратным знаком), I – интеграл площадей, K – третий квадратичный по скоростям интеграл. Компоненты скоростей соответствуют цилиндрическим координатам, v_m – меридиональная скорость. $\varphi^*(R, z)$ – вспомогательная безразмерная функция, удовлетворяющая уравнениям

$$\beta^2 \frac{\partial \varphi^*}{\partial R} = z^2 \frac{\partial \varphi}{\partial R} - Rz \frac{\partial \varphi}{\partial z}, \quad \beta^2 \frac{\partial \varphi^*}{\partial R} = (R^2 + \beta^2) \frac{\partial \varphi}{\partial z} - Rz \frac{\partial \varphi}{\partial R},$$

для интегрируемости которых потенциал φ должен удовлетворять уравнению в частных производных, приведённому в [1].

4. Уравнения движения. Известно (см., например, [4]), что траектории звёзд в стационарных ротационно-симметричных моделях описываются дифференциальными уравнениями второго порядка в цилиндрических координатах.

$$\ddot{\mathbf{r}} = \nabla \varphi(\mathbf{r}), \quad \nabla = \left(\frac{\partial}{\partial R}, \frac{1}{R} \frac{\partial}{\partial \theta}, \frac{\partial}{\partial z} \right)^T. \quad (4)$$

Эти соотношения имеют смысл уравнений движения звезды в гравитационном поле галактики, задаваемым потенциалом φ . Представим теперь уравнение (4) в виде системы пяти дифференциальных уравнений первого порядка, используя интеграл площадей, и приведём

их покомпонентно [4]:

$$\begin{cases} \dot{R} = v_R, & \dot{v}_R = \frac{I^2}{R^3} + \frac{\partial \varphi}{\partial R}, \\ \dot{\theta} = \frac{I}{R^2}, \\ \dot{z} = v_z, & \dot{v}_z = \frac{\partial \varphi}{\partial z}. \end{cases} \quad (5)$$

5. Выбор начальных условий с использованием обобщённой диаграммы Линдблада. Множество возможных траекторий имеет мощность континуума. Поэтому для того, чтобы рассчитать траектории, необходимо выбрать характерные случаи. Поскольку проблема начальных значений решается эмпирически, следует попытаться так изменять начальные условия при переходе от одного случая к другому, чтобы каждый раз получать орбиты с новыми свойствами. Здесь предпринята попытка создать набор начальных условий с помощью диаграммы Линдблада [6], которая описывает движение звёзд в пространстве интегралов движения. На первом этапе ограничимся двумерной диаграммой, по осям которой отложены значения I и E .

Важной характеристикой гравитационного поля является круговая скорость $V_c(R)$. Воспользуемся её обобщением на пространственный случай в виде квазикруговой скорости [3], которая при $z = 0$ равна круговой:

$$V_q^2(R, z) \equiv -R \frac{\partial \varphi(R, z)}{\partial R}, \quad V_q^2(R, 0) = V_c^2(R). \quad (6)$$

Фиксируются некоторые значения γ , после чего для каждого такого γ определяются R_1, R_2, \dots , для которых $V_q(R, z_k)$ достигает максимума по R . Выбор z_k может быть таким:

$$z_1 = 0, \quad z_2 = R_1/2, \quad z_3 = R_2/2, \dots$$

Для каждой пары (R_k, z_k) находится потенциал $\varphi_k = \varphi(R_k, z_k)$ по (2). На двумерной диаграмме Линдблада, согласно [6], строим характеристические параболы, изображающие звёзды с нулевой меридиональной скоростью

$$E = \varphi_k - I^2/2R_k^2, \quad k = 1, 2, \dots \quad (7)$$

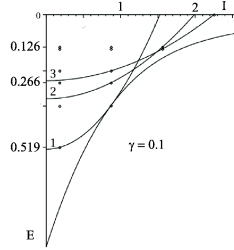


Рис. 1. Диаграмма Линдблада для $\gamma = 0,1$. Номера парабол соответствуют индексам k в формуле (7)

Изображаем на диаграмме Линдблада (рис. 1) последовательность I_l таким образом:

$$I_l = I_c(R_1) + (l - 2)\Delta I, \quad l = 1, 2, 3, 4,$$

где $I_c(R_1)$ – значение интеграла площадей в точке касания первой параболы с огибающей всего семейства характеристических парабол, $\Delta I = (I_{max} - I_c(R_1)) / 2$ задаёт шаг по оси I . Значения E_{kl} выбираются на пересечениях k -ой параболы с l -ой ординатой $I = I_l$. Значения I_l и E_{kl} дают компоненты скорости $(v_\theta)_l$ и $(v_m)_{kl}$, соответственно. Будем задавать скорость таким образом:

$$v_R^2 = \varkappa v_m^2, \quad \varkappa = 0; 0,25; 0,5; 0,75; 1,$$

выбирая как положительные, так и отрицательные значения v_R . Значение v_θ получаем из (3). В силу зеркальной симметрии, ограничимся лишь случаем $v_z \geq 0$.

6. Вычисления траекторий звёзд. Вместе с выбранными начальными условиями уравнения движения (5) образуют задачу Коши. Для их интегрирования используется метод Рунге – Кутты четвёртого порядка с переменным шагом. В качестве языка программирования выбирается C# [8], а для графического представления – программный пакет Maple (версии 13).

Время t , как и все переменные и функции, безразмерно. Вообще говоря, единица времени следует из (1), $u_t = u_R/u_v$, однако более наглядно и уместно использовать T – период прохождения звезды по круговой орбите с заданным радиусом R . Поскольку модель стационарна, в ней должен сохраняться интеграл энергии E . Но фактиче-

ски, вследствие округления чисел в памяти машины, ошибки исходных данных, а также погрешности метода интегрирования, текущее вычисленное значение энергии будет в общем случае отличаться от начального значения. Это даёт нам относительную погрешность интеграла энергии $\delta E = (E - E_0)/E_0$, где E_0 – значение интеграла энергии в начальной точке, E – в текущей точке, вычисляемое по формуле (3).

Также уместен дополнительный эмпирический контроль точности вычислений в случае, когда заранее известна траектория движения звезды, например, колебания по диаметру окружности в плоскости $z = 0$. В этом случае $\delta E \leq 0,5 \cdot 10^{-5}$ для 100 колебаний.

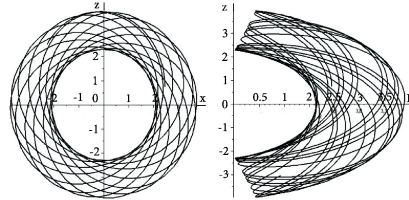


Рис. 2. Проекция траекторий движения звёзд на меридиональную плоскость: слева – второй группы, справа – четвёртой

7. Анализ расчётов траекторий. В данной работе были проведены расчёты около 400 траекторий звёзд. Их можно разделить на следующие группы в зависимости от начальных условий:

1. Звёзды, падающие с контура нулевой меридиональной скорости (начальная меридиональная скорость равна нулю).
2. Звёзды, движущиеся в меридиональной плоскости с неотрицательными начальными значениями v_R .
3. Звёзды, движущиеся в меридиональной плоскости с отрицательными начальными значениями v_R .
4. Звёзды, чьи траектории, согласно [7], образуют сильно изогнутый подковообразный «ящик».
5. Звёзды, движущиеся в экваториальной плоскости (начальная скорость $v_z = 0$ при значении γ , близком к нулю).
6. Звёзды, улетающие из галактики; их скорость превышает скорость освобождения (см., например, [5]).

Заключение. В данной работе вычислены траектории звёзд в модели галактики, допускающей три интеграла. В дальнейшем представляет интерес рассмотреть ограничение, которое накладывает третий интеграл на начальные условия, построить обобщённую трёхмерную диаграмму Линдблада, используя это ограничение. Также следует рассмотреть аналитический способ построения траекторий, предложенный Кузминым [1].

Литература

1. Кузмин Г.Г. Модель стационарной Галактики, допускающая трёхосное распределение скоростей // Астрон. журн., 1956. Т. 33. Вып. 1. С. 27–44.
2. Кузмин Г.Г., Кутузов С.А. Модели стационарных самогравитирующих звёздных систем с осевой симметрией // Тартуские сообщения, 1963. № 6. С. 10–15.
3. Кутузов С.А. Методы построения моделей распределения масс в галактиках / Докторская диссертация. Л.: ЛГУ, 1991. 296 с.
4. Кутузов С.А., Осипков Л.П. Методы расчёта галактических орбит // Движение искусственных и естественных небесных тел. Свердловск: Изд-во Урал. ун-та. 1981. С. 46–62.
5. Кутузов С.А., Распопова Н.В. Рельеф поля сил и орбиты в модели галактики // Вестник СПбУ. Сер. 10, 2008. Вып. 1. С. 32–42.
6. Линдblad Б. Динамика Галактики. Стрoение звездных систем. М.: ИЛ, 1962. С. 39–132.
7. Осипков Л.П., Кутузов С.А., Мюлляр А.А. Морфология галактических орбит шаровых скоплений // Вестник СПбУ. Сер. 1, 2001. Вып. 3. С. 81–88.
8. Шилдт Г. Полный справочник по $C\#$ / Пер. с англ. М.: Издат. дом «Вильямс», 2004. 752 с.

Пронин А.О.

Санкт-Петербургский государственный университет

Об устойчивости механического поведения нелинейно-упругих тел¹

Рекомендовано к публикации профессором Далем Ю.М.

Введение. Зарождение и слияние микропор является одним из характерных механизмов начальной стадии разрушения материалов. Поэтому анализ формирования новых и развития имеющихся поверхностей в твердых телах играет фундаментальную роль в механике разрушения. В качестве одной из основных расчетных моделей для этих исследований используется толстостенная сфера под действием равномерной нагрузки. Для описания механических свойств металлов и сплавов хорошо подходит деформационная теория Девиса – Надаи [1]. В рамках данной теории задача о полой сфере рассматривалась в [2, 3]. Интеграл уравнения равновесия в этих работах выписан в общем виде в текущих координатах, поэтому по форме совпадает с решением А.А. Ильюшина для малых деформаций. В статье [4] найдено решение задачи о больших логарифмических деформациях толстостенной сферы, окончательные выражения построены в отсчетной конфигурации для степенного закона деформирования. В настоящей работе на основе решения [4] проведено исследование кавитационной неустойчивости сплошного шара при растяжении и дана оценка критической нагрузки порообразования.

Постановка задачи. На основе решения задачи о полой сфере требуется исследовать напряженно-деформированное состояние сплошного шара при всестороннем растяжении. Для этой цели сначала рассматривается задача о толстостенной сфере под действием постоянной нормальной растягивающей нагрузки q , приложенной на ее внешней поверхности. Внутренний и внешний радиусы сферы в начальный момент времени обозначены через a и b соответственно. Механическое поведение материала сферы подчиняется соотношениям деформационной теории Девиса – Надаи.

Основные соотношения деформационной теории Девиса. Введем лагранжевы сферические координаты ρ, θ, φ с началом

¹Работа выполнена при финансовой поддержке СПбГУ, проект № 9.0.165.2009

в центре сферы. Из условий сферической симметрии и отсутствия касательной нагрузки вытекает, что напряжения $\sigma_{\rho\rho}$ и $\sigma_{\theta\theta} = \sigma_{\varphi\varphi}$ являются главными напряжениями, тангенциальные перемещения $u_\theta = u_\varphi = 0$, а радиальное перемещение u зависит только от координаты ρ . В нашем случае истинные деформации, отнесенные к сферической системе координат ρ, θ, φ , определяются выражениями [1, 2]

$$\varepsilon_{\rho\rho}(\rho) = \ln \left(1 + \frac{du}{d\rho} \right), \quad \varepsilon_{\theta\theta}(\rho) = \varepsilon_{\varphi\varphi}(\rho) = \ln \left(1 + \frac{u}{\rho} \right).$$

Согласно деформационной теории Девиса – Надаи, материал тела считается несжимаемым, а напряжения и деформации связаны между собой известными зависимостями

$$\sigma_{\rho\rho} - \sigma = \frac{2\sigma_i}{3\varepsilon_i} \varepsilon_{\rho\rho}, \quad \sigma_{\theta\theta} - \sigma = \frac{2\sigma_i}{3\varepsilon_i} \varepsilon_{\theta\theta}, \quad \sigma_{\varphi\varphi} - \sigma = \frac{2\sigma_i}{3\varepsilon_i} \varepsilon_{\varphi\varphi}, \quad (1)$$

где $\sigma = (\sigma_{\rho\rho} + \sigma_{\theta\theta} + \sigma_{\varphi\varphi})/3$ – среднее нормальное напряжение, σ_i – интенсивность напряжений (независящая от вида напряженного состояния), ε_i – интенсивность деформаций. В качестве закона деформирования материала принята степенная зависимость

$$\sigma_i = B \varepsilon_i^\mu, \quad 0 < \mu \leq 1.$$

Решение задачи о полой сфере. Решение задачи о толсто-стенной сфере под действием гидростатического давления в рамках деформационной теории Девиса – Надаи представлено в статье [4]. Из условия несжимаемости материала следует, что

$$u(\rho) = \sqrt[3]{\rho^3 + \alpha^3} - \rho,$$

где α^3 – произвольная постоянная интегрирования. При этом истинные деформации определяются выражениями

$$\varepsilon_{\theta\theta}(\rho) = \varepsilon_{\varphi\varphi}(\rho) = -\frac{1}{2} \varepsilon_{\rho\rho}(\rho) = \frac{1}{3} \ln \left(1 + \frac{\alpha^3}{\rho^3} \right) = \text{sign } u \frac{1}{2} \varepsilon_i.$$

Компоненты напряжений, согласно [4], вычисляются по формулам

$$\sigma_{\rho\rho}(\rho) = \text{sign } u \left(\frac{2}{3} \right)^{1+\mu} B \int_x^{\alpha} \frac{|\ln x|^\mu}{x^2 - x} dx + q,$$

$$\sigma_{\theta\theta}(\rho) = \sigma_{\varphi\varphi}(\rho) = \text{sign } u B \left(\frac{2}{3} |\ln x| \right)^\mu + \sigma_{\rho\rho}(\rho)$$

при краевом условии

$$q = \text{sign } u \left(\frac{2}{3} \right)^{1+\mu} B \int_{x_b}^{x_a} \frac{|\ln x|^\mu}{x^2 - x} dx, \quad (2)$$

$$\text{где } x = \frac{\rho^3 + \alpha^3}{\rho^3}, \quad x_a = \frac{a^3 + \alpha^3}{a^3}, \quad x_b = \frac{b^3 + \alpha^3}{b^3}.$$

Тривиальное решение. В случае сплошного шара (при $a = 0$) под действием внешней нагрузки q в теле может реализоваться однородное напряженное состояние $\sigma_{\rho\rho} = \sigma_{\theta\theta} = \sigma_{\varphi\varphi} = \sigma = q$. При этом задача имеет тривиальное решение $\varepsilon_i = \sigma_i = 0$. Перемещения точек среды в таком случае отсутствуют: $u = 0$, $\alpha^3 = 0$. В то же время, как видно из (1), напряжения $\sigma_{\rho\rho} = \sigma_{\theta\theta} = \sigma_{\varphi\varphi}$ могут (с математической точки зрения) неограниченно возрастать при увеличении q . Естественно, в действительности они возрастают до тех пор, пока не достигнут некоторого предельного значения. Ниже показано, что такое решение, начиная с некоторого q^* , не единственно.

Оценка критической нагрузки при порообразовании. Исследование подобной задачи для некоторых гиперупругих материалов показало существование бифуркационного решения, соответствующего образованию в теле шаровой полости. Один из возможных вариантов его построения – переход к пределу в решении для сферы с концентрической полостью при стремлении радиуса полости к нулю. Воспользуемся этим подходом. Зафиксируем внешний начальный радиус сферы $b = 1$ и построим ряд кривых деформирования при различных начальных размерах отверстия. На рис. 1 представлены зависимости $q/B - \delta V$, изображенные линиями 1–6, последовательно соответствующими значениям $a = 0,9; 0,7; 0,5; 0,3; 0,1; 0,01$ для $\mu = 2/3$. Здесь δV – абсолютное приращение объема полости. Все кривые 1–6 начинаются от нуля, быстро возрастают до максимума, а затем постепенно убывают. По мере уменьшения начального размера отверстия величина критической нагрузки растет, стремясь к некоторому пределу, а точка экстремума приближается к оси ординат. В пределе при $a \rightarrow 0$ получается бифуркационное решение (соответствующее сплошному шару), которое практически

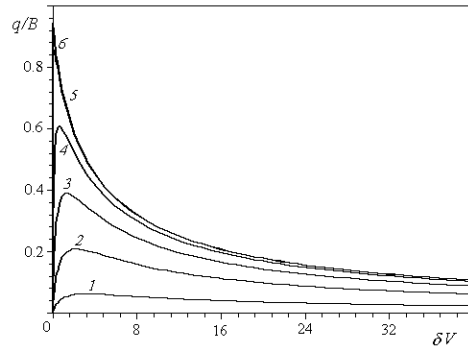


Рис. 1. Зависимости приращения объема поры от внешней нагрузки при различных начальных размерах поры

совпадает с убывающей ветвью кривой b . Таким образом, при нагружении шара с предсуществующей микропорой при достижении нагрузкой некоторого критического значения q^* начинается внезапный рост поры, по мере которого напряжения в теле падают.

Тем не менее с математической точки зрения вопрос об ограниченности критического значения нагрузки q^* при $a = 0$ остается открытым. В связи с этим докажем следующую теорему.

Теорема. *Образование поры в сплошном шаре конечных размеров происходит при конечном значении внешней нагрузки.*

Для доказательства этой теоремы достаточно получить верхнюю границу указанной нагрузки, которая определяется выражением (2) при $a = 0$, $b < \infty$, т.е. показать, что

$$\left(\frac{2}{3}\right)^{1+\mu} B \int_{x_b}^{\infty} \frac{(\ln x)^\mu}{x^2 - x} dx = q^* < \text{const} < \infty \quad \text{при } x_b > 1.$$

Из графика подынтегральной функции $f(x) = \frac{(\ln x)^\mu}{x^2 - x}$ видно, что на интервале $(1, \infty)$ она монотонно убывает, оставаясь при этом выше нуля. Для исследования сходимости указанного интеграла правомерно воспользоваться свойством определенных интегралов [5], согласно которому, если $f(x) \leq g(x)$ на $x \in [x_1, x_2]$, то

$\int_{x_1}^{x_2} f(x) dx \leq \int_{x_1}^{x_2} g(x) dx$. Теперь необходимо найти такую функцию $g(x)$, которая бы убывала на бесконечности. Можно проверить, что при $x > 1$ функция $f(x)$ ограничена сверху функцией $g(x)$:

$$f(x) = \frac{(\ln x)^\mu}{x^2 - x} < (x - 1)^{\mu-2} = g(x).$$

Вычислим интеграл от функции $g(x)$ для $\mu < 1$

$$\int_{x_b}^{\infty} g(x) dx = \int_{x_b}^{\infty} (x - 1)^{\mu-2} dx = \frac{1}{(1 - \mu)(x_b - 1)^{1-\mu}}.$$

В силу того что $x_b > 1$ и $\mu < 1$, последнее выражение ограничено, а следовательно, конечен и интеграл

$$q^* = \left(\frac{2}{3}\right)^{1+\mu} B \int_{x_b}^{\infty} \frac{(\ln x)^\mu}{x^2 - x} dx < \left(\frac{2}{3}\right)^{1+\mu} \frac{B}{(1 - \mu)(x_b - 1)^{1-\mu}} < \infty.$$

Для линейно-упругого тела, при $\mu = 1$, функцию $f(x)$ необходимо ограничить сверху несколько иначе (так как в предыдущем случае интеграл оказывается неограниченным), например,

$$f(x) = \frac{\ln x}{x^2 - x} < (x - 1)^{-3/2} = g(x).$$

Тогда, поскольку

$$\int_{x_b}^{\infty} g(x) dx = \int_{x_b}^{\infty} (x - 1)^{-3/2} dx = \frac{2}{(x_b - 1)^{1/2}},$$

то

$$q^* = \left(\frac{2}{3}\right)^2 B \int_{x_b}^{\infty} \frac{\ln x}{x^2 - x} dx < \frac{8B}{9(x_b - 1)^{1/2}} < \infty.$$

Таким образом, теорема доказана.

Значение критической нагрузки можно уточнить на основании следующего утверждения, вытекающего из теоремы подобия

А.А. Ильюшина: при фиксированном отношении $b/a = \text{const}$ максимальные значения внешней нагрузки q одинаковы. Отсюда следует, что усилия, необходимые для образования поры в теле конечных размеров, равны максимальному значению нагрузки для бесконечного тела с конечным отверстием при тех же механических характеристиках материала, поскольку в обоих случаях $b/a = \infty$. Для приближенной нижней оценки этой величины можно взять значение нагрузки, соответствующее конкретному достаточно большому относительному увеличению объема поры Δ в неограниченном теле, например,

$$q^* \approx \left(\frac{2}{3}\right)^{1+\mu} B \int_1^{10} \frac{(\ln x)^\mu}{x^2 - x} dx.$$

Таким образом, показано, что для сплошного шара существует два решения задачи: тривиальное и сингулярное. Согласно первому, напряжения в теле могут неограниченно возрастать при нулевых деформациях. Бифуркационное решение соответствует образованию в шаре сферической полости при достижении нагрузкой критического значения q^* . Для $q \geq q^*$ это решение является энергетически более выгодным.

Литература

1. Надаи А. Пластичность и разрушение твердых тел. Т. 1. М.: ИЛ, 1954. 647 с.
2. Hill R. The mathematical theory of plasticity. Oxford: Clarendon Press, 1950. 356 p.
3. Tvergard V., Huang Y., Hutchinson J.W. Cavitation instabilities in a power hardening elastic-plastic solids // Europ. J. of Mechanics. Ser. A: Solids, 1992. Vol. 11, № 2. P. 215–231.
4. Даль Ю.М., Пронина Ю.Г. Деформация шаровой поры в нелинейно-упругом теле // Изв. РАН. Сер. физическая, 2006. Т. 70, № 9. С. 1341–1343.
5. Смирнов В.И. Курс высшей математики. Т. 1. М.–Л.: Гос. изд-во ТТЛ, 1951. 472 с.

Степанова В.А.

Санкт-Петербургский государственный университет

Нелинейная задача для плоскости с прямолинейной межфазной трещиной¹

Рекомендовано к публикации профессором Мальковым В.М.

Введение. Методы теории функций комплексной переменной широко используются для решения плоских задач линейной теории упругости. Они особенно эффективны в сингулярных краевых задачах для областей с трещинами и сосредоточенными воздействиями [1]. Для высокоэластичных резино-подобных материалов большие деформации реализуются в упругой области деформирования, поэтому применение уравнений теории упругости здесь обосновано. Приведены некоторые соотношения нелинейной плоской задачи для материала Джона [2], и рассмотрена задача о действии нормального давления на берегах межфазной трещины в двухкомпонентной плоскости. Принятие модели гармонического материала позволило использовать преимущества метода комплексных потенциалов и получить точное решение задачи во всей области [3].

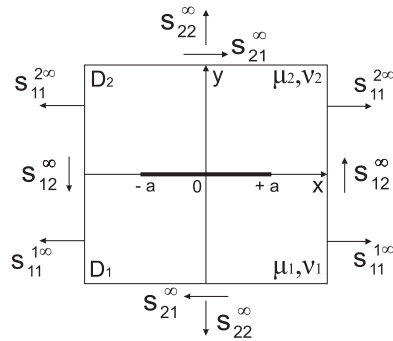


Рис. 1. Двухкомпонентная плоскость с прямолинейной межфазной трещиной

Соотношения плоской деформации. Рассматривается нелинейная задача плоской деформации двухкомпонентной пластины с прямолинейной межфазной трещиной (рис. 1). Концам трещины на

¹Работа выполнена при финансовой поддержке РФФИ, проект № 09-01-00656

линии сопряжения полуплоскостей $y = 0$ отвечают значения $x = \pm a$. На бесконечности заданы напряжения и углы поворота. Уравнения равновесия для тензора условных напряжений и уравнения совместности деформаций для градиента деформации имеют вид

$$\operatorname{div} \mathbf{S} = \mathbf{0}, \quad \operatorname{rot} \mathbf{G}^T = \mathbf{0}.$$

Граничные условия на берегах трещины

$$\mathbf{s}_n = \mathbf{n} \mathbf{S} = \mathbf{q}, \quad \mathbf{q} = p \mathbf{n} J \mathbf{G}^{-1},$$

где \mathbf{n} – внешняя нормаль к недеформированной поверхности тела. Специфика задачи состоит в том, что условия на трещине зависят от деформации, хотя величина давления p постоянна.

Значения условных напряжений на трещине в комплексном виде

$$s_{22} - is_{21} = p(g_{11} + ig_{21}). \quad (1)$$

Гармонический материал Джона. В качестве гармонического материала принята модель материала Джона [2], для которой упругий потенциал имеет вид

$$\Phi = 2\mu[F(L) - k], \quad L = \lambda_1 + \lambda_2, \quad k = \lambda_1\lambda_2, \quad (2)$$

где μ – модуль сдвига, L и k – инварианты тензора кратности удлинений, $F(L)$ – некоторый функционал, зависящий от первого инварианта.

Для материала (2) закон упругости для тензора условных напряжений в комплексной форме имеет вид

$$\begin{aligned} s_{11} + is_{12} &= 2\mu \left[\frac{2}{L} F'(L) \frac{\partial \zeta}{\partial z} - \frac{\partial \zeta}{\partial z} + \frac{\partial \zeta}{\partial \bar{z}} \right], \\ s_{22} - is_{21} &= 2\mu \left[\frac{2}{L} F'(L) \frac{\partial \zeta}{\partial z} - \frac{\partial \zeta}{\partial z} - \frac{\partial \zeta}{\partial \bar{z}} \right], \end{aligned} \quad (3)$$

где z и ζ – комплексные переменные отсчетной и текущей конфигураций. Введем комплексную функцию σ и напряжения представим выражениями

$$s_{11} + is_{12} = \frac{\partial \sigma}{\partial z} - \frac{\partial \sigma}{\partial \bar{z}}, \quad s_{22} - is_{21} = \frac{\partial \sigma}{\partial z} + \frac{\partial \sigma}{\partial \bar{z}}, \quad (4)$$

из (3) и (4) получим уравнения для искомым потенциалов

$$\frac{\partial \sigma}{\partial z} + 2\mu \frac{\partial \zeta}{\partial z} = 4\mu \frac{1}{L} F'(L) \frac{\partial \zeta}{\partial z}, \quad \frac{\partial \sigma}{\partial \bar{z}} + 2\mu \frac{\partial \zeta}{\partial \bar{z}} = 0. \quad (5)$$

На практике вместо функционала $F(L)$ обычно задают зависимость $L(|\varphi'(z)|)$. Далее будем использовать следующую модель материала [3]

$$L(|\varphi'(z)|) = 2|\varphi'(z)| \left[b + c \frac{1}{|\varphi'(z)|^2} \right].$$

Решения уравнений (5) имеют вид

$$\zeta = \overline{\psi(z)} + b\varphi(z) + \frac{cz}{\varphi'(z)},$$

$$\sigma = (1 - 2\mu b)\varphi(z) - 2\mu \overline{\psi(z)} - 2\mu \frac{cz}{\varphi'(z)},$$

где $\psi(z)$ и $\varphi(z)$ – неизвестные аналитические функции, а постоянные b и c однозначно определяются из условий перехода закона упругости в закон Гука при малых деформациях.

Формулы для напряжений и деформаций

$$s_{11} + is_{12} = (1 - 2\mu b)\varphi'(z) - 2\mu \frac{c}{\varphi'(z)} + 2\mu \overline{\psi'(z)} - 2\mu \frac{cz\overline{\varphi''(z)}}{\varphi'^2(z)},$$

$$s_{22} - is_{21} = (1 - 2\mu b)\varphi'(z) - 2\mu \frac{c}{\varphi'(z)} - 2\mu \overline{\psi'(z)} + 2\mu \frac{cz\overline{\varphi''(z)}}{\varphi'^2(z)},$$

$$g_{11} + ig_{21} = b\varphi'(z) + \frac{c}{\varphi'(z)} + \overline{\psi'(z)} - \frac{cz\overline{\varphi''(z)}}{\varphi'^2(z)},$$

$$g_{22} - ig_{12} = b\varphi'(z) + \frac{c}{\varphi'(z)} - \overline{\psi'(z)} + \frac{cz\overline{\varphi''(z)}}{\varphi'^2(z)}.$$

Из этих соотношений получим

$$(s_{22} - is_{21}) + 2\mu(g_{11} + ig_{21}) = \varphi'(z),$$

$$(s_{11} + is_{12}) + 2\mu(g_{22} - ig_{12}) = \varphi'(z);$$

$$s_{11} + s_{22} + i(s_{12} - s_{21}) = 2(1 - 2\mu b)\varphi'(z) - 4\mu \frac{c}{\varphi'(z)},$$

$$s_{11} - s_{22} + i(s_{12} + s_{21}) = 4\mu \overline{\psi'(z)} - 4\mu \frac{cz\overline{\varphi''(z)}}{\varphi'^2(z)}.$$

Плоскость с межфазной трещиной, нагруженной нормальным давлением. Учитывая формулы (8), из условий (1) получим

$$s_{22} - is_{21} = \frac{p}{2\mu + p} \varphi'(z), \quad g_{11} + ig_{21} = \frac{1}{2\mu + p} \varphi'(z). \quad (10)$$

Введем две новые функции комплексной переменной $h(z)$ и $r(z)$, которые аналитичны во всей плоскости, исключая линию раздела [3]. Граничные условия на трещине примут вид

$$h^+(x_1) - h^-(x_1) = \Delta s(x_1), \quad r^+(x_1) - r^-(x_1) = \Delta g(x_1). \quad (11)$$

Комплексные потенциалы выразим через функции $h(z)$ и $r(z)$:

- верхняя полуплоскость

$$\begin{aligned} \varphi_2'(z) &= \frac{(h + 2\mu_1 r)(z)}{1 + 2(\mu_1 - \mu_2)b_2}, \\ \frac{c_1}{\varphi_1'(z)} + \overline{\psi_1'(z)} - \frac{c_1 z \overline{\varphi_1''(z)}}{\overline{\varphi_1'^2(z)}} &= \frac{[b_2 h - (1 - 2\mu_2 b_2) r](z)}{1 + 2(\mu_1 - \mu_2)b_2}, \end{aligned} \quad (12)$$

- нижняя полуплоскость

$$\begin{aligned} \varphi_1'(z) &= \frac{(h + 2\mu_2 r)(z)}{1 + 2(\mu_2 - \mu_1)b_1}, \\ \frac{c_2}{\varphi_2'(z)} + \overline{\psi_2'(z)} - \frac{c_2 z \overline{\varphi_2''(z)}}{\overline{\varphi_2'^2(z)}} &= \frac{[b_2 h - (1 - 2\mu_2 b_2) r](z)}{1 + 2(\mu_2 - \mu_1)b_1}. \end{aligned} \quad (13)$$

Подставив (10) в (11), получим

$$h^+ - h^- = p(r^+ - r^-) = \frac{p}{2\mu_2 + p} [\varphi_2'(z)]^+ - \frac{p}{2\mu_1 + p} [\varphi_1'(z)]^-.$$

Откуда следует

$$\begin{aligned} h(z) - pr(z) &= h(\infty) - pr(\infty), \\ h(\infty) &= 2\mu_2(1 - 2\mu_2 b_2) + 2\mu_1(1 - 2\mu_1 b_1), \\ r(\infty) &= 2\mu_2 b_2 + 2\mu_1 b_1 - 1. \end{aligned}$$

Чтобы получить граничную задачу для одной из функций $h(z)$ или $r(z)$, запишем условия на трещине

$$\begin{aligned}(s_{22} - is_{21})^+ &= \frac{p}{2\mu_2 + p} [\varphi_2'(z)]^+, \\(s_{22} - is_{21})^- &= \frac{p}{2\mu_1 + p} [\varphi_1'(z)]^-. \end{aligned} \quad (14)$$

Заменим напряжения в (14) на выражения (6), сложим уравнения и в полученном уравнении заменим комплексные функции $\varphi_1'(z)$ и $\varphi_2'(z)$ на выражения (12) и (13).

$$r^+(x_1) + \alpha r^-(x_1) = d, \quad (15)$$

$$\begin{aligned}\alpha &= \frac{p + 2\mu_2}{p + 2\mu_1} \frac{1 + 2(\mu_1 - \mu_2)b_2}{1 + 2(\mu_2 - \mu_1)b_1} \frac{1 - (p + 2\mu_1)b_1}{1 - (p + 2\mu_2)b_2}, \\d &= -\frac{1 - (p + 2\mu_1)b_1 - (p + 2\mu_2)b_2}{(p + 2\mu_1)[1 - (p + 2\mu_2)b_2][1 + 2(\mu_2 - \mu_1)b_1]} [h(\infty) - pr(\infty)].\end{aligned}$$

Решение уравнения (15) имеет вид

$$r(z) = \frac{d}{1 + \alpha} [1 - (z - 2i\beta a)X_0(z)] + X_0(z)(A_0z + B_0), \quad (16)$$

$$X_0(z) = \frac{1}{\sqrt{z^2 - a^2}} \left(\frac{z + a}{z - a} \right)^{i\beta}, \quad \beta = \frac{\ln \alpha}{2\pi}. \quad (17)$$

В результате решения задачи получены напряжения и деформации для каждой из полуплоскостей, а также на трещине. Найдены формулы перемещений материала на трещине. Расчеты проведены для однородной плоскости со следующими заданными параметрами: $\mu_1 = \mu_2 = 10 \text{ кг/см}^2$, $\lambda_1 = \lambda_2 = 25 \cdot 10^3 \text{ кг/см}^2$, $p = 5 \text{ кг/см}^2$, $a = 1$.

На рис. 2 пунктиром показан график перемещений берегов трещины, полученный по линейной теории, а сплошной кривой – по нелинейной. Перемещения в первом случае получены по формуле

$$u_2 = -\frac{p}{\mu} \frac{\lambda + 2\mu}{(\lambda + \mu)} \sqrt{1 - x^2}.$$

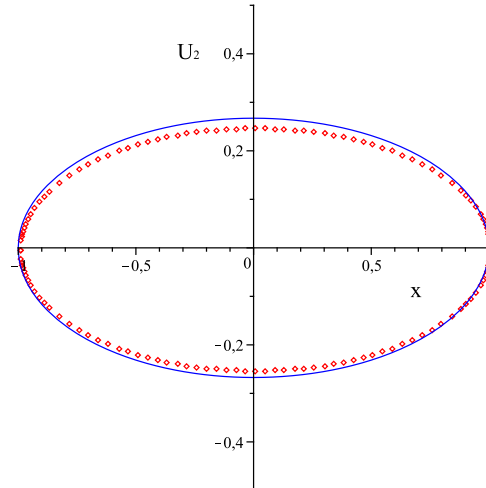


Рис. 2. Перемещения берегов трещины, полученные по линейной и нелинейной теориям

В случае нелинейной теории перемещения получены путем интегрирования второй формулы в (10) по переменной x , $z = x + iy$, на трещине $y = 0$. Перемещения не имеют особенности у концов трещины. Из формулы (17) для функции $X_0(z)$ следует, что напряжения и перемещения имеют осцилляцию у концов межфазной трещины. Для однородной плоскости с трещиной осцилляции нет.

Литература

1. Мусхелишвили Н.И. Некоторые основные задачи математической теории упругости. М.: Наука, 1966. 708 с.
2. John F. Plane strain problems for a perfectly elastic material of a harmonic type // Communications on pure and applied mathematics, 1960. Vol. XIII, P. 239–290.
3. Мальков В.М., Малькова Ю.В. Плоская задача нелинейной теории упругости для гармонического материала // Вестник СПбГУ. Сер. 1, 2008. Вып. 3. С. 114–126.

Телевный Д.С.

Санкт-Петербургский государственный университет

Моделирование триодной системы с полевым острием

Рекомендовано к публикации доцентом Виноградовой Е.М.

Введение. Эмиссионные устройства на основе полевого катода применяются в электронных микроскопах, плоских дисплеях на основе полевой эмиссии, системах диагностики поверхности, приборах микро- и нанoeлектроники. Таким устройствам характерны малые размеры эмитирующей области, компактность, экономичность (отсутствие расхода энергии на принудительный нагрев) [1].

Процесс эмиссии определяют: форма полевого катода, радиус кривизны острия, расположение электродов в системе. Также используют дополнительный электрод – модулятор, который позволяет при небольших потенциалах влиять на эффективность эмиссии.

Физическая постановка задачи. Рассмотрим триодную цилиндрическую систему, состоящую из плоского катода, модулятора в виде круговой диафрагмы и плоского анода (рис. 1). Внутренняя область системы заполнена двумя различными диэлектриками с диэлектрическими проницаемостями ε_1 и ε_2 . На оси системы расположено острие, моделируемое заряженной нитью длиной z_0 . Катод имеет нулевой потенциал, на модуляторе задан потенциал V_1 , на аноде – потенциал V_2 . Требуется найти функцию распределения потенциала в области триодной системы с полевым острием.

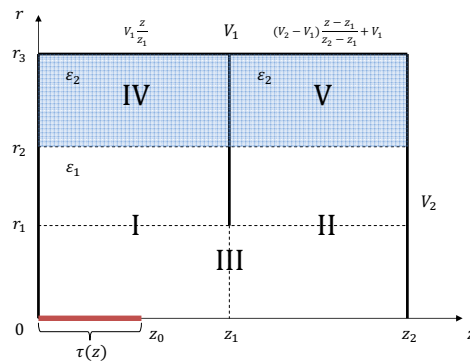


Рис. 1. Сечение системы с учетом аксиальной симметрии

Математическая постановка задачи. Введем некоторую совокупность зарядов, которые будут представлять заряженную нить вдоль оси z с некоторой неизвестной плотностью $\tau(z)$. Тогда, учитывая аксиальную симметрию, уравнение для потенциала запишется в виде уравнения Пуассона

$$\Delta U(r, z) = -\frac{\rho(r, z)}{\varepsilon_0}, \quad (1)$$

с граничными условиями

$$\begin{aligned} U(r, 0) &= 0, \quad r \in (0, r_3), \\ U(r, z_1) &= V_1, \quad r \in (r_1, r_3), \\ U(r, z_2) &= V_2, \quad r \in (0, r_3), \\ U(r_3, z) &= V_1 \frac{z}{z_1}, \quad z \in (0, z_1), \\ U(r_3, z) &= \frac{V_2 - V_1}{z_2 - z_1} (z - z_1) + V_1, \quad z \in (z_1, z_2), \end{aligned} \quad (2)$$

где $\rho(r, z)$ – объемная плотность заряда, создаваемая заряженной нитью.

Решение уравнения (1) представим как сумму решения уравнения Лапласа с граничными условиями (2) и решения уравнения Пуассона с однородными граничными условиями. Первое было найдено [2], поэтому будем искать решение второго.

Представим неизвестную функцию $U(r, z)$ в уравнении (1) в виде [3]

$$U(r, z) = \sum_{n=1}^{\infty} V_n(z) J_0 \left(\frac{\omega_n}{a} r \right), \quad (3)$$

где $V_n(z)$ – неизвестная функция, a – граница области по r . Будем искать распределение потенциала в первой и третьей областях, так как только в этих областях присутствует заряженная нить.

Подставляя (3) в (1) и пользуясь свойством ортогональности функции Бесселя J_0 , приходим к неоднородному дифференциальному уравнению второго порядка

$$V_m''(z) - \left(\frac{\omega_m}{a} \right)^2 V_m(z) = -\frac{2}{a^2 J_1^2(\omega_m)} \int_0^{r_0} r \frac{\rho(r, z)}{\varepsilon_0} J_0 \left(\frac{\omega_m}{a} r \right) dr,$$

где r_0 – предполагаемый радиус острия ($0 < r_0 \ll a$).

Представим объемную плотность в следующем виде

$$\rho(r, z) = \frac{\tau(z) \delta(r - r_0)}{2\pi r}.$$

Тогда при достаточно малом значении r_0 функция $J_0\left(\frac{\omega_m}{a}r_0\right) \approx 1$.

$$\varphi(z) = \begin{cases} -\frac{1}{a^2 J_1^2(\omega_m)} \frac{\tau(z)}{2\pi\epsilon_0}, & 0 < z < z_0, \\ 0, & z \geq z_0, \end{cases}$$

$$V_m''(z) - \left(\frac{\omega_m}{a}\right)^2 V_m(z) = \varphi(z).$$

Решение уравнения (5) будем искать в виде

$$V_m(z) = A_m(z) \operatorname{ch}\left(\frac{\omega_m}{a}z\right) + B_m(z) \operatorname{sh}\left(\frac{\omega_m}{a}z\right).$$

Функции $A_m(z)$ и $B_m(z)$ при использовании метода вариации постоянных должны определяться из уравнений

$$\begin{aligned} A_m' \operatorname{ch}\left(\frac{\omega_m}{a}z\right) + B_m' \operatorname{sh}\left(\frac{\omega_m}{a}z\right) &= 0, \\ \frac{\omega_m}{a} \left(A_m' \operatorname{ch}\left(\frac{\omega_m}{a}z\right) + B_m' \operatorname{sh}\left(\frac{\omega_m}{a}z\right) \right) &= \varphi(z). \end{aligned}$$

Определяя вид функции $V_m(z)$ с использованием однородных граничных условий и учитывая уравнение (3), получим распределение потенциала в I и III областях в общем виде

$$U(r, z) = \begin{cases} -\sum_{l=1}^{\infty} \left(\frac{a \operatorname{sh}\left(\frac{\omega_l}{a}(z^* - z)\right)}{\omega_l \operatorname{sh}\left(\frac{\omega_l}{a}z^*\right)} \int_0^{z_0} \operatorname{sh}\left(\frac{\omega_l}{a}\eta\right) \varphi(\eta) d\eta + \right. \\ \left. + \frac{a}{\omega_l} \int_z^{z_0} \operatorname{sh}\left(\frac{\omega_l}{a}\eta\right) \varphi(\eta) d\eta \right) \times J_0\left(\frac{\omega_l}{a}r\right), & z < z_0, \\ -\sum_{l=1}^{\infty} \left(\frac{a \operatorname{sh}\left(\frac{\omega_l}{a}(z^* - z)\right)}{\omega_l \operatorname{sh}\left(\frac{\omega_l}{a}z^*\right)} \int_0^{z_0} \operatorname{sh}\left(\frac{\omega_l}{a}\eta\right) \varphi(\eta) d\eta \right) \times \\ \times J_0\left(\frac{\omega_l}{a}r\right), & z \geq z_0, \end{cases} \quad (4)$$

где z^* – граница области по z .

Разложим неизвестную функцию $\tau(\eta)$ в ряд Фурье со следующими граничными условиями $\tau(0) = 0$ и $\tau(z_0) = 0$. Выражение для функции $\varphi(\eta)$ будет иметь вид

$$\varphi(\eta) = -\frac{1}{2\pi\varepsilon_0 a \omega_l J_1^2(\omega_l)} \sum_{s=1}^{\infty} \tau_s \sin\left(\frac{\pi s}{z_0} \eta\right). \quad (5)$$

Подставляем (5) в (4) и интегрируя в указанных пределах, получим решение уравнения Пуассона с однородными граничными условиями

$$U(r, z) = \sum_{s_1=1}^{\infty} \sum_{l=1}^{\infty} \frac{\tau_{s_1}}{2\pi\varepsilon_0 a \omega_l J_1^2(\omega_l)} \times$$

$$\times \left[\frac{\frac{\omega_l}{a} \sin\left(\frac{\pi s_1}{z_0} z\right)}{\left(\frac{\pi s_1}{z_0}\right)^2 + \left(\frac{\omega_l}{a}\right)^2} - \frac{(-1)^{s_1} \frac{\pi s_1}{z_0} \operatorname{sh}\left(\frac{\omega_l}{a}(z^* - z_0)\right)}{\operatorname{sh}\left(\frac{\omega_l}{a} z^*\right)} \frac{\operatorname{sh}\left(\frac{\omega_l}{a} z\right)}{\left(\frac{\pi s_1}{z_0}\right)^2 + \left(\frac{\omega_l}{a}\right)^2} \right] \times$$

$$\times J_0\left(\frac{\omega_l}{a} r\right), \quad z < z_0,$$

$$U(r, z) = \sum_{s_1=1}^{\infty} \sum_{l=1}^{\infty} \frac{\tau_{s_1}}{2\pi\varepsilon_0 a \omega_l J_1^2(\omega_l)} \frac{(-1)^{s_1+1} \frac{\pi s_1}{z_0} \operatorname{sh}\left(\frac{\omega_l}{a} z_0\right) \operatorname{sh}\left(\frac{\omega_l}{a}(z^* - z)\right)}{\left(\frac{\pi s_1}{z_0}\right)^2 + \left(\frac{\omega_l}{a}\right)^2 \operatorname{sh}\left(\frac{\omega_l}{a} z^*\right)} \times$$

$$\times J_0\left(\frac{\omega_l}{a} r\right), \quad z \geq z_0.$$

В результате получили функции распределения потенциала с учетом введенных зарядов в двух областях. Используя метод перекрытия и условия непрерывности нормальной составляющей вектора электрического смещения [4, 2], вычислим соотношения между неизвестными коэффициентами, входящие в уравнения для потенциала. Также введем два дополнительных уравнения, аппроксимирующие форму острия, для полноты системы коэффициентов.

Решая такую систему, определим значения неизвестных коэффициентов и получим функции для распределения потенциала во всей области системы.

Представление численных результатов. На основе проведенных теоретических выкладок написана программа, которая используется для решения системы относительно неизвестных коэффициентов. В результате расчетов получено следующее распределение потенциала в системе с представленными параметрами (рис. 2): $r_1 = 0,5$; $r_2 = 0,75$; $r_3 = 1$; $z_1 = 0,5$; $z_2 = 1$; $\varepsilon_1 = 1$; $\varepsilon_2 = 4,75$; $V_1 = 1$; $V_2 = 1$; $z_0 = 0,4$. Для численных расчетов использовались безразмерные единицы измерения. Форма острия задается в виде эллипса.

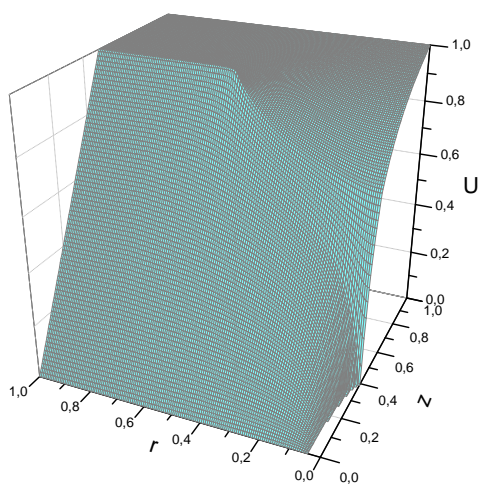


Рис. 2. Распределение электростатического потенциала во всей области системы

Из рис. 2 видно выполнение граничных условий и нулевую эквипотенциальную линию острия.

На рис. 3. изображены эквипотенциальные линии потенциала системы. Жирной линией обозначена эквипотенциальная линия острия, которая определяет форму катода.

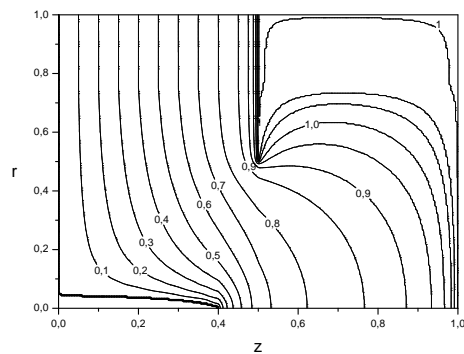


Рис. 3. Эквипотенциальные линии электростатического потенциала во всей области системы

Заключение. В данной работе рассмотрена задача моделирования триодной электронно-оптической системы с полевым острием. Решено уравнение Пуассона. Получены в аналитическом виде функции распределения потенциала во всей области системы. По полученным результатам численного эксперимента построены поверхность распределения потенциала и представлен график эквипотенциальных линий.

Литература

1. David S.Y. Microgating carbon nanotube field emitters by in situ growth inside open aperture arrays // Applied physics letters, 2002. Vol. 80. P. 2988–2990.
2. Телевный Д.С. Расчет электростатического потенциала в электронно-оптической системе с модулятором // Процессы управления и устойчивость: Труды 41-й международной научной конференции аспирантов и студентов / Под ред. Н. В. Смирнова, Г. Ш. Тамасяна. СПб.: Издат. Дом С.-Петерб. гос. ун-та, 2010. С. 225–230.
3. Миролобов Н.Н. Методы расчета электростатических полей. М.: Изд-во Высшая Школа, 1963. 209 с.
4. Виноградова Е.М. Математическое моделирование электронно-оптических систем. СПб.: Изд-во СПбГУ, 2005. 110 с.

Уланов Е.А., Утешев А.Ю.
Санкт-Петербургский государственный университет

Аналитическое решение обобщенной задачи Ферма – Торричелли – Штейнера

1. Постановка задачи. На плоскости даны точки $\mathcal{P}_1, \mathcal{P}_2, \mathcal{P}_3$, не лежащие на одной прямой, заданы также положительные числа m_1, m_2, m_3 . Требуется определить точку \mathcal{P} плоскости, для которой сумма взвешенных расстояний $m_1|\mathcal{P}_1\mathcal{P}| + m_2|\mathcal{P}_2\mathcal{P}| + m_3|\mathcal{P}_3\mathcal{P}|$ будет наименьшей.

Эта задача под названием «задача об оптимальном расположении станции» была поставлена в 1846 году [1, 2]. Она представляет собой обобщение более известной в литературе [1, 3] проблемы Ферма – Торричелли – Штейнера, получающейся из нее в частном случае при $m_1 = m_2 = m_3 = 1$. Решение обеих задач традиционно [1–3] приводится с использованием либо геометрических построений, либо механической модели «грузы на нитях». Ни в одной публикации авторы не нашли представления координат оптимальной точки \mathcal{P} в виде явных функций координат точек $\mathcal{P}_1, \mathcal{P}_2, \mathcal{P}_3$. Именно в такой постановке задача решается в данной статье.

2. Координаты оптимальной точки. Пусть (x_i, y_i) – координаты точки \mathcal{P}_i . Тогда целевой функцией задачи оптимизации является

$$f(x, y) = \sum_{i=1}^3 m_i \sqrt{(x - x_i)^2 + (y - y_i)^2}. \quad (1)$$

Координаты ее стационарной точки удастся определить с помощью подходящей замены переменных из системы линейных уравнений, приведенной в [4].

2.1. Общий случай. Веса m_1, m_2, m_3 произвольны.

Теорема 1. Координаты стационарной точки функции (1) задаются формулами

$$x = \frac{-n_0(k_{x,1} + k_{x,2}) - \mathcal{M}\mathcal{X}\mathcal{T}}{2d\sqrt{n_0}}, \quad y = \frac{n_0(k_{y,1} + k_{y,2}) - \mathcal{M}\mathcal{Y}\mathcal{T}}{2d\sqrt{n_0}} \quad (2)$$

при соответствующем критическом значении, равном \sqrt{d} . Здесь

$$d = \sqrt{n_0}|\mathcal{S}| - (m_1^2 r_{23} + m_2^2 r_{13} + m_3^2 r_{12}) + \frac{1}{2}(m_1^2 + m_2^2 + m_3^2)(r_{12} + r_{13} + r_{23}), \quad (3)$$

r_{12}, r_{13}, r_{23} – квадраты длин сторон треугольника $\mathcal{P}_1\mathcal{P}_2\mathcal{P}_3$:

$$r_{23} = (x_2 - x_3)^2 + (y_2 - y_3)^2, \quad r_{13} = (x_1 - x_3)^2 + (y_1 - y_3)^2, \\ r_{12} = (x_2 - x_1)^2 + (y_2 - y_1)^2,$$

$$\mathcal{S} = \begin{vmatrix} 1 & 1 & 1 \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{vmatrix}, \quad (4)$$

$$n_0 = -m_1^4 - m_2^4 - m_3^4 + 2(m_1^2 m_2^2 + m_3^2 m_2^2 + m_1^2 m_3^2), \quad (5)$$

$$k_{x,1} = \begin{vmatrix} 1 & 1 & 1 \\ y_1 & y_2 & y_3 \\ y_1^2 & y_2^2 & y_3^2 \end{vmatrix}, \quad k_{x,2} = \begin{vmatrix} 1 & 1 & 1 \\ x_1 & x_2 & x_3 \\ x_1 y_1 & x_2 y_2 & x_3 y_3 \end{vmatrix}, \\ k_{y,1} = \begin{vmatrix} 1 & 1 & 1 \\ x_1 & x_2 & x_3 \\ x_1^2 & x_2^2 & x_3^2 \end{vmatrix}, \quad k_{y,2} = \begin{vmatrix} 1 & 1 & 1 \\ y_1 & y_2 & y_3 \\ x_1 y_1 & x_2 y_2 & x_3 y_3 \end{vmatrix}, \\ \mathcal{X} = \begin{pmatrix} x_1 - x_2 - x_3 & x_3 & x_2 \\ x_3 & -x_1 + x_2 - x_3 & x_1 \\ x_2 & x_1 & -x_1 - x_2 + x_3 \end{pmatrix}, \\ \mathcal{Y} = \begin{pmatrix} y_1 - y_2 - y_3 & y_3 & y_2 \\ y_3 & -y_1 + y_2 - y_3 & y_1 \\ y_2 & y_1 & -y_1 - y_2 + y_3 \end{pmatrix}, \\ \mathcal{M} = (m_1^2, m_2^2, m_3^2), \quad \mathcal{R} = (r_{23}, r_{13}, r_{12}), \\ \mathcal{T} = \mathcal{S}\mathcal{M}^\top - \sqrt{n_0}\mathcal{R}^\top.$$

Дополнительно предполагается, что треугольник $\mathcal{P}_1\mathcal{P}_2\mathcal{P}_3$ обходится по часовой стрелке, т.е. величина (4) отрицательна и сумма любых двух весов из m_1, m_2, m_3 больше третьего. При условии нахождения точки (2) внутри треугольника $\mathcal{P}_1\mathcal{P}_2\mathcal{P}_3$ в этой точке достигается минимум функции f .

Проанализируем формулу (3) для квадрата критического значения минимизируемой функции. Входящее в ее состав выражение для n_0 (5) допускает разложение на линейные множители:

$$n_0 \equiv (m_1 + m_2 + m_3)(m_1 + m_2 - m_3)(m_1 - m_2 + m_3)(-m_1 + m_2 + m_3).$$

Если обозначить $\mathfrak{P} = \frac{1}{2}(m_1 + m_2 + m_3)$, то

$$\sqrt{n_0} = 4\sqrt{\mathfrak{P}(\mathfrak{P} - m_1)(\mathfrak{P} - m_2)(\mathfrak{P} - m_3)},$$

и радикал в правой части равенства имеет вид формулы Герона площади треугольника с длинами сторон, равными m_1, m_2, m_3 . Если учесть, что сомножитель при $\sqrt{n_0}$ в формуле (3) представляет величину удвоенной площади треугольника $\mathcal{P}_1\mathcal{P}_2\mathcal{P}_3$, то можно сделать вывод об абсолютной симметрии выражения для критического значения функции f относительно входящих в него наборов весов и длин: оптимизационная задача, для которой стороны треугольника имеют длины m_1, m_2, m_3 , а вес, соответствующий вершине между сторонами m_i и m_j , взят равным $\sqrt{r_{ij}}$, будет иметь то же критическое значение для целевой функции.

2.2. Случай одинаковых весов. Пусть теперь $m_1 = m_2 = m_3 = 1$.

Теорема 2. Координаты стационарной точки функции $f = \sum_{i=1}^3 \sqrt{(x - x_i)^2 + (y - y_i)^2}$ задаются формулами

$$x = \frac{\sqrt{3}(x_1r_{23} + x_2r_{13} + x_3r_{12}) - (x_1 + x_2 + x_3)\mathcal{S} - 3(k_{x,1} + k_{x,2})}{2\sqrt{3}d},$$

$$y = \frac{\sqrt{3}(y_1r_{23} + y_2r_{13} + y_3r_{12}) - (y_1 + y_2 + y_3)\mathcal{S} + 3(k_{y,1} + k_{y,2})}{2\sqrt{3}d},$$

при соответствующем критическом значении, равном \sqrt{d} . Здесь

$$d = \frac{1}{2}(r_{12} + r_{13} + r_{23}) + \sqrt{3}|\mathcal{S}|.$$

Все остальные обозначения совпадают с использованными в теореме 1. Дополнительно предполагается, что треугольник $\mathcal{P}_1\mathcal{P}_2\mathcal{P}_3$ обходится по часовой стрелке.

При условии, что величины всех углов треугольника меньше $\frac{2\pi}{3}$, стационарная точка лежит внутри треугольника, она совпадает с

точкой Ферма – Торричелли – Штейнера и в ней достигается минимум целевой функции.

3. Случай переменных весов. Изучим теперь влияние параметров задачи на положение стационарной точки. В случае одинаковых весов этими параметрами выступают координаты вершин треугольника. Динамика стационарной точки при изменении координат одной вершины известна [3]: стационарная точка остается внутри треугольника, если величина всех его углов остается меньше $\frac{2\pi}{3}$.

Если же величина одного угла, возрастая, стремится к $\frac{2\pi}{3}$, то стационарная точка движется в сторону соответствующей вершины и попадает в нее, когда величина угла становится равной $\frac{2\pi}{3}$. При дальнейшем увеличении угла, стационарная точка покидает треугольник и перестает соответствовать минимальному значению целевой функции — в этом случае последнее достигается в вершине треугольника, соответствующей тупому углу.

В общем случае переменных весов анализ динамики стационарной точки усложняется. Мы приведем его в частном случае однопараметрической зависимости.

Пусть увеличивается вес m_1 при неизменных m_2, m_3 . Стационарная точка стремится к вершине \mathcal{P}_1 . Формулы (2) позволяют найти условие на массу m_1 , при выполнении которого эти точки совпадут:

$$r_{12}r_{13}(m_1^2 - m_2^2 - m_3^2)^2 - 4m_2^2m_3^2[(x_2 - x_1)(x_3 - x_1) + (y_2 - y_1)(y_3 - y_1)]^2 = 0.$$

Оно эквивалентно

$$m_1^2 = m_2^2 + m_3^2 \pm 2m_2m_3 \cos(\widehat{\mathcal{P}_3\mathcal{P}_1\mathcal{P}_2}).$$

При выполнении условия

$$m_1^2 < m_2^2 + m_3^2 - 2m_2m_3 |\cos(\widehat{\mathcal{P}_3\mathcal{P}_1\mathcal{P}_2})|$$

стационарная точка будет находиться внутри треугольника и, следовательно, будет обеспечивать минимум целевой функции. При нарушении этого неравенства минимум достигается при помещении точки \mathcal{P} в вершину \mathcal{P}_1 (см. рис. 1).

Пусть теперь при неизменном m_1 синхронно увеличиваются веса m_2 и m_3 : $m_2 = m_3 = t$. Переходя в (2) к пределу при $t \rightarrow \infty$,

получаем, что координаты стационарной точки стремятся к

$$x_* = \frac{x_2 + x_3}{2} + \frac{x_2 - x_3}{2} \frac{r_{13} - r_{12}}{r_{23}}, \quad y_* = \frac{y_2 + y_3}{2} + \frac{y_2 - y_3}{2} \frac{r_{13} - r_{12}}{r_{23}},$$

которые задают точку \mathcal{P}_* на прямой $\mathcal{P}_2\mathcal{P}_3$, являющуюся основанием высоты, опущенной на сторону треугольника из вершины \mathcal{P}_1 .

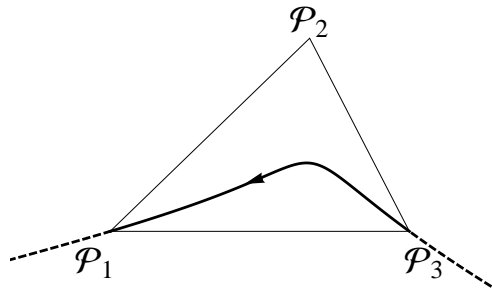


Рис. 1. Динамика стационарной точки при увеличении веса в одной вершине

В отличие от предыдущего случая эта точка не достигается при конечных величинах m . В случае, если \mathcal{P}_* принадлежит отрезку $\mathcal{P}_2\mathcal{P}_3$, то стационарная точка может подойти к \mathcal{P}_* сколь угодно близко, но только при неограниченном возрастании m . В случае, если \mathcal{P}_* не принадлежит отрезку $\mathcal{P}_2\mathcal{P}_3$ (один из углов $\widehat{\mathcal{P}_1\mathcal{P}_2\mathcal{P}_3}$ или $\widehat{\mathcal{P}_1\mathcal{P}_3\mathcal{P}_2}$ – тупой), то стационарная точка при некотором конечном значении m попадет в вершину тупого угла (см. рис. 2).

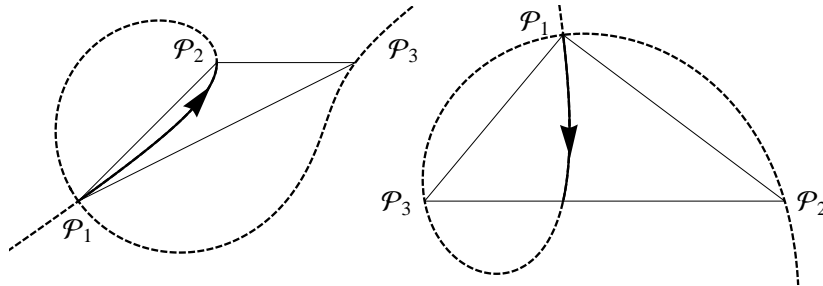


Рис. 2. Динамика стационарной точки при синхронном увеличении весов в двух вершинах

В [4] для этих кривых указывались алгебраические представления вида $F(x, y) = 0$, при F – полиноме третьей степени. Эти урав-

нения получаются из параметрического задания (2) исключением параметра (m_1 – в первом случае, m – во втором случае). Пунктиром показаны участки кривых, выходящие за пределы треугольника – они не имеют «физического смысла».

4. Заключение. Найдены аналитические выражения для координат стационарной точки для суммы взвешенных расстояний обобщенной задачи Ферма – Торричелли – Штейнера. С их помощью проведен анализ геометрического расположения точки в зависимости от параметров задачи.

Литература

1. Martini H. Fermat–Torricelli problem [Электронный ресурс] // Springer Online Reference Works [сайт]. [2001]. URL: <http://eom.springer.de/f/f130050.htm> (дата обращения: 25.02.2011)
2. Launhardt W. Kommercielle Tracirung der Verkehrswege // Zeitschrift f. Architekten u.Ingenieur-Vereinis im Konigreich Hannover, 1872. Т. 18. Р. 516–534.
3. Курант Р., Роббинс Г. Что такое математика? М.: МЦНМО, 2007. 568 с.
4. Уланов Е. А. Обобщение задачи Ферма–Торричелли–Штейнера // Процессы управления и устойчивость: Труды 41-й международной научной конференции аспирантов и студентов / Под ред. Н.В. Смирнова, Г.Ш. Тамасяна. СПб.: Издат. Дом С.-Петерб. гос. ун-та, 2010. С. 237–242.

Целищев С.О., Еремин А.С.
Санкт-Петербургский государственный университет

Об устойчивости структурного метода решения систем ОДУ

Введение. Эффективность специальных методов решения структурно разделённых систем обыкновенных дифференциальных уравнений (СОДУ) (см., напр., [1–3]), получивших название структурных методов, по критерию «вычислительные затраты/точность» была наглядно продемонстрирована на численных примерах в работах [4, 5]. Однако остаётся открытым вопрос о правомочности применения данных методов с точки зрения численной устойчивости получаемых решений.

Как известно, численный метод решения СОДУ, имеющей устойчивые точные решения, должен, хотя бы при некоторой длине шага интегрирования, давать устойчивое численное решение. Обычно, устойчивость определяется путём рассмотрения так называемого тестового уравнения Далквиста [6]

$$y'(x) = \lambda y(x), \quad (1)$$

для которого достаточно просто получить зависимость приближения $\bar{y}(x+h)$ одношаговым методом к решению $y(x+h)$ от начального условия $y(x)$ в виде

$$\bar{y}(x+h) = R(h\lambda)y(x). \quad (2)$$

Очевидно, что метод будет давать устойчивое численное решение в случае выполнения условия $|R(h\lambda)| \leq 1$.

Областью устойчивости метода называют область на комплексной плоскости, для которой выполняется неравенство $|R(z)| \leq 1$. Она показывает, как при фиксированном собственном числе λ ограничен обеспечивающий устойчивость метода шаг h . Как правило, интерес представляет левая полуплоскость, соответствующая при $h > 0$ собственному числу λ с неположительной действительной частью, т. е. устойчивому решению.

Система для анализа. Структурные методы, применяемые к скалярному уравнению, и в частности к (1), равносильны некоторым

классическим методам и не демонстрируют своей эффективности. В силу этого необходимо найти некоторую систему ОДУ, которая, во-первых, проявит свойства структурных методов, а во-вторых, позволит исследовать их устойчивость.

Дж. Дорманд с соавторами в [7] рассматривают устойчивость метода типа Ньюстрёма на примере уравнения

$$y''(x) = \lambda^2 y(x). \quad (3)$$

При этом, вероятно, начальные условия выбраны так, что решение не содержит одновременно компонентов, соответствующих обоим собственным числам $\pm\lambda$. Расчёт устойчивости метода, представленного в [7], при таком предположении дал результаты, полностью совпадающие с полученными в статье.

Отметим, что уравнение (3) легко сводится к системе первого порядка

$$\begin{cases} y_1'(x) = y_2(x), \\ y_2'(x) = \lambda^2 y_1(x), \end{cases}$$

которая равносильна системе

$$\begin{cases} y_1'(x) = \lambda y_2(x), \\ y_2'(x) = \lambda y_1(x). \end{cases} \quad (4)$$

Система (4), при предположении о согласованности начальных условий, подходит для исследования устойчивости структурных методов, так как они проявляют свою эффективность при применении к ней.

В [3] даны расчётные схемы различных структурных методов. Рассмотрим методы порядков со второго по четвёртый применительно к рассматриваемой системе.

Метод второго порядка. Поскольку интегрирование уравнений системы (4) структурным методом ведётся по-разному, то логично ожидать получение разных функций устойчивости для первого и для второго уравнений. Любой классический двухэтапный метод типа Рунге – Кутты второго порядка имеет функцию устойчивости [6]

$$R_{2,K}(z) = 1 + z + \frac{z^2}{2}, \quad (5)$$

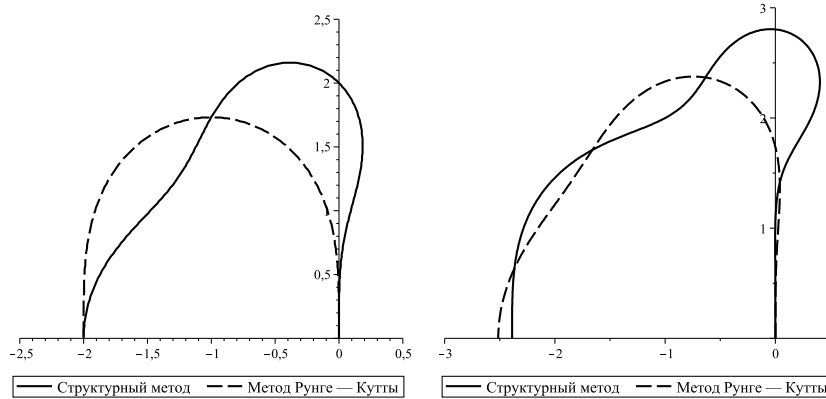


Рис. 1. Области устойчивости методов 2-го порядка для первого уравнения системы (4)

Рис. 2. Области устойчивости методов 3-го порядка для второго уравнения системы (4)

где индекс 2 показывает порядок метода, а K – то, что функция относится к классическому методу. После соответствующих построений получим функции устойчивости структурного метода для обоих уравнений системы (4):

$$R_{2,1}(z) = 1 + z + \frac{z^2}{2} + \frac{z^3}{4}, \quad (6)$$

$$R_{2,2}(z) = 1 + z + \frac{z^2}{2}.$$

Сравнение областей устойчивости (рис. 1) показывает, что по обоим компонентам системы структурный метод при действительных собственных числах не хуже, а по первому уравнению в области комплексных собственных чисел где-то хуже, а где-то лучше. Весьма интересным оказывается то, что степень функции устойчивости $R_{2,1}(z)$ оказывается выше классической. Связано это с тем, что из-за особенностей структурного метода число «эффективных этапов» (число этапов в равносильной схеме, построенной по классическому образцу) для второго уравнения равно двум, а для первого – трём (реально же схема содержит два этапа по первому уравнению и один по второму).

Метод третьего порядка. Классическому трёхэтапному методу типа Рунге – Кутты третьего порядка точности соответствует функция устойчивости [6]

$$R_{3,K}(z) = 1 + z + \frac{z^2}{2} + \frac{z^3}{6}. \quad (7)$$

Функции устойчивости структурного метода третьего порядка, имеющего по два этапа для каждого уравнения системы (4), принимают вид

$$\begin{aligned} R_{3,1}(z) &= 1 + z + \frac{z^2}{2} + \frac{z^3}{6}, \\ R_{3,2}(z) &= 1 + z + \frac{z^2}{2} + \frac{z^3}{6} + \frac{z^4}{18}. \end{aligned} \quad (8)$$

Как видим, степени полученных областей устойчивости вновь соответствуют числу «эффективных этапов»: три по первому и четыре по второму. Это даёт основания считать, что в методах более высоких порядков степень функций устойчивости превзойдёт классическую ещё больше, что и подтверждается нижеследующими результатами.

Сравнивая области устойчивости (рис. 2), наблюдаем некоторое сжатие по действительной оси, но вновь расширение по мнимой.

Метод четвёртого порядка. Структурный метод четвёртого порядка требует совершить по три этапа по каждому из уравнений системы (4). Число эффективных этапов составляет пять по первому и шесть по второму уравнению. В этом случае мы вправе ожидать несовпадение обеих функций устойчивости с классической, соответствующей четырёхэтапному методу четвёртого порядка [6]

$$R_{4,K}(z) = 1 + z + \frac{z^2}{2} + \frac{z^3}{6} + \frac{z^4}{24}. \quad (9)$$

Действительно,

$$\begin{aligned} R_{4,1}(z) &= 1 + z + \frac{z^2}{2} + \frac{z^3}{6} + \frac{z^4}{24} + \frac{z^5}{144}, \\ R_{4,2}(z) &= 1 + z + \frac{z^2}{2} + \frac{z^3}{6} + \frac{z^4}{24} + \frac{z^5}{144} + \frac{z^6}{864}. \end{aligned} \quad (10)$$

Полученные области устойчивости в обоих случаях целиком покрывают классическую (рис. 3, 4). Данный результат показывает,

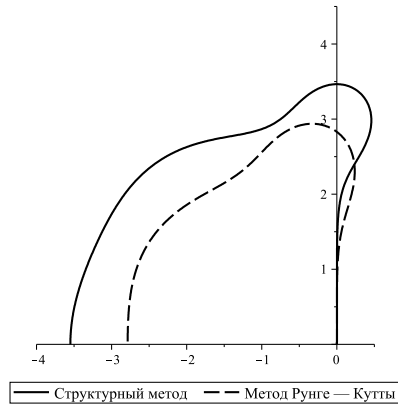


Рис. 3. Области устойчивости методов 4-го порядка для первого уравнения системы (4)

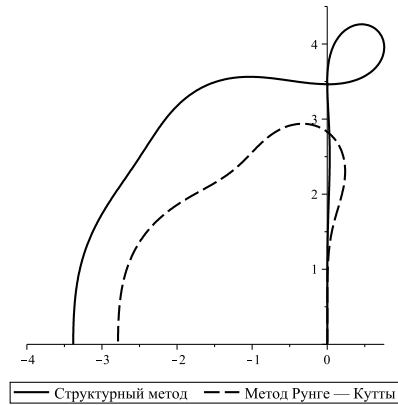


Рис. 4. Области устойчивости методов 4-го порядка для второго уравнения системы (4)

что рассматриваемый метод не только экономичнее, но и более устойчив.

Заключение. Опираясь на результаты, полученные в статье, можно ожидать, что устойчивость методов пятого порядка в сравнении с классическими окажется также не хуже. Особый интерес представляет четырёхэтапный метод из [8].

Система (4) содержит всего два уравнения. Однако при применении структурных методов к системам большей размерности число эффективных этапов, а следовательно, и степень функций устойчивости могут возрастать. Анализ классических областей устойчивости показывает, что с ростом степени функции $R(z)$ область устойчивости в среднем расширяется. Это даёт основания полагать, что применение структурных методов к системам большей размерности только увеличит их устойчивость по всем уравнениям.

Кроме того, интерес представляет рассмотрение устойчивости структурных методов на системах наиболее общего вида, содержащих так называемую *общую* группу уравнений, не обладающих структурными особенностями. Однако полученные в данной статье результаты дают основания полагать, что и для уравнений общей группы области устойчивости будут не меньше, а почти наверняка (в зависимости от вида системы) и больше классических.

Литература

1. Олемской И.В. Численный метод интегрирования систем обыкновенных дифференциальных уравнений // Математические методы анализа управляемых процессов / Под. ред. Боярского А.Я. Л.: Изд-во ЛГУ, 1986. С. 157–160.
2. Олемской И.В. Структурный подход в задаче конструирования явных одношаговых методов // Журнал вычислительной математики и математической физики, 2003. Т. 43, № 7. С. 961–974.
3. Олемской И.В. Методы интегрирования систем структурно разделённых дифференциальных уравнений. СПб: Изд-во СПбГУ, 2009. 180 с.
4. Еремин А.С. Вложенный метод типа Дормана – Принса для структурно разделённых систем // Процессы управления и устойчивость: Труды 40-й международной научной конференции аспирантов и студентов / Под ред. Н.В. Смирнова, Г.Ш. Тамасяна. СПб: Издат. Дом С.-Петербур. гос. ун-та, 2009. С. 160–165.
5. Еремин А.С., Олемской И.В. Вложенный метод интегрирования систем структурно разделённых обыкновенных дифференциальных уравнений // Журнал вычислительной математики и математической физики, 2010. Т. 50, № 3. С. 434–448.
6. Хайрер Э., Ваннер Г. Решение обыкновенных дифференциальных уравнений. Том 2: Жёсткие и дифференциально-алгебраические задачи / Пер. с англ. Е.Л. Старостина, И.А. Кульчицкой, А.В. Тыглияна и С.С. Филиппова под ред. С.С. Филиппова. М.: Мир, 1999. 685 с.
7. Dormand J.R., El-Mikkaway M.E.A., Prince P.J. High-order embedded Runge–Kutta–Nystrom formulae // IMA Journal of Numerical Analysis, 1987. Vol. 7. P. 423–430.
8. Олемской И.В. Четырёхэтапный метод пятого порядка точности численного интегрирования систем специального вида // Журнал вычислительной математики и математической физики, 2002. Т. 42, № 8. С. 1179–1190.

Шиманчук Д.В., Белова О.А.
Санкт-Петербургский государственный университет

Стабилизация орбитального движения космического аппарата в окрестности коллинеарной точки либрации

Введение. Известно, что некоторые математические модели, такие как задача двух тел и задача трёх тел, являются полезными для разработки проектов, связанных с исследованием космического пространства. Одной из наиболее известных моделей движения является модель ограниченной круговой задачи трёх тел [1].

Уравнения движения космического аппарата (КА), получаемые в рамках ограниченной задачи трёх тел Солнце – Земля – КА, как показали Эйлер и Лагранж [2], имеют частные решения, сохраняющие положение в пространстве конфигураций. Они и получили название либрационных.

С использованием теоретических свойств точек либрации связаны многие проекты, как освоения космического пространства, мониторинга околоземного космического пространства, так и защиты Земли от потенциально опасных космических объектов.

Уравнения движения и их характеристика. Пусть КА движется в окрестности коллинеарных точек либрации L_1 или L_2 . Тогда в рамках круговой ограниченной задачи трёх тел Солнце – Земля – КА, следуя [3], уравнения управляемого движения КА во вращательной системе координат имеют вид [4]:

$$\begin{cases} \dot{x}_1 = x_2 + y_1, \\ \dot{x}_2 = -x_1 + y_2, \\ \dot{x}_3 = y_3, \end{cases} \begin{cases} \dot{y}_1 = -\frac{3x_1}{\|\mathbf{x}\|^3} + 2x_1 + y_2 + u_1, \\ \dot{y}_2 = -\frac{3x_2}{\|\mathbf{x}\|^3} - x_2 - y_1 + u_2, \\ \dot{y}_3 = -\frac{3x_3}{\|\mathbf{x}\|^3} - x_3 + u_3, \end{cases} \quad (1)$$

где центр инерции Земли совпадает с началом системы координат, а ось Ox_1 направлена вдоль оси, соединяющей центры масс Земли и Солнца. Здесь (x_1, x_2, x_3) – вектор координат КА, (y_1, y_2, y_3) – вектор импульсов, $\|\cdot\|$ – евклидова норма. Точки либрации L_1

и L_2 во вращающейся системе неподвижны и имеют координаты $\mathbf{x}^* = (1, 0, 0)$, $\mathbf{y}^* = (0, 1, 0)$ и $\mathbf{x}^{**} = (-1, 0, 0)$, $\mathbf{y}^{**} = (0, -1, 0)$ соответственно. За единицу расстояния принята величина, равная 10^{-2} а.е., за единицу времени – 58,0916 суток.

Неуправляемая система (1) является гамильтоновой, где функция Гамильтона представляется в виде

$$H = \frac{1}{2} \|\mathbf{y}\|^2 - \frac{3}{\|\mathbf{x}\|} - \frac{3}{2} x_1^2 + \frac{1}{2} \|\mathbf{x}\|^2 + x_2 y_1 - x_1 y_2.$$

Система линеаризованных уравнений в окрестности коллинеарных точек либрации L_1 или L_2 имеет вид

$$\begin{cases} \dot{x}_1 = x_2 + y_1, \\ \dot{x}_2 = -x_1 + y_2, \\ \dot{x}_3 = y_3, \end{cases} \quad \begin{cases} \dot{y}_1 = 8(x_1 \mp 1) + (y_2 \mp 1), \\ \dot{y}_2 = -4x_2 - y_1, \\ \dot{y}_3 = -4x_3. \end{cases} \quad (2)$$

Замечание. Выбор знака в выражении для \dot{y}_1 соответственно зависит от рассматриваемой окрестности коллинеарной точки либрации L_1 или L_2 .

Матрица линеаризованной системы (2) имеет набор собственных значений [4]

$$\lambda_{1,2} = \pm \sqrt{1 + 2\sqrt{7}}, \quad \lambda_{3,4} = \pm i \sqrt{2\sqrt{7} - 1}, \quad \lambda_{5,6} = \pm 2i.$$

Замечание. Из положительности собственного значения $\lambda_1 = \sqrt{1 + 2\sqrt{7}}$ следует неустойчивость точки либрации L_1 и L_2 .

Так как последние два уравнения, соответствующие пространственным переменным (x_3, y_3) , не зависят от остальных, удобно рассмотреть задачу нахождения законов управления отдельно для системы плоского движения:

$$\begin{cases} \dot{x}_1 = x_2 + y_1, \\ \dot{x}_2 = -x_1 + y_2, \end{cases} \quad \begin{cases} \dot{y}_1 = 8(x_1 \mp 1) + (y_2 \mp 1) + u_1, \\ \dot{y}_2 = -4x_2 - y_1 + u_2, \end{cases} \quad (3)$$

и системы пространственных переменных:

$$\begin{cases} \dot{x}_3 = y_3, \\ \dot{y}_3 = -4x_3 + u_3. \end{cases} \quad (4)$$

Замечание. Для неуправляемой системы (4) ($u_3 = 0$) в силу интеграла энергии имеет место устойчивость по Ляпунову стационарного решения.

Из проведенного анализа системы уравнений орбитального движения КА можно сделать вывод, что неустойчивость коллинеарной точки либрации, с одной стороны, требует стабилизирующего управления, а с другой стороны, может быть использована как положительный фактор в задачах маневрирования [5, 6].

Функция опасности, область управляемости. Выше было показано, что стационарное решение уравнений неуправляемого движения, соответствующее коллинеарной точке либрации, является неустойчивым положением равновесия. Поэтому весьма актуальной будет задача стабилизации КА в окрестности точки либрации.

В связи с этим возникает вопрос: поскольку на практике управляющее воздействие ограничено количеством используемых ресурсов, то как оценивается через это количество область управляемости, т.е. область, в которой возможно удержание КА при ограничениях на управления.

Матрицу линеаризованной подсистемы плоского движения можно привести к диагональной ($\det \neq 0$), тогда для общего решения данной системы можем записать

$$\mathbf{z} = \sum_{i=1}^4 c_i \mathbf{b}_i \exp(\lambda_i t), \quad (5)$$

где $\mathbf{z} = (x_1 \mp 1, x_2, y_1, y_2 \mp 1)$, c_i – произвольные постоянные, $\mathbf{b}_i = (b_i^1, b_i^2, b_i^3, b_i^4)$ – собственный вектор соответствующий собственному значению λ_i линеаризованной неуправляемой системы (3), а сама система векторов \mathbf{b}_i , вообще говоря, ортонормирована.

Пусть $d_i = \mathbf{b}_i \mathbf{z}$, $i = 1, \dots, 4$, тогда в силу неуправляемой системы (3) справедливы равенства

$$\dot{d}_i = \lambda_i d_i. \quad (6)$$

Если в (6) положить $i = 1$, то

$$\dot{d}_1 = \lambda_1 d_1. \quad (7)$$

Данное дифференциальное уравнение имеет решение

$$d_1 = c_1 \exp(\lambda_1 t). \quad (8)$$

Легко видеть, что решение $|d_1| \rightarrow \infty$ при $t \rightarrow \infty$, когда $c_1 \neq 0$.

Отсюда следует, что определенная функция d_1 экспоненциально возрастает (когда $c_1 \neq 0$) и её рост определяется равенством (8).

Если же $i = 2, 3, 4$, то

$$\begin{cases} \dot{d}_2 = \lambda_2 d_2, & d_2 = c_2 \exp(\lambda_2 t), \\ \dot{d}_3 = \lambda_3 d_3, & d_3 = c_3 \exp(\lambda_3 t), \\ \dot{d}_4 = \lambda_4 d_4, & d_4 = c_4 \exp(\lambda_4 t). \end{cases}$$

Функция d_2 экспоненциально стремится к нулю, когда $c_2 \neq 0$. Из совместного рассмотрение функций d_3 и d_4 заключаем, что они описывают колебательный процесс.

Выражение $d_1 = \mathbf{b}_1 \mathbf{z}$, соответствующее компоненте решения (5), которая обеспечивает неустойчивость коллинеарной точки, будем называть *функцией опасности* [7].

Тогда, если нам удастся некоторым образом обеспечить выполнение ограничения функции d_1 в окрестности коллинеарной точки либрации, т.е. выполнение неравенства $|d_1| \leq d_0$, где d_0 – достаточно малая величина, то КА будет находиться в некоторой окрестности коллинеарной точки либрации достаточно длительное время (быстрого ухода по экспоненте, по крайней мере, в линейном приближении не будет).

В случае управляемого движения КА для производной функции d_1 можем записать (в силу линеаризованной системы уравнений движения (3))

$$\dot{d}_1 = \lambda_1 d_1 + u, \quad (9)$$

где $u = b_1^3 u_1 + b_1^4 u_2$.

Откуда для удержания (уменьшения модуля d_1) необходимо и достаточно, чтобы

$$\lambda_1 |d_1| < |u| \leq u_0.$$

Смысл данного неравенства заключается в том, что по значению величины d_1 можно судить о возможности стабилизации КА в окрестности коллинеарной точки либрации или, начиная с некоторого момента времени отклонения КА, стабилизация станет невозможной по причине ограниченности ($|u| \leq u_0$) возможностей ресурсов управления.

Проведенное исследование дает математическую базу для построения (с помощью функции опасности) не только алгоритмов ста-

билизации движения КА, но, например, и поиска траекторий возвращения КА в окрестность точки либрации [5, 6].

Управления, стабилизирующие орбитальное движение КА. Ранее было показано, если удастся с помощью управляющего воздействия демпфировать значение функции опасности, то КА будет находиться на траектории, принадлежащей окрестности коллинеарной точки либрации. В общем случае, если функция опасности равна нулю, то КА находится на гало-орбите, совершая некоторое условно периодическое движение.

Таким образом, управляя значением (единственной) введенной функции, можно будет решить задачу по стабилизации орбитального движения КА в окрестности точки либрации. Получаем задачу нахождения управления, при котором решение уравнения (9) $d_1 \rightarrow 0$, при $t \rightarrow \infty$, т.е. нахождения управления, демпфирующего функцию опасности (7).

Примером таких управлений могут быть управления с компонентами [6]:

$$\begin{cases} u_1 = k_1 d_1, \\ u_2 = k_2 d_1, \end{cases} \quad \text{при} \quad b_1^3 k_1 + b_1^4 k_2 < -\lambda_1, \quad (10)$$

$$\begin{cases} u_1 = k \left(\frac{b_1^1}{b_1^3} (x_1 \mp 1) + y_1 \right), \\ u_2 = k \left(\frac{b_1^2}{b_1^4} x_2 + (y_2 \mp 1) \right), \end{cases} \quad \text{при} \quad k < -\lambda_1, \quad (11)$$

где k , k_1 и k_2 – вещественные константы, удовлетворяющие неравенствам в (10) и (11).

Численные эксперименты. Приведем несколько примеров с предложенными законами управления. Проинтегрируем нелинейную систему (1) на промежутке времени, соответствующем 50 единицам (около 8 лет), и с начальными данными:

$$\begin{aligned} x_{1_0} &= 1, & y_{1_0} &= 0,01, \\ x_{2_0} &= 0, & y_{2_0} &= 1, \\ x_{3_0} &= 0, & y_{3_0} &= 0,05. \end{aligned}$$

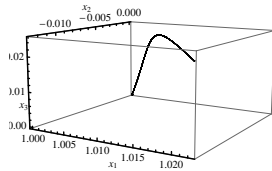


Рис. 1. Траектория КА без управления

Выбор таких начальных данных связан с предположением, что КА, находясь в фазовом пространстве с координатами точки L_1 , получил дополнительный импульс вдоль составляющих y_1 и y_3 вектора \mathbf{y} . На рис. 1 представлена траектория неуправляемого движения КА на промежутке времени $t \in [0, 1]$, за которое наблюдается уход на расстояние порядка 10^4 км.

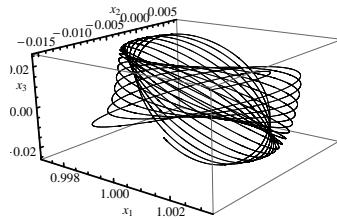


Рис. 2. Траектория движения КА с управлением (10) при $k_1 = -5, k_2 = -10$

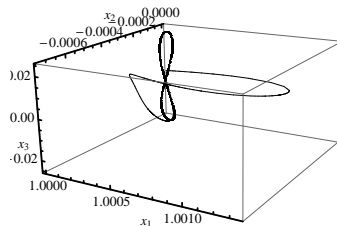


Рис. 3. Траектория движения КА с управлением (11) при $k = -5$

Заключение. Таким образом была показана возможность стабилизации КА в окрестности коллинеарной точки либрации с помощью управляющего воздействия, которое демпфирует функцию опасности.

Необходимо отметить, возможность осуществления такого рода управлений на практике, например, в случае использования таких технических систем управления как солнечный парус. Его применение становится возможным в силу того, что управляющие ускорения оказываются достаточно малыми по величине (порядка 10^{-6} м/с²).

Литература

1. Брюно А.Д. Ограниченная задача трех тел. М.: Наука, 1990. 295 с.
2. Маркеев А.П. Точки либрации в небесной механике и космодинамике. М.: Наука, 1978. 312 с.
3. Hill G.W. Researches in the lunar theory // Amer. Journal of Mathematics, 1878. Vol. 1. P. 5–26, 129–147, 245–260.
4. Шмыров В.А. Стабилизация управляемого орбитального движения космического аппарата в окрестности коллинеарной точки либрации L_1 // Вестник СПбГУ. Сер. 10, 2005. Вып. 2. С. 193–199.
5. Шиманчук Д.В. О возможности построения траектории возвращения КА в окрестность коллинеарных точек либрации L_1 и L_2 // Процессы управления и устойчивость: Труды 41-й международной научной конференции аспирантов и студентов / Под ред. Н.В. Смирнова, Г.Ш. Тамасяна. СПб.: Издат. Дом С.-Петербур. гос. ун-та, 2010. С. 212–216.
6. Шиманчук Д.В., Шмыров А.С. Построение траектории возвращения в окрестность коллинеарной точки либрации системы Солнце–Земля // Устойчивость и процессы управления: Тезисы докладов всероссийской конференции, посвященной 80-ти летию со дня рождения В.И. Зубова, 2010. С. 103–104.
7. Шмыров В.А. Управление орбитальным движением космического аппарата в окрестности коллинеарной точки либрации. Диссертация на соискание уч. степени канд. физ.-мат. наук, СПбГУ, 2005. 100 с.
8. Шиманчук Д.В. Моделирование орбитального управляемого движения космического аппарата в окрестности коллинеарной точки либрации L_1 // Вестник СПбГУ. Сер. 10, 2010. Вып. 3. С. 86–92.

Широколюбов А.Ю.

Санкт-Петербургский государственный университет

Математическая модель оптимизации параметров ускорителя на бегущей волне в случае кусочно-гладких управлений

Рекомендовано к публикации профессором Овсянниковым Д.А.

1. Введение. Проблемы оптимизации движения пучков заряженных частиц исследованы во многих работах (например, [1, 2]). Но, в основном, управления искались в классе кусочно-непрерывных функций, и алгоритмы не учитывали гладкость функций.

В данной статье рассматривается математическая модель совместной оптимизации программного и возмущенных движений [3] в ускорителе на бегущей волне в случае кусочно-гладких управлений и рассмотрен специальный тип вариации управления.

2. Постановка задачи. Рассмотрим продольное движение электронов в поле бегущей волны. Введем цилиндрическую систему координат, ось Oz которой совмещена с осью симметрии структуры. Уравнения движения частиц без учета их взаимодействия имеют вид [1]

$$\frac{d\gamma}{d\xi} = -\alpha(\xi) \sin \varphi,$$
$$\frac{d\varphi}{d\xi} = 2\pi \left(\frac{1}{\beta_{\Phi}(\xi)} - \frac{\gamma}{\sqrt{\gamma^2 - 1}} \right).$$

Здесь $\xi = \frac{z}{\lambda}$ – безразмерное расстояние; λ – длина волны основной гармоники ускоряющего поля; $\gamma = \frac{W_n}{W_0}$ – приведенная энергия частицы; $W_n = mc^2$ и $W_0 = m_0c^2$ – соответственно полная энергия и энергия покоя частицы; m и m_0 – соответственно масса и масса покоя частицы; c – скорость света в вакууме; φ – фаза частицы.

Функции $\alpha(\xi)$ и $\beta_{\Phi}(\xi)$ определяют законы изменения соответственно амплитуды напряженности ускоряющей волны и её фазовой скорости.

Под программным движением будем понимать решение системы

$$\begin{aligned}\frac{d\bar{\gamma}}{d\xi} &= -\alpha(\xi) \sin \bar{\varphi}, \\ \frac{d\bar{\varphi}}{d\xi} &= 2\pi \left(\frac{1}{\beta_{\Phi}(\xi)} - \frac{\bar{\gamma}}{\sqrt{\bar{\gamma}^2 - 1}} \right)\end{aligned}$$

при начальных условиях

$$\begin{aligned}\bar{\gamma}(0) &= \bar{\gamma}_0, \\ \bar{\varphi}(0) &= \bar{\varphi}_0.\end{aligned}$$

Движение частиц пучка опишем уравнениями в отклонениях от программного движения:

$$\begin{aligned}\frac{d\hat{\gamma}}{d\xi} &= \frac{d\gamma}{d\xi} - \frac{d\bar{\gamma}}{d\xi} = \alpha(\xi) (\sin \bar{\varphi} - \sin(\hat{\varphi} + \bar{\varphi})), \\ \frac{d\hat{\varphi}}{d\xi} &= \frac{d\varphi}{d\xi} - \frac{d\bar{\varphi}}{d\xi} = 2\pi \left(-\frac{\hat{\gamma} + \bar{\gamma}}{\sqrt{(\hat{\gamma} + \bar{\gamma})^2 - 1}} + \frac{\bar{\gamma}}{\sqrt{\bar{\gamma}^2 - 1}} \right).\end{aligned}$$

Здесь $\hat{\gamma} = \gamma - \bar{\gamma}$; $\hat{\varphi} = \varphi - \bar{\varphi}$; γ и φ – приведенная энергия и фаза частицы пучка.

Далее примем следующие обозначения:

$$\begin{aligned}x &= \bar{\gamma}(\xi), \\ u_1 &= \alpha(\xi), \quad u_2 = \bar{\varphi}(\xi) \\ y_1 &= \hat{\gamma}(\xi), \quad y_2 = \hat{\varphi}(\xi).\end{aligned}$$

Функции $u_1(\xi)$, $u_2(\xi)$ будем называть *управлениями*.

В итоге получим управляемую динамическую систему, описываемую системой обыкновенных дифференциальных уравнений:

$$\frac{dx}{d\xi} = -u_1(\xi) \sin(u_2(\xi)), \quad (1)$$

$$\begin{aligned}\frac{dy_1}{d\xi} &= u_1(\xi) (\sin(u_2(\xi)) - \sin(y_2 + u_2)), \\ \frac{dy_2}{d\xi} &= 2\pi \left(-\frac{x + y_1}{\sqrt{(x + y_1)^2 - 1}} + \frac{x}{\sqrt{x^2 - 1}} \right)\end{aligned} \quad (2)$$

с начальными условиями

$$\begin{aligned}x(0) &= x_0, \\y(0) &= y_0 \in M_0.\end{aligned}\tag{3}$$

Здесь $\xi \in T_0 = [0, L]$ – независимая переменная; $x \in \Omega_x \subseteq R^1$; $y \in \Omega_y \subseteq R^2$ – вектор фазовых переменных размерности 2; $u \in U \subseteq R^2$ – двумерная вектор-функция управления; L – фиксированная величина.

Предполагается, что множества Ω_x, Ω_y – открытые, множество U – компактное. Множество ненулевой меры $M_0 \subset \Omega_y$ будем полагать компактным.

Предположим также, что допустимые управления $u = u(\xi)$, $\xi \in T_0$, составляют некоторый класс кусочно-гладких на интервале $[0, L]$ функций, принимающих значения из компактного множества U .

Решения подсистемы (2) с начальными условиями (3) при фиксированном программном движении будем называть *возмущенным движением* [3].

Введем следующие функционалы, определенные на траекториях системы (1), (2):

$$I_1(u) = \int_0^L \varphi_1(\xi, x(\xi), u(\xi)) d\xi + g_1(x(L)),\tag{4}$$

$$I_2(u) = \int_0^L \int_{M_{\xi, u}} \varphi_2(\xi, x(\xi), y(\xi), u(\xi)) dy_\xi d\xi + \int_{M_{L, u}} g_2(y_L) dy_L,\tag{5}$$

где множество $M_{\xi, u}$ есть сечение в момент ξ пучка траекторий подсистемы (2) при управлении $u(\xi)$ и соответствующем программном движении $x(\xi)$, исходящих из множества M_0 ,

$$M_{\xi, u} = \{y_\xi \mid y_\xi = y(\xi, x(\xi), y_0, u(\xi)), y_0 \in M_0, x(0) = x_0\}.$$

Функционал (4) характеризует динамику программного движения, а функционал (5) оценивает поведение пучка траекторий в целом.

Рассмотрим связку функционалов (4) и (5):

$$I(u) = c_1 I_1(u_1) + c_2 I_2(u_2),\tag{6}$$

где c_1 и c_2 – неотрицательные константы.

Функционал (6) позволяет одновременно учитывать динамику программного движения и пучка траекторий.

3. Вариация функционала. Рассмотрим классическую вариацию функционала (6). Пусть варьируемое управление строится по правилу [3, 4]

$$u_\varepsilon(\xi) = u(\xi + \varepsilon\delta(\xi)), \quad \xi \in T_0. \quad (7)$$

Здесь $\varepsilon \in [0, 1]$ – параметр, характеризующий малость вариации; $\delta(\xi)$ – непрерывно дифференцируемая функция, удовлетворяющая условиям

$$0 \leq \xi + \varepsilon\delta(\xi) \leq L, \quad \xi \in T_0.$$

Очевидно, что если $u(\xi)$ – допустимое управление, то управление $u_\varepsilon(\xi)$ при $\varepsilon \in [0, 1]$ будет допустимым.

При вариации управления по формуле (7) множество значений управления останется неизменным. Следовательно, значения управления будут просто перемешиваться. Перемешивать между собой будем значения не на всем промежутке $[0, L]$ в целом, а по участкам. Обозначим точки разбиения k_i . При этом $0 = k_0 < k_1 < \dots < k_{N-1} < k_N = L$, $i = \overline{0, N}$, где N – число участков.

Очевидно, что $\operatorname{div}_x f(x(\xi), u_1(\xi), u_2(\xi)) = 0$, где $f(x(\xi), u_1(\xi), u_2(\xi))$ – правая часть подсистемы (1), и $\operatorname{div}_y F(x(\xi), y(\xi), u_1(\xi), u_2(\xi)) = 0$, где $F(x(\xi), y(\xi), u_1(\xi), u_2(\xi))$ – правая часть подсистемы (2). Следовательно, якобианы преобразования пучков траекторий по сечениям равны единице, т.е. $\left| \frac{\partial \tilde{x}_\xi}{\partial x_\xi} \right| = 1$ и $\left| \frac{\partial \tilde{y}_\xi}{\partial y_\xi} \right| = 1$.

Выбирая варьируемое управление по формуле (7), используя представление

$$\Delta u(\xi) = \frac{du(\xi)}{d\xi} \varepsilon \delta(\xi) + o(\varepsilon),$$

учитывая вышесказанное и следуя [3], представим классическую ва-

риацию функционала (6) следующим образом:

$$\begin{aligned} \delta I(u, \Delta u) = & \varepsilon \sum_{i=1}^N \int_{k_{i-1}}^{k_i} \left\{ c_1 \left(-\psi^* \frac{\partial f}{\partial u} + \frac{\partial \varphi_1}{\partial u} \right) - \right. \\ & \left. - c_2 \int_{M_{\xi, u}} \left(\mu^* \frac{\partial F}{\partial u} + \nu^* \frac{\partial f}{\partial u} - \frac{\partial \varphi_2}{\partial u} \right) dy_{\xi} \right\} \frac{du}{d\xi} \delta_i(\xi) d\xi, \end{aligned}$$

где «*» – знак транспонирования, вспомогательная функция $\psi(\xi)$ удовлетворяет вдоль траекторий подсистемы (1) следующему обыкновенному дифференциальному уравнению при условии на правом конце

$$\psi(L) = - \left(\frac{\partial g_1(x(L))}{\partial x} \right)^*,$$

а вспомогательные вектор-функция $\mu(\xi, y_{\xi})$ и функция $\nu(\xi, y_{\xi})$, в свою очередь, удовлетворяют вдоль траекторий системы (1), (2) следующим системам уравнений

$$\begin{aligned} \frac{d\mu}{d\xi} = & - \left(\frac{\partial F(\xi, x(\xi), y(\xi), u(\xi))}{\partial y} \right)^* \mu + \left(\frac{\partial \varphi_2(\xi, x(\xi), u(\xi))}{\partial y} \right)^*, \\ \frac{d\nu}{d\xi} = & - \left(\frac{\partial f(\xi, x, u)}{\partial x} \right)^* \nu - \left(\frac{\partial F(\xi, x, u)}{\partial x} \right)^* \mu + \left(\frac{\partial \varphi_2(\xi, x, u)}{\partial x} \right)^* \end{aligned}$$

с условиями на правом конце

$$\begin{aligned} \mu(L, y_L) = & - \left(\frac{\partial g_2(y(L))}{\partial y} \right)^*, \\ \nu(L, y_L) = & - \left(\frac{\partial g_2(y(L))}{\partial x} \right)^*. \end{aligned}$$

При этом функции $\psi(\xi)$, $\mu(\xi, y_{\xi})$, $\nu(\xi, y_{\xi})$ удовлетворяют следующим равенствам:

$$\begin{aligned} \psi(k_i - 0) &= \psi(k_i + 0), \\ \mu(k_i - 0) &= \mu(k_i + 0), \\ \nu(k_i - 0) &= \nu(k_i + 0), \quad i = \overline{1, N-1}. \end{aligned} \tag{8}$$

Формулы (8) определяют условия непрерывности функций $\psi(\xi)$, $\mu(\xi, y_\xi)$, $\nu(\xi, y_\xi)$ и являются аналогом условий Вейерштрасса – Эрдемана [5] в вариационном исчислении.

4. Заключение. На основе полученной вариации исследуемого функционала можно строить алгоритмы численного решения в задаче оптимизации динамики пучков заряженных частиц.

Литература

1. Овсянников Д.А. Математические методы управления пучками. Л.: Изд-во ЛГУ, 1980. 228 с.
2. Владимирова Л.В., Овсянников А.Д., Рубцова И.Д. Об управлении пучком электронов в ускорителе на бегущей волне // Вопросы механики и процессов управления, 2004. Вып. 22. С. 82–91.
3. Овсянников А.Д. Управление программным и возмущенными движениями // Вестник СПбГУ. Сер. 10, 2006. Вып. 4. С. 111–124.
4. Аргучинцев А.В. Оптимальное управление начально-краевыми условиями гиперболических систем. Иркутск: Изд-во ИГУ, 2004. 40 с.
5. Сейдж Э.П., Уайт Ч.С. Оптимальное управление системами. М.: Радио и связь, 1982. 392 с.

Ярыш Г.В., Пронина Ю.Г.
Санкт-Петербургский государственный университет

Механохимическая коррозия линейного участка трубопровода¹

1. Введение. *Коррозией* называют самопроизвольное разрушение металлических материалов, вызванное химическим или электрохимическим воздействием окружающей среды. Коррозионные процессы отличаются широким разнообразием условий и сред, в которых они протекают. По характеру коррозионного разрушения различают *местную коррозию* (охватывающую отдельные участки металла) и *сплошную* (когда поражается вся поверхность металла, находящаяся под воздействием среды). Сплошная коррозия бывает равномерной, неравномерной и избирательной. В некоторых системах металл – среда механические напряжения повышают интенсивность коррозионного процесса. В научной литературе такую коррозию называют *механохимической*.

В данной работе рассматривается равномерная поверхностная механохимическая коррозия упругой трубы. Для скорости равномерной коррозии v в настоящее время большинство ученых (см. [1-3]) принимают линейную зависимость от напряжений σ

$$v = (a + m \sigma) e^{-bt}.$$

Экспоненциальный множитель здесь означает возможность затухания коррозии с течением времени t при образовании плотной пленки окислов. В работе [4] выведено решение задачи о механохимической коррозии трубы, когда в качестве эквивалентного напряжения σ выбрано максимальное нормальное напряжение. В инженерных расчетах вместо σ часто удобнее использовать интенсивность напряжений. В представленной статье производится оценка прочности линейной части трубопровода с использованием именно интенсивности напряжений как основной прочностной характеристики.

2. Постановка задачи. Рассматривается плоская деформация длинного толстостенного полого цилиндра, находящегося под действием внутреннего p_r и внешнего p_R давления коррозионных сред.

¹Работа выполнена при финансовой поддержке СПбГУ (проекты № 9.0.165.2009, 9.37.129.2011) и РФФИ (проекты № 10-01-00093, 11-01-00230)

Его внутренний и внешний радиусы в начальный момент времени $t = 0$ обозначены соответственно через r_0 и R_0 . Под действием среды материал цилиндра корродирует равномерно по внешней и внутренней поверхностям, т.е. его поперечное сечение меняет свои размеры, оставаясь при этом концентрическим кольцом. Через промежуток времени t его внутренний радиус увеличивается от r_0 до $r = r_0 + \delta_r$, а внешний радиус уменьшается от R_0 до $R = R_0 - \delta_R$. Обозначим скорость коррозии материала цилиндра с внутренней стороны через v_r , а с наружной – через v_R . Как было отмечено выше, наилучшее совпадение с экспериментальными данными обеспечивается следующими выражениями для скоростей коррозии:

$$v_r = \frac{d(r_0 + \delta_r)}{dt} = [a_r + m_r \sigma_i(r)] e^{-bt}, \quad (1)$$

$$v_R = -\frac{d(R_0 - \delta_R)}{dt} = [a_R + m_R \sigma_i(R)] e^{-bt}, \quad (2)$$

где b, a_r, a_R, m_r, m_R – постоянные, определяемые опытным путем, σ_i – интенсивность напряжений.

Требуется определить интенсивность напряжений в любой момент времени и оценить долговечность изделия.

3. Вывод основного разрешающего уравнения. Согласно решению задачи Ламе о толстостенной трубе, интенсивность напряжений на ее внутренней и внешней поверхностях вычисляется по формулам

$$\sigma_i(r) = \sqrt{3} \frac{R^2 |p_r - p_R|}{(R^2 - r^2)} = \sqrt{3} \frac{\eta^2 |p_r - p_R|}{(\eta^2 - 1)}, \quad (3)$$

$$\sigma_i(R) = \sqrt{3} \frac{|p_r - p_R|}{(\eta^2 - 1)} = \sigma_i(r) - \sqrt{3} |p_r - p_R|, \quad (4)$$

где

$$\eta = \frac{R}{r}. \quad (5)$$

Поскольку интенсивность напряжений максимальна на внутренней поверхности трубы, целесообразно наблюдать именно за величиной $\sigma_i(r)$. Выберем ее в качестве основной переменной и в дальнейшем будем обозначать σ_i . С использованием (4) уравнение (2)

перепишем в виде

$$\frac{d(R_0 - \delta_R)}{dt} = -[A_R + m_R \sigma_i(R)] e^{-bt}, \quad (6)$$

где $A_R = a_R - \sqrt{3} m_R |p_r - p_R|$.

Из (3) находим обратную зависимость

$$\eta = \sqrt{\frac{\sigma_i}{\sigma_i - \sqrt{3} |p_r - p_R|}}. \quad (7)$$

Исключив из системы уравнений (1), (6) интенсивность напряжений и проинтегрировав полученное уравнение по t от 0 до t , с учетом обозначения (5), приходим к соотношению

$$\begin{aligned} r_0 + \delta_r = \\ = \frac{m_R \left[r_0 - \frac{a_r}{b} (e^{-bt} - 1) \right] + m_r \left[R_0 - \frac{A_R}{b} (e^{-bt} - 1) \right]}{\eta m_r + m_R}. \end{aligned} \quad (8)$$

Дифференцирование выражения (3) по t и последующие преобразования с использованием (5), (7) и (8) позволяют получить разрешающее дифференциальное уравнение

$$\begin{aligned} \frac{d\sigma_i}{dt} = & \frac{2\sqrt{\sigma_i (\sigma_i - \sqrt{3} |p_r - p_R|)} \left(m_r \sqrt{\sigma_i} + m_R \sqrt{\sigma_i - \sqrt{3} |p_r - p_R|} \right)}{\sqrt{3} |p_r - p_R| e^{bt}} \times \\ & \times \frac{[A_R + m_R \sigma_i] \sqrt{\sigma_i - \sqrt{3} |p_r - p_R|} + [a_r + m_r \sigma_i] \sqrt{\sigma_i}}{m_R \left[r_0 + \frac{a_r}{-b} (e^{-bt} - 1) \right] + m_r \left[R_0 + \frac{A_R}{-b} (e^{-bt} - 1) \right]}. \end{aligned} \quad (9)$$

4. Решение основного уравнения. Разделив переменные и проинтегрировав уравнение (9) по t от 0 до t и, соответственно, по σ_i от $\sigma_0 = \sigma_i|_{t=0}$ до σ_i , для $b \neq 0$ находим

$$t = -\frac{1}{b} \ln \left[1 - \frac{b(m_R r_0 + m_r R_0)}{m_R a_r - m_r A_R} (\exp \{ (m_R a_r - m_r A_R) F(\sigma_i) \} - 1) \right],$$

где

$$F(\sigma_i) = \int_{\sigma_0}^{\sigma_i} \frac{\sqrt{3} |p_r - p_R| / 2}{(A_R + m_R \sigma) \sqrt{\sigma - \sqrt{3} |p_r - p_R|} + (a_r + m_r \sigma) \sqrt{\sigma}} \times \frac{d\sigma}{\sqrt{\sigma} \left(m_r \sqrt{\sigma} + m_R \sqrt{\sigma - \sqrt{3} |p_r - p_R|} \right) (\sigma - \sqrt{3} |p_r - p_R|)}.$$

При $b = 0$ интеграл уравнения (9) имеет вид

$$t = \frac{m_R r_0 + m_r R_0}{m_R a_r - m_r A_R} (\exp \{ (m_R a_r - m_r A_R) F(\sigma_i) \} - 1),$$

где функция $F(\sigma_i)$ остается без изменений.

В случае односторонней коррозии решение существенно упрощается. Например, при коррозионном износе с внутренней стороны

$$t = -\frac{1}{b} \ln [1 - b F_r(\sigma_i)],$$

где

$$F_r(\sigma_i) = \frac{R_0}{a_r} \sqrt{\frac{\sigma_i - \sqrt{3} |p_r - p_R|}{\sigma_i}} - \frac{r_0}{a_r} + \frac{\sqrt{3} |p_r - p_R| R_0 m_r / 2}{\sqrt{a_r^3 (a_r + m_r \sqrt{3} |p_r - p_R|)}} \times \times \ln \left(\frac{a_r \sqrt{3} |p_r - p_R| - (2 a_r + m_r \sqrt{3} |p_r - p_R|) \sigma}{a_r + m_r \sigma} + \right. \\ \left. + \frac{2 \sqrt{a_r (a_r + m_r \sqrt{3} |p_r - p_R|)} \sqrt{\sigma (\sigma - \sqrt{3} |p_r - p_R|)}}{a_r + m_r \sigma} \right) \Bigg|_{\sigma_0}^{\sigma_i}.$$

С помощью полученных решений можно определить долговечность линейного участка трубопровода, например, при хрупком разрушении: теоретический момент разрушения определяется точкой пересечения кривой изменения предела прочности материала $\sigma_s(t)$ и построенных интегральных кривых основного разрешающего уравнения.

5. Некоторые результаты расчетов. По выведенным формулам проведен ряд расчетов. В приведенных ниже примерах не указаны численные значения и единицы измерения, поскольку характер построенных зависимостей сохраняется при изменении масштабов времени, шкалы напряжений или линейных размеров. Представленные модели коррозионного износа применимы как для скоростей коррозии измеряемых в [мм/сек], так и для скоростей с единицами измерения [мм/год] (см. [2, 3]).

Сначала исследовался вопрос о том, какая коррозия – внутренняя или внешняя – более опасна с точки зрения прочности. Для примера на рис. 1 представлены графики изменения интенсивности напряжений σ_i с течением времени, где кривая 1 соответствует внешнему износу, 2 – внутреннему, 3 – двухстороннему; линия 4 изображает убывающий со временем предел прочности $\sigma_s(t)$. Начальные ско-

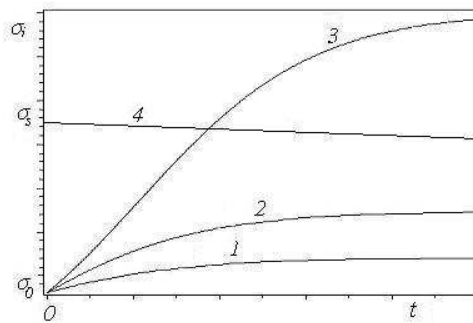


Рис. 1. Сопоставление двухсторонней, внутренней и внешней коррозии

рости коррозии, отличные от нуля, и прочие данные во всех трех случаях положены одинаковыми. Срок службы изделия в каждой ситуации определяется точкой пересечения соответствующей кривой σ_i с линией 4. Несложно заметить, что внутренняя коррозия опаснее, чем внешняя. Также видно, что при одновременном протекании коррозионных процессов (изнутри и снаружи) интенсивность напряжений существенно превосходит соответствующую величину, которая получилась бы в результате простой суперпозиции этих процессов по-отдельности.

На рис. 2 изображены зависимости σ_i от t для различных показателей затухания коррозии b , при прочих равных условиях. Линии 1–4 последовательно соответствуют значениям $b = \{0; 0, 2; 0, 5; 1\}$ кривая 5 – пределу прочности $\sigma_s(t)$. Из рис. 2 видно, что повышение показателя

теля затухания коррозии приводит к снижению роста напряжений и,

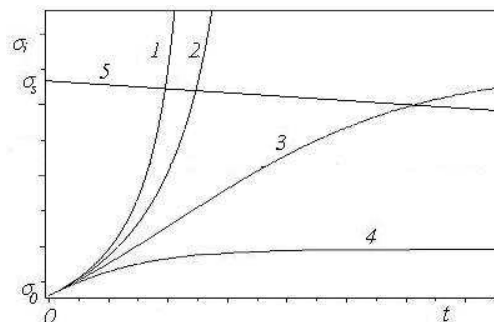
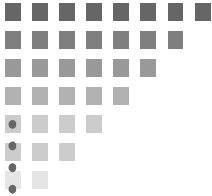


Рис. 2. Рост напряжений σ_i при различных показателях b

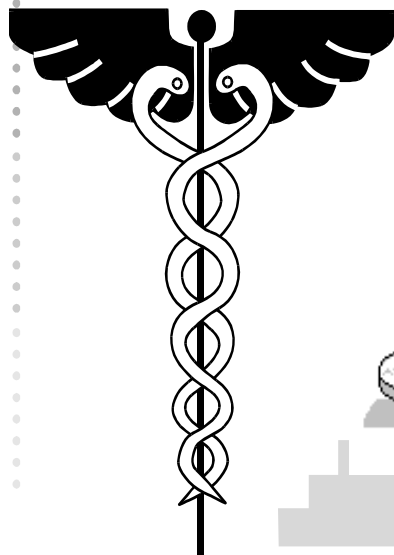
следовательно, к увеличению долговечности изделия. В некоторых случаях коррозия может практически прекратиться до достижения критического значения. В таких ситуациях долговечность трубопровода определяется скоростью снижения предела прочности материала или случайными обстоятельствами.

Литература

1. Долинский В.М. Расчет нагруженных труб, подверженных коррозии // Химическое и нефтяное машиностроение, 1967. № 2. С. 9–10.
2. Павлов П.А., Кадырбеков Б.А., Колесников В.А. Прочность сталей в коррозионных средах. Алма-ата: Наука, 1987. 272 с.
3. Наумова Г.А., Овчинников И.Г. Расчеты на прочность сложных стержневых систем и трубопроводных конструкций с учетом коррозионных повреждений. Саратов: Изд-во СГТУ, 2000. 227 с.
4. Пронина Ю.Г. Механохимическая коррозия полого цилиндра из идеального упругопластического материала под действием постоянного давления // Вестник СПбГУ. Сер. 1, 2006. Вып. 3. С. 121–130.



*3. Математические
модели медико-
биологических систем*



Абдрафиков В.А.

Санкт-Петербургский государственный университет

Анализ морфологии клеток крови

Рекомендовано к публикации профессором Андриановым С.Н.

Введение. Общий клинический анализ крови широко используется как один из самых важных методов обследования при большинстве заболеваний. Изменения, происходящие в крови, чаще всего неспецифичны, но в то же время отражают изменения, происходящие в целом организме. Клеточный состав крови здорового человека довольно постоянен. Поэтому различные его изменения, наступающие при заболеваниях, могут иметь важное диагностическое значение.

Эритроциты – это клетки крови человека, а также позвоночных и некоторых беспозвоночных животных, которые содержат в себе гемоглобин и выполняют транспортировку кислорода от легких к тканям и углекислого газа от тканей к легким. Эритроциты человека имеют форму двояковогнутого диска, красную окраску и размер 7–8 мкм в диаметре. При различных заболеваниях крови возможно изменение цвета эритроцитов, их размеров, количества, а также формы (см. рис. 1).

Многие специалисты проводят анализ крови посредством визуального наблюдения в микроскоп. Задача состоит в подсчете количества эритроцитов, а также оценке форм и размеров. Этот процесс требует от специалиста высокой концентрации и не гарантирует получение объективного результата. Именно поэтому представляется перспективным использовать для этой цели комплексы интеллектуального анализа данных.

1. Постановка задачи. Программно-аппаратный комплекс для проведения анализа крови должен состоять из тринокулярного микроскопа с объективами 10x, 40x, 100x; камеры с возможностью захвата изображения с разрешением не менее 1280×1024 и вычислительной машины для обработки видеоданных и получения результатов анализа.

Итак, исходными данными для текущей задачи является отснятое под микроскопом изображение мазка крови. Главная цель – опре-

делить последовательность алгоритмов теории обработки и распознавания изображений с целью подсчета количества эритроцитов и анализа их морфологии. На рис. 1 приведены различные виды патологий эритроцитов.

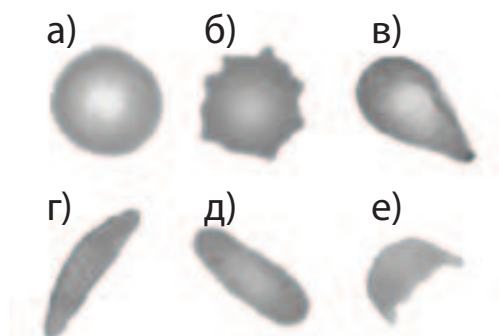


Рис. 1. Эритроциты: а) – нормоцит, б) – акантоцит, в) – дакриоцит, г) – дрепаноцит, д) – эллиптоцит, е) – шистоцит

Будем считать, что изображение представляет из себя матрицу размерности $w \times h$, т. е. $F = f_{x,y}$, где $x = \overline{1, w}, y = \overline{1, h}$, и каждый пиксель $f_{x,y}$ определяет значение яркости в точке (x, y) . Для определенности предполагаем, что F – полутоновое изображение (этой информации достаточно для оценки морфологии эритроцитов).

2. Выделение эритроцитов. Пусть F – изображение мазка крови, полученное с микроскопа. Сначала выполняем бинаризацию, т. е. разделение изображения на пиксели, принадлежащие объектам и фону.

$$F^*(x, y) = \begin{cases} 1, & \text{если } F(x, y) \leq \tau, \\ 0, & \text{если } F(x, y) > \tau, \end{cases}$$

где τ – порог бинаризации, который вычисляется автоматически, с помощью метода, описанного в [1].

Далее необходимо построить контура объектов на бинарном изображении F^* . Будем считать, что пиксели со значением 1 принадлежат объектам, 0 – фону.

Пусть дан объект на плоскости xy . Начиная с произвольной точки (x_0, y_0) , обойдем границу, например, против часовой стрелки, и обозначим координаты встречающихся точек границы $(x_0, y_0), \dots, (x_{K-1}, y_{K-1})$. Запишем эти координаты в форме $x(k) = x_k, y(k) = y_k$, тогда контур объекта можно представить следующим образом $s(k) = [x(k), y(k)]$. Процесс построения контуров подробно описан в [2].

3. Формирование вектора признаков. В данном разделе рассматривается процедура описания формы и размеров объектов, которую можно разбить на несколько этапов.

3.1. Калибровка. Определение реальных размеров эритроцитов на изображении требует знания калибровочного коэффициента. Для его нахождения необходимо отснять изображение калибровочной линейки на рабочем увеличении и определить количество пикселей в единице длины, например, количество пикселей в 1 мкм. Учитывая калибровочный коэффициент, можно оценить реальные размеры объекта на цифровом изображении.

3.2. Вычисление параметров формы и размера. Для каждого контура $s(k) = [x(k), y(k)], k = 0, K-1$, вычисляем следующие параметры:

- площадь $A = \sum_{k=0}^{K-1} (x(k)y(k+1) - x(k+1)y(k)) / 2$;
- периметр $P = \sum_{k=0}^{K-1} \sqrt{(x(k) - x(k+1))^2 + (y(k) - y(k+1))^2}$ (для более точного вычисления периметра необходимо производить предварительное сглаживание контура, например, с помощью метода бегущего среднего);
- фактор формы круга $S_f = 4\pi A/P^2$.

Эти параметры потребуются для распознавания нормацитов.

3.3. Фурье-дескрипторы. Границу объекта можно представить в виде последовательности комплексных чисел $z(k) = x(k) + jy(k)$. Дискретное преобразование Фурье конечной последовательности $z(k)$ задается уравнением

$$a(u) = \sum_{k=0}^{K-1} z(k)e^{-j2\pi uk/K}$$

при $u = 0, \dots, K-1$. Комплексные коэффициенты $a(u)$ называются Фурье-дескрипторами границы. Обратное преобразование Фурье,

примененное к этим коэффициентам, позволяет восстановить границу

$$z(k) = \frac{1}{K} \sum_{u=0}^{K-1} a(u) e^{j2\pi uk/K}.$$

Как показано в [3], высокочастотные составляющие описывают мелкие детали, тогда как низкочастотные компоненты определяют общую форму границы. Поэтому целесообразно использовать низкочастотные компоненты Фурье-преобразования для формирования векторов признаков.

Сами по себе коэффициенты преобразования Фурье не инвариантны относительно параллельного переноса и масштаба.

Нулевой коэффициент $a(0)$ отвечает за положение контура на плоскости xy , остальные коэффициенты не зависят от перемещения. Для получения инвариантных относительно масштаба дескрипторов формы можно поделить коэффициенты $a(u)$ на $|a(1)|$, т.е. $\hat{a}(u) = a(u)/|a(1)|$ [4].

Далее формируем вектор \mathbf{w} , содержащий абсолютные значения низкочастотных составляющих $\hat{a}(u)$. Таким образом, $\mathbf{w} = |\hat{a}(m)|$, где $m = [1, p] \cup [K - p, K - 1]$, $m \in \mathbb{Z}$, и p – индекс для отсекаания высокочастотных составляющих преобразования Фурье. Низкочастотные компоненты располагаются ближе к началу и концу вектора $\hat{a}(u)$. Для данной задачи значение индекса $p=15-20$.

4. Распознавание. Процедура распознавания заключается в сравнении вектора признаков объекта с векторами эталонных объектов, принадлежащих различным классам. Для формирования эталонных векторов признаков проводится процесс обучения.

4.1. Обучение. Процедура обучения производится в ручном режиме, т.е. для каждого класса объектов $c_i \in C$, C – множество классов, на которые можно разделить объекты данного типа. Пользователь указывает объекты, которые по его мнению принадлежат данному классу. Получаем обучающую выборку.

Таким образом, для каждого класса c_i формируется набор векторов признаков $\mathbf{w}_n^{(i)}$, $n = \overline{1, N}$, где N – количество объектов в классе, i – индекс класса. Фактически, для каждого класса получаем матрицу признаков $W_{M \times N}^{(i)}$, где M – количество компонентов вектора признаков \mathbf{w} (по построению $M = 2p$).

4.2. Выбор компонентов для классификации. Некоторые элементы вектора $\mathbf{w}_n^{(i)}$ могут быть нехарактерными для каждого класса c_i , а следовательно, отрицательно сказываться на качестве распознавания. Для решения данной проблемы действуем согласно следующему правилу. В матрице $W^{(i)}$ оставляем столбцы с теми индексами, которые позволяют отличить хотя бы один из классов от остальных.

Для всех классов c_i вычисляем среднее значение и среднеквадратичное отклонение по каждому столбцу $\mathbf{q}_m^{(i)}$, $m = \overline{1, M}$, матрицы $W^{(i)}$. Таким образом, среднее значение

$$d_m^{(i)} = \sum_{n=1}^N q_{m,n}^{(i)} / N$$

и среднеквадратичное отклонение

$$s_m^{(i)} = \sqrt{\sum_{n=1}^N (q_{m,n}^{(i)} - d_m^{(i)})^2 / N}$$

где N – количество объектов в классе, M – количество столбцов в матрице $W^{(i)}$.

Далее строим отрезки вида $x_m^{(i)} = [d_m^{(i)} - s_m^{(i)}, d_m^{(i)} + s_m^{(i)}]$. Если среди отрезков $x_m^{(i)}$, $i = \overline{1, J}$, где J – количество классов, есть хотя бы один отрезок, который не пересекается с остальными, то признак с индексом m оставляем для формирования итогового вектора. Сформированный таким способом вектор обозначим $\mathbf{v}^{(i)}$.

Затем вычисляем центральный вектор для каждого класса $\bar{\mathbf{v}}^{(i)}$ (среднее арифметическое по всем векторам в классе).

4.3. Классификация. После того, как произведен поиск объектов на изображении, для каждого контура вычисляются параметры площади (для площади нужно учитывать калибровочный коэффициент) и формы, по которым производится отбор нормальных эритроцитов (см. рис. 1 а)). Площадь нормоцита $\approx 35\text{--}55 \text{ мкм}^2$ и фактор формы круга не менее 0,7.

Затем для объекта формируется вектор признаков $\bar{\mathbf{v}}$ и находятся расстояния до эталонных векторов $\bar{\mathbf{v}}^{(i)}$, соответствующих классам $c_i \in C$. Данному объекту приписывается тот класс, расстояние до

центра которого наименьшее. Для вычисления расстояния используется Евклидова метрика.

Заключение. Представленный алгоритм позволяет классифицировать эритроциты и определять их концентрацию в исследуемом образце.

Алгоритм реализован в среде MATLAB и протестирован на наборе из 50 изображений, отснятых на разных препаратах (объекты на этих изображениях также были классифицированы вручную специалистом). Ошибка классификации составляет 7–12% в зависимости от построения обучающей выборки.

Дальнейшее улучшение алгоритма должно быть связано с классификацией других видов патологий эритроцитов, основным признаком которых является внутренняя текстура, а не форма.

Литература

1. Otsu N. A threshold selection method from gray-level histograms // IEEE Trans. Sys., Man., Cyber., 1979. № 9. P. 62–66.
2. Kimbahune V., Uke N. Blood cell image segmentation and counting. // International Journal of Engineering Science and Technology, 2011. Vol. 3, № 3. P. 2448–2453.
3. Gonzalez R.C., Woods R.E., Eddins S.L. Digital image processing using MATLAB, 2nd ed. Knoxville: Gatesmark Publishing. 2009, 618 p.
4. Яне Б. Цифровая обработка изображений. М.: Техносфера, 2007. 583 с.

Кузнецов С.Р.

Санкт-Петербургский государственный университет

**Пространственная математическая модель
развития дисбактериоза на слизистых
оболочках человека**

Рекомендовано к публикации профессором Шишкиным В.И.

Введение. В организме, главным образом в кишечнике, обитает множество бактерий, которые составляют нормальную микрофлору и выполняют важные для хозяина функции: препятствуют внедрению и размножению патогенной микрофлоры, синтезируют значительную часть витаминов и т.д. Имунная система хозяина не агрессивна к нормальной микрофлоре. Это свойство называется оральной толерантностью. В последние 50 лет отмечено резкое повышение заболеваемости населения эндогенными инфекциями, в частности, дисбактериозом, вызванными нарушением толерантности организма к нормальной микрофлоре в связи с повсеместным применением антибиотиков [1, 2]. Такие аутоинфекции обычно не приводят к гибели человека, однако существенно снижают его «качество жизни» и могут провоцировать ряд других тяжелых заболеваний.

С подобными инфекциями медицинская наука пока справляется плохо, да и сам механизм оральной толерантности до недавнего времени был мало понятен. Ситуация начала меняться в связи с открытием так называемых образраспознающих рецепторов (ОРР). Данные рецепторы имеются практически на всех клетках организма и способны различать патогенные организмы не только по отдельным их белкам, но и по их совокупности – по характерным для специфического патогена молекулярным образам (pathogen-associated molecular patterns) [3]. Микроорганизмы нормальной микрофлоры живут и размножаются в просвете и в слизи и имеют тесный контакт с клетками эпителия. Под действием различных стрессогенных факторов (аномальные температуры, голод, воздействие антибиотиков и т.п.) на поверхности бактерий образуются новые, меняющие их фенотип антигенные структуры, например, белки теплового шока. Появление большого количества таких структур может приводить к тому, что ОРР эпителиальных клеток перестают «узнавать» микроорганизмы собственной микрофлоры и начинают развивать им-

мунный ответ к ним, выпуская вначале в слизь различные антимикробные факторы (дефензины, лизоцим, белки комплимента и др.), а затем вовлекая в воспалительный процесс клетки естественного и адаптивного иммунитета [1, 2]. В нашей модели рассмотрим начальный этап этого процесса, проходящий еще без участия иммунных клеток. Этот этап очень важен, так как именно на этой стадии запускаются те воспалительные механизмы, которые потом переходят в хроническую эндогенную инфекцию.

Описание модели. Рассмотрим участок слизистой оболочки кишечника. Она состоит из эпителия, под которым находится слой соединительной ткани. Со стороны просвета кишки эпителий покрыт слоем слизи толщиной 0,1–0,3 мм. Именно в этом слое проходят процессы, которые и представляют интерес. Так как эпителий и слизь в разных участках кишечника имеют одинаковое строение, то можно описать эти процессы с помощью одномерных уравнений в частных производных, выбрав систему координат, в которой ось Ox совпадает с нормалью к слою эпителиальных клеток. Пусть толщина слизи равна l , точка $x = 0$ соответствует границе слоя слизи и полости, а точка $x = l$ соответствует границе слоя слизи и эпителия. Тогда интересующие нас процессы можно описать следующими уравнениями:

$$\begin{cases} \frac{\partial a}{\partial t} = D_1 \frac{\partial^2 a}{\partial x^2} - \lambda a, \\ \frac{\partial u}{\partial t} = D_1 \frac{\partial^2 u}{\partial x^2} - \lambda u, \\ \frac{\partial c_0}{\partial t} = D_2 \frac{\partial^2 c_0}{\partial x^2} + c_0 (\alpha - \beta(c + c_0)) - (k + \varphi)ac_0 - \\ \qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad - (m + \psi)uc_0 + \mu c, \\ \frac{\partial c}{\partial t} = D_2 \frac{\partial^2 c}{\partial x^2} + c (\alpha - \beta(c + c_0)) + (ka + mu)c_0 - \\ \qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad - \varphi ca - \psi cu - \mu c. \end{cases} \quad (1)$$

Здесь $a = a(x, t)$ описывает концентрацию внешнего раздражителя – антибиотика; $u = u(x, t)$ – концентрация противомикробных факторов, выделяемых эпителиальными клетками; $c_0 = c_0(x, t)$ – концентрация популяции бактерий нормальной микрофлоры, оказавшейся устойчивой к действию антибиотика; $c = c(x, t)$ – та часть

популяции c_0 , которая изменила свои фенотипические характеристики в результате стресса, вызванного действием антибиотика; $D_{1,2}$ – коэффициенты диффузии (были взяты одинаковые коэффициенты диффузии для антибиотика и противомикробных факторов, так как антибиотик изготавливается из сходных биологических веществ, имеющих противомикробную активность); λ – скорость выведения антибиотика и противомикробных факторов; второе слагаемое третьего и четвертого уравнения описывает размножение бактерий по закону логистического роста (можно также рассматривать эти слагаемые, как частный случай модели, описывающей конкуренцию двух видов [4]); k и m характеризуют интенсивность перехода бактерий в новое измененное состояние под действием антибиотика и противомикробных факторов; φ и ψ – бактерицидное действие антибиотика и противомикробных факторов; наконец, μ описывает возвращение «измененных» бактерий в нормальное (первоначальное) состояние.

Граничные условия. Будем считать, что живые бактерии не покидают слизи, а выведение антибиотика и противомикробных факторов полностью описывается членами λa и λu в (1). При этом источник антибиотиков находится в полости (точка $x = 0$), а источник противомикробных факторов – в эпителиальных клетках ($x = l$). Тогда граничные условия можно описать таким образом:

Левая граница	Правая граница
$a(x, t) \Big _{x=0} = \begin{cases} A, & t \in [0, t_a], \\ 0, & t > t_a, \end{cases}$	$\frac{\partial a}{\partial x} = 0,$
$\frac{\partial u}{\partial x} = 0,$	$u(x, t) \Big _{x=l} = \gamma c(l, t), \quad (2)$
$\frac{\partial c_0}{\partial x} = 0,$	$\frac{\partial c_0}{\partial x} = 0,$
$\frac{\partial c}{\partial x} = 0,$	$\frac{\partial c}{\partial x} = 0,$

где A – доза антибиотика, t_a – время его действия, а γ характеризует интенсивность выделения противомикробных факторов эпителиальными клетками в результате контакта с «измененными» бактериями.

Начальные условия. Будем полагать, что бактерии нормальной микрофлоры в начальный момент времени распределены по слизи равномерно, и их концентрация равна $C = \alpha/\beta$ (точка равновесия

для логистического уравнения в отсутствии конкуренции и раздражителей). Тогда начальные условия будут иметь следующий вид:

$$a(x, 0) \equiv 0, \quad u(x, 0) \equiv 0, \quad c_0(x, 0) \equiv C, \quad c(x, 0) \equiv 0. \quad (3)$$

Результаты. Описанная модель исследовалась численно с помощью метода конечных разностей с явной одношаговой схемой [5], реализованного на языке программирования Python. Коэффициенты были взяты из доступной научной литературы [6–13] или были рассчитаны (см. таблицу). Отметим, что D_1 вычислено на основе данных о диффузии молекул различных типов в слизи (измерения производились для NH_3 , HCl , производных полистирола); φ и A рассчитаны на основе данных о действии тобрамицина на *Pseudomonas aeruginosa*; γ рассчитан на основе данных об уровне лизоцима в цервикальной слизи в норме и при воспалении (от 6,4 до 391,4 мкг/мл).

Таблица. Значения коэффициентов

	Размерность	Допустимые значения	Рассчетные значения	Источник
D_1	см ² /с	$10^{-10} \dots 10^{-5}$	$2,7 \times 10^{-9}$	[6, 7, 13]
D_2	см ² /с	$10^{-7} \dots 10^{-3}$	$2,7 \times 10^{-7}$	[6–9]
λ	1/с	$5 \times 10^{-4} \dots 1,6 \times 10^{-3}$	$8,3 \times 10^{-4}$	[6, 13]
α	1/с	$0 \dots 4,2 \times 10^{-4}$	$2,7 \times 10^{-4}$	[10, 11]
C	1/см ³	$10^7 \dots 10^8$	10^8	[10]
β	см ³ /с	$10^{-12} \dots 10^{-11}$	$2,7 \times 10^{-12}$	рассчитан
k	1/(г·с)	нет данных	$1,39 \times 10^2$	рассчитан
m	1/(г·с)	нет данных	$1,39 \times 10^2$	рассчитан
φ	1/(г·с)	$0 \dots 1,97 \times 10^2$	$1,94 \times 10^2$	[11]
ψ	1/(г·с)	нет данных	83,4	рассчитан
μ	1/с	нет данных	$8,3 \times 10^{-5}$	рассчитан
A	г/см ³	$1,2 \times 10^{-7} \dots 1,04 \times 10^{-5}$	10^{-5}	[11]
t_a	с	$4,32 \times 10^4 \dots 8,64 \times 10^4$	$4,32 \times 10^4$	
γ	г	$0 \dots 3,9 \times 10^{-12}$	$2,6 \times 10^{-12}$	[12]
l	см	$9 \times 10^{-5} \dots 3 \times 10^{-4}$	2×10^{-4}	[6, 13]

Динамика модели позволяет увидеть, как происходит нарушение оральной толерантности. Во время действия антибиотика в кишечнике формируется субпопуляция «измененных» микроорганизмов нормальной микрофлоры. Эти бактерии благодаря высокой численности интенсивно активируют эпителий, который начинает выделять большие количества антимикробных факторов. После того, как

действие антибиотика заканчивается, воспалительный процесс продолжает поддерживаться антимикробными факторами эпителия, в результате чего популяция «измененных» бактерий не возвращается в нормальное состояние, что и является первым этапом в нарушении оральной толерантности и развитии дисбактериоза.

Модель исследовалась на чувствительность к изменению параметров γ и D_1 . Интересно, что наиболее высокий уровень воспаления наблюдается в довольно узком промежутке средних значений коэффициента γ . При высоких значениях этого коэффициента воспалительный процесс становится незначительным. Эти данные могут говорить о том, что при ослабленном иммунитете человек становится особенно подверженным дисбактериозу. Также присутствует зависимость между уровнем воспаления и коэффициентом диффузии D_1 . При $D_1 < 6 \times 10^{-11}$ см²/с заболевание вообще отсутствует; максимальный уровень воспаления наблюдается при значениях D_1 от $1,4 \times 10^{-10}$ до 10^{-9} см²/с; а при еще большем увеличении значения коэффициента уровень воспаления начинает постепенно снижаться. Это говорит о важности показателя вязкости слизи при развитии дисбактериоза и может указывать на новые пути в направлении поиска средств борьбы с данным заболеванием.

Заключение. Данные, полученные с помощью модели, не противоречат ожидаемым результатам. Модель может быть использована для дальнейшего исследования заболеваний, связанных с конфликтом организма и его микрофлоры. Заметим, однако, что значения коэффициентов модели нуждаются в большем уточнении, для чего требуется проведение специальных экспериментов.

Автор выражает благодарность д. м. н. В.И. Шишкину и к. ф.-м. н. В.М. Лыкову за помощь в проведении исследования.

Литература

1. Лебедев К.А., Понякина И.Д. Роль белков теплового шока в формировании конфликта организма человека с его микрофлорой // Физиология человека, 2007. Т. 33, № 3. С. 100–107.
2. Лебедев К.А., Понякина И.Д. Конфликт организма человека с его микрофлорой // Природа, 2007. № 7. С. 20–28.

3. Marca H.M. et al. Sensing and signaling by the immune system: NVVI–Dutch Society for Immunology Course, Lunteren, April 2-3, 2009 // *Immunology Letters*, 2010. Vol. 128, Iss. 1. P. 1–3.
4. Мюррей Дж. Математическая биология. Т. I. М.-Ижевск: Институт компьютерных исследований, 2009. 776 с.
5. Вержбицкий В.М. Основы численных методов: Учебник для вузов. 2-е изд., перераб. М.: Высшая школа, 2005. 840 с.
6. *Mucos Immunology (Third Edition)*. Ed. by Mestecky J. et al. Amsterdam: Elsevier Academic Press, 2005. 2064 p.
7. Celli J.P. et al. Helicobacter pylori moves through mucus by reducing mucin viscoelasticity // *Proceedings of the National Academy of Sciences USA*, 2009. Vol. 106, № 34. P. 14321–14326.
8. Ben-Jacob E. et al. Modeling branching and chiral colonial patterning of lubricating bacteria // *Mathematical models for biological pattern formation* / Ed. by Maini Ph.K., Othmer H.G. New York: Springer, 2001. P. 211–253.
9. Shigematsu M. et al. Spirochaete-like swimming mode of Campylobacter jejuni in a viscous environment // *Journal of Medical Microbiology*, 1998. Vol. 47, № 6. P. 521–526.
10. Kirschner D., Freter R. *Mathematical Models of Colonization and Persistence in Bacterial Infections* // *Persistent bacterial infections* / Ed. by Nataro J.P. et al. Washington, D.C.: ASM Press, 2000. P. 79–100.
11. Corvaisier S. et al. Comparisons between Antimicrobial Pharmacodynamic Indices and Bacterial Killing as Described by Using the Zhi Model // *Antimicrobial Agents and Chemotherapy*, 1998. Vol. 42, № 7. P. 1731–1737.
12. Eggert-Kruse W. et al. Antimicrobial activity of human cervical mucus // *Human Reproduction*, 2000. Vol. 15, № 4. P. 778–784.
13. Laia S.K. et al. Mucus-penetrating nanoparticles for drug and gene delivery to mucosal tissues // *Advanced Drug Delivery Reviews*, 2009. Vol. 61, Iss. 2. P. 158–171.

Мамочев В.А.

Санкт-Петербургский государственный университет

Применение статистического метода в задаче классификации¹

Рекомендовано к публикации доцентом Иголкиным В.Н.

В работах [1–3] была описана одна модификация статистического подхода для решения задачи классификации базы данных. Ниже указанный подход применяется для изучения базы данных пациентов психиатрического отделения психиатрической больницы св. Николая Чудотворца. Приводятся результаты применения предлагаемого метода и анализируются полученные результаты, подтверждающие эффективность нового подхода и возможность его применения в совокупности с другими известными методами.

Постановка задачи. Пусть имеются два множества

$$A = \{X_1^1, \dots, X_{n_1}^1\} \quad \text{и} \quad B = \{X_1^2, \dots, X_{n_2}^2\},$$

которые состоят из некоторого количества векторов размерности p . Требуется построить решающее правило (идентификатор) так, чтобы «оптимальным» образом разделить эти два множества. Далее найденный идентификатор используется для построения прогноза.

Напомним вкратце основные моменты алгоритма [1–3].

Алгоритм. Суть метода состоит в следующем: в пространстве признаков выбирается некоторое «хорошее» направление l , все векторы проектируются на это направление $y_i^j = (X_i^j, l)$, и решающая функция строится уже для этих одномерных статистических величин. Для этого сначала строятся выборочные распределения по полученным одномерным выборкам. Затем по одному из методов классификации находится решающая функция. Рассмотрим метод построения направления l . Введем величины

$$M_1 = \frac{1}{n_1} \sum_{j=1}^{n_1} X_j^1, \quad M_2 = \frac{1}{n_2} \sum_{j=1}^{n_2} X_j^2, \quad S = \sum_{i=1}^2 \sum_{j=1}^{n_i} (X_j^i - M_i)(X_j^i - M_i)^T.$$

¹Работа выполнена при финансовой поддержке РФФИ, проект № 09-01-00360

Величина σ_y^2 пропорциональна среднему квадрату отклонения y_j^i , $i = 1, 2$, $j = 1, \dots, n_i$. Направление l находится по формуле [2]

$$l = S^{-1}(M_1 - M_2). \quad (1)$$

В результате p -мерная задача классификации сводится к одномерной: имеем две обучающие последовательности y_j^k , $k = 1, 2$, $j = 1, \dots, n_k$, и вектор l , на который нужно проектировать наблюдения и классифицировать.

На последнем шаге алгоритма производится классификация наблюдений, ранее это осуществлялось на основе функций плотностей f_1 и f_2 , построенных таблично. В настоящей модификации алгоритма предлагается аппроксимировать выше описанные плотности распределения полиномами, построенными методом наименьших квадратов. Правило разделения данных двух классов будет являться отношением правдоподобий для этих двух полиномов.

Практическое применение. Данным методом были исследованы несколько баз данных. Ниже более подробно описаны результаты анализа базы психиатрического отделения. Изучается эффективность лечения двух групп пациентов (пациенты-мужчины (ПМ) и пациенты-женщины (ПЖ)). Классическим статистико-вероятностным методом (см. [4]) были получены три набора наиболее существенных параметров. Результаты применения вышеописанного подхода приведены в таблицах 1–4 для групп ПМ и ПЖ, соответственно.

База мужчин содержит данные о 523 пациентах, из них 297 входят в группу A (назовем их здоровыми), а 226 – в группу B (больные). База женщин содержит данные о 393 пациентках, из них 192 входят в группу C (назовем их здоровыми), а 201 – в группу D (больные). Каждый пациент характеризуется точкой в 58-мерном пространстве – это набор его параметров. Они описывают пациента либо качественно, либо количественно.

База с данными пациентов-мужчин обрабатывалась тремя методами (наборами параметров):

Таблица. Методы и набор параметров

Метод	Набор параметров
метод М1	8, 28, 37
метод М2	28, 29, 37
метод М3	28, 32, 37

Методом М1 получены результаты: $l = (-0,00245; 0,00016; -0,0026)$, $F(x) = -2234734,5x^3 - 0,4 \cdot 10^{-5}x^2 - 11814,3x + 109,62$, где l – вектор направления (1), $F(x) = f_1(x) - f_2(x)$.

Методом М2 получены результаты: $l = (-0,00248; -0,00025; -0,00267)$, $F(x) = 7869312,7x^3 - 0,7 \cdot 10^{-5}x^2 + 6976,46x + 80,5$.

Методом М3 получены результаты: $l = (-0,00003; -0,00249; -0,00264)$, $F(x) = 55071914,6x^3 - 0,00024 \cdot 10^{-5}x^2 - 9142,38x + 2,07$.

Таблица 1. Результаты для ПМ

Метод	Всего точек (a/b)	Идентифицировано правильно	Процент
М1	523 (297/226)	360 (173/187)	70,49
М2	523 (297/226)	360 (173/187)	70,49
М3	523 (297/226)	332 (149/183)	65,57

Таблица 2. Результаты для ПЖ

Метод	Всего точек (c/d)	Идентифицировано правильно	Процент
М1	393 (192/201)	245 (93/152)	54,07
М2	393 (192/201)	260 (132/128)	66,22
М3	393 (192/201)	275 (137/138)	70

Из приведенных результатов следует, что для ПЖ более эффективен М3, а для ПМ – методы М1 и М2.

Таблица 3. Результаты для ПМ методами М1 и М2

М1, М2	Всего	a	a %	b	b %	%	Множества
+	212	173	58,2	39	17,24	77,15	A
-	311	124	41,8	187	82,76	66,44	B
итог	523	297	58,2	226	82,76	70,48	

В табл. 3 представлены результаты подстановки данных пациентов-мужчин в $F(x) = f_1(x) - f_2(x)$, где $x = (X, l)$, $X \in A$. Из 523 точек 212 дали положительный знак, а 311 – отрицательный. Из 212 точек, идентифицированных как точки множества A, в действительности принадлежат этому множеству 173 точки (что составляет 58,2% от количества точек в множестве A). Соответственно, 39 точек принадлежат множеству B (что составляет 17,24% от количества точек в множестве B). Таким образом, если пациент оказался в группе «+», то с вероятностью, равной $58,2/(58,2 + 17,24)100\% = 77,15\%$, он принадлежит множеству A (см. табл. 3 столбец 6). Аналогично, если пациент оказался в группе «-», то с вероятностью 66,44% он принадлежит B. Аналогичные табл. 3 результаты анализа метода М3 для пациентов-мужчин приводятся в табл. 4.

Таблица 4. Результаты для ПМ метода МЗ

МЗ	Всего	a	a %	b	b %	%	Множества
+	192	149	50,16	43	19,03	72,49	A
–	331	148	49,84	183	80,96	61,9	B
итог	523	297	50,16	226	80,97	65,57	

Заключение. В работе представлен анализ базы данных (мужчин и женщин) шизофренического отделения. Полученные результаты оказались сравнимыми с результатами [5]. В дальнейшем планируется обобщение алгоритма на двумерный случай.

Литература

1. Иголкин В.Н., Ковригин Ф.Б., Старшинов А.И., Хохлов В.А. Статистическая классификация, основанная на выборочных распределениях. Л.: Изд-во ЛГУ, 1978. 102 с.
2. Демьянов В.Ф., Иголкин В.Н., Приставко В.Т. и др. Разделение баз данных результатов медицинских исследований лечения. Учебно-методическое пособие. СПб.: Изд-во СПбГУ, 2007. 88 с.
3. Мамочев В.А. Один статистический подход в задаче классификации // Процессы управления и устойчивость: Труды 41-й международной научной конференции аспирантов и студентов / Под ред. Н.В. Смирнова, Г.Ш. Тамасяна. СПб.: Издат. Дом С.-Петерб. гос. ун-та, 2010. С. 281–284.
4. Kokorina A.V. Ranking the parametrs in classification databases // Longevity, Aging and Degradation Models. Vol. 2. СПб: Изд-во СПбГПУ, 2004. С. 191–193.
5. Ананьев К.И. Оптимизационные методы в задачах идентификации и ранжирования // Процессы управления и устойчивость: Труды 41-й международной научной конференции аспирантов и студентов / Под ред. Н. В. Смирнова, Г. Ш. Тамасяна. СПб.: Издат. Дом С.-Петерб. гос. ун-та, 2010. С. 257–260.

Морозов П.Д.
Санкт-Петербургский государственный университет

Моделирование колебаний базилярной мембраны на основании теории Бекешы

Рекомендовано к публикации профессором Михеевым С.Е.

Анатомия ушной улитки и принцип ее работы. Ушная улитка представляет собой костный спиральный канал, имеющий два с половиной оборота и длину около 32 мм (рис. 1). Вид ушной человеческой улитки в поперечнике показан на рис. 2.

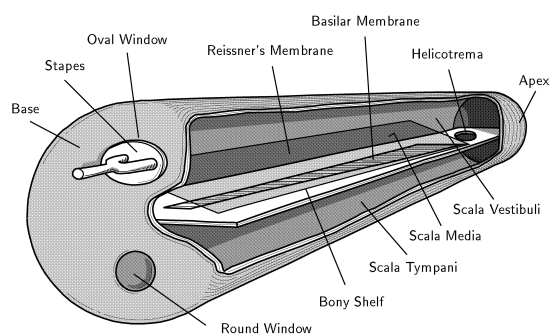


Рис. 1. Модель развернутой улитки

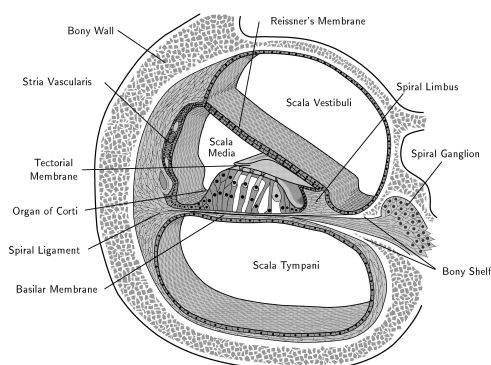


Рис. 2. Поперечный разрез настоящей человеческой улитки уха

Внутри улитка имеет сложную структуру. По всей длине она разделена двумя мембранами на три полости. Верхняя и нижняя полости называются *scala vestibuli* (в русской литературе именуется как лестница преддверия) и *scala tympani* (барабанная лестница) соответственно, и заполнены жидкой перилимфой. В начале *scala vestibuli* находится овальное окно, затянутое мембраной, через которую стремечко передает колебания во внутреннее ухо. В начале *scala tympani* находится круглое окно, выходящее обратно в среднее ухо. Разделяющая базилярную и преддверия лестницы полость носит название *scala media* (срединная полость). Она заполнена жидкой эндолимфой [1]. Дно *scala media* образует базилярная пластинка (базилярная мембрана), отделяющая его от *scala tympani*. Лестницы изолированы друг от друга на всем протяжении, кроме вершины улитки, где они соединены через отверстие – геликотрему. Звуковые волны в среднем ухе переходят в колебания стремечка, которые через овальное окно приводят в движение (внутри *scala vestibuli*) перилимфу, а та, в свою очередь, через рейснерову мембрану – эндолимфу в *scala media*. Вибрации обеих лимф не отделимы от колебаний волокон базилярной мембраны [2]. Базилярная мембрана состоит из нескольких тысяч слабо связанных между собой поперечных волокон, имеющих длину 32 мм, ширину у стремечка 0,05 мм (этот конец узкий, легкий и жесткий), а у геликотремы 0,5 мм (этот конец толще и мягче). На внутренней стороне базилярной мембраны находится орган Корти, а в нем специализированные слуховые рецепторы – порядка 23500 волосковых клеток. Колебания волокон возбуждают волосковые клетки. Они трансформируют механические колебания в электрические нервные импульсы. В поперечном направлении орган Корти состоит из одного ряда внутренних волосковых клеток и трех рядов наружных волосковых клеток. Между ними образуется тоннель. Волокна слухового нерва пересекают тоннель и контактируют с волосковыми клетками. Слуховой нерв представляет собой перекрученный ствол, сердцевина которого состоит из волокон, отходящих от верхушки улитки, а наружные слои – от нижних ее участков. Войдя в ствол мозга, нейроны взаимодействуют с клетками различных уровней, поднимаясь к коре и перекрещиваясь по пути так, что слуховая информация от левого уха поступает в основном в правое полушарие, где происходит главным образом обработка эмоциональной информации, а от правого уха – в левое полушарие, где в основном обрабатывается смысловая информация. В коре основные зоны слуха находятся в височной области.

Теории Гельмгольца и Бекеша. Существует несколько теорий для объяснения происходящих во внутреннем ухе процессов рецепции звуков. Основными являются теории Гельмгольца [3] и Бекеша [4]. Гельмгольц предложил резонансную теорию слуха. Согласно ей, волокна базилярной мембраны представляют собой набор из большого числа резонаторов, каждый из которых отзывается на колебания определенной частоты и возбуждает соответствующие нервные окончания слухового нерва. Таким образом, происходит вибрация тех участков базилярной мембраны, собственные частоты которых соответствуют частотам поступающих звуковых сигналов. По аналогии со струнными инструментами звуки высокой частоты приводят в колебательное движение участок базилярной мембраны с короткими волнами у основания улитки, а звуки низкой частоты вызывают колебания участка мембраны с длинными волокнами у вершины улитки. Теория Гельмгольца впервые позволила объяснить основные свойства уха – способность определения высоты, громкости и тембра и получила подтверждение в клинике. Кроме того, стало понятно почему ухо не различает звуков, полученных смешением одного и того же набора гармонических колебаний, но взятых в разных фазах. Дальнейшие исследования показали, что волокна основной мембраны связаны между собой, и, кроме того, погружение их в жидкость сильно гасит их колебания, так что отдельные резонансные колебания практически невозможны. Вместе с тем современные знания не подтверждают возможность резонирования отдельных «струн» базилярной мембраны. Прямое изучение механических свойств базилярной мембраны показало, что ей не свойственна высокая механическая избирательность. Звуковые волны различных частот вызывают колебания мембраны на участках большой протяженности.

В настоящее время все большее признание получает теория «бегущей волны» Бекеша. Согласно ей, звуки определенной высоты вызывают на базилярной мембране «бегущую волну» (рис. 3).

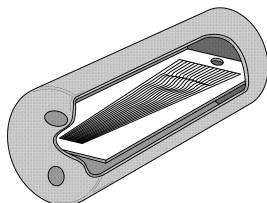


Рис. 3. «Бегущая волна»

Бекеша предложил модель возбуждения основной мембраны как целой натянутой гибкой перепонки, окруженной каналами с жидкостью, а также показал, что на такой мембране в зависимости от частоты колебаний, сообщаемых лестнице преддверия в ее начале, образуется в определенном месте область с колебаниями большой амплитуды.

Опытами Бекеша было доказано, что от частоты воспринимаемого звука зависит, какая группа волокон будет возбуждена в кортиевом органе. Низкие звуки вызывают прогибание мембраны у верхушки улитки, звуки высокой частоты – в области основного завитка улитки. Гребень «бегущей волны» смещает базилярную мембрану, что вызывает наибольшую деформацию волосковых клеток спирального органа. Таким образом, число различных типов раздражений, производимых звуками в слуховом органе, очень велико. Благодаря этому, ухо человека способно различать весьма тонкие особенности в звуках по их силе, частоте и спектральному составу.

В теории Бекеша и в последовавших за ней официально признанных моделях слуховой улитки есть общая черта: в них основную роль играют колебательные моды и физико-химические эффекты, обусловленные упруго-пластическими свойствами мембранного механизма аппарата слухового восприятия. Жидкости в слуховой улитке учитываются лишь как «передатчики давления» от овального окна к поверхности мембраны. Явно свойства перилимфы и эндолимфы в уравнения не входят, а влияние жидкостей учитывается косвенно с помощью функций Грина, задающих линейный закон передачи на мембрану звукового давления в овальном окне и давления в соседних точках мембраны:

$$p = \rho \frac{\partial \phi}{\partial t},$$

где p – звуковое давление, ρ – плотность базилярной мембраны, t – время, $\phi(x, y, t)$ – потенциал скоростей, определяющийся из соотношений

$$v_x = -\frac{\partial \phi}{\partial x}, \quad v_y = -\frac{\partial \phi}{\partial y}.$$

В связи с этим представляют интерес такие модели слухового восприятия, в которых бы гидродинамические эффекты в перилимфе учитывались более обоснованно наряду с учетом упругих свойств

мембраны. В частности, пассивные однокамерные модели, описанные в [5, 6], качественно согласовывались с данными Бекеши и сформировали теоретическую основу для механизма движения звуковой волны.

При моделировании однокамерной модели улитки основным уравнением является уравнение Лапласа, соответствующее в физическом смысле несжимаемости жидкости внутри канала:

$$\nabla^2 \phi = 0.$$

Граничным условием для него будет уравнение, накладываемое на колебания базилярной мембраны:

$$2\rho \frac{\partial \phi}{\partial t} = S(x)\delta + \beta(x) \frac{\partial \delta}{\partial t} + M(x) \frac{\partial \delta^2}{\partial t^2} \quad (1)$$

при $y = l$, где $\frac{\partial \delta}{\partial t} = v_y = -\frac{\partial \phi}{\partial y}$, а $S(x)$, $\beta(x)$, $M(x)$ – жесткость, затухание и масса базилярной мембраны, соответственно.

Согласно экспериментальным данным, масса базилярной мембраны крайне мала, вследствие чего можно отбросить содержащее ее слагаемое. Аналогично поступим с затуханием. Это существенно упростит уравнение (1). Воспользуемся следующими связями для получения уравнений для волнового числа:

$$\frac{\partial^2 \phi}{\partial t^2} = -\omega^2 \phi, \quad (2)$$

$$\frac{\partial \phi}{\partial y} = kB \sinh(ky) e^{\omega t - ky} = k \tanh(ky) \phi, \quad (3)$$

$$\frac{\partial^2 \phi}{\partial y \partial t} = i\omega k \tanh(ky) \phi, \quad (4)$$

$$\frac{\partial^3 \phi}{\partial y \partial t^2} = -\omega^2 k \tanh(ky) \phi. \quad (5)$$

Из (1)–(5) имеем

$$k(x) \tanh k(x)h = \frac{2\rho\omega^2}{S(x)}. \quad (6)$$

По мере прохождения по каналу улитки звуковые волны замедляются и становятся короче. Поэтому появляется зависимость волнового

числа от переменной x . Когда $t \ll 1$, алгебраический тангенс близок к t . Поэтому из (6) получаем

$$k(x) \approx \omega \sqrt{\frac{2\rho}{hS(x)}}.$$

Окончательно потенциал скоростей рассчитывается по формуле

$$\phi(x, y, t) = B \cosh(ky) e^{i(\omega t - kx)}.$$

Результат. Для рассмотренной модели получена схема для отыскания приближенного решения уравнения Лапласа и граничных условий.

Литература

1. Пальчун В.Т., Магомедов М.М., Лучихин Л.А. Оториноларингология. М.: Медицина, 2002. 576 с.
2. Камачкин А.М., Михеев С.Е., Евстафьева В.В. Модели колебаний в нелинейных системах. СПб.: Изд-во СПбГУ, 2004. 194 с.
3. Helmholtz H.L.F. On the sensations of tone. Dover, New York, 1954 (original German publication 1885).
4. Bekesy G. Experiments in hearing. Mc.Grow-Hill, 1960. 745 p.
5. Zwislocki J.J. Theory of the acoustical action of the cochlea // Journal of the Acoustical Society of America, 1950. Vol. 22. P. 778–784.
6. Ranke O.F. Theory of operation of the cochlea: A contribution to the hydrodynamics of the cochlea // Journal of the Acoustical Society of America, 1950. Vol. 22. P. 772–777.

Морозов П.Д., Михеев В.С.
Санкт-Петербургский государственный университет

Акустическая модуляция речи

Рекомендовано к публикации профессором Михеевым С.Е.

Задача. В радиотехнике под амплитудной модуляцией понимается такое кодирование звукового сигнала в радиоволны, при котором амплитуда входного сигнала переходит в амплитуду гармонического электромагнитного колебания несущего сигнала (при частотной модуляции, соответственно, амплитуда переходит в частоту). Наша задача заключалась в проведении экспериментов, связанных с распознаванием сигнала после применения амплитудной (частотной) модуляции не на электромагнитных, а на акустических гармонических колебаниях высокой частоты.

Описание эксперимента. Имеющаяся в наличии аппаратура (PC Genuine Intel(R) CPU 2160 @ 1,80 ГГц, Sound Blaster, динамики УТТ SP-80) позволяла генерировать различимый звук лишь до частот порядка 14 кГц. Поэтому, несмотря на возможность программного инструментария, как личного, так и фирменных звуковых редакторов, создавать файлы, которые могли бы на соответствующей аппаратуре порождать звуки до 96 кГц, ни интересный диапазон слышимых более высоких звуков 14 ÷ 20 кГц, ни неслышимый ультразвуковой 20 ÷ 48 кГц остались неисследованными.

Для исследования фразы записывались с помощью микрофона и звукового редактора Sound Forge 7.0 в WAV-файл. Формат WAV был выбран в силу простоты структуры и как содержащий информацию о звуковом давлении в виде простой оцифровки усредненного давления на фиксированном временном интервале, называемом *сэмплом*. Частота дискретизации, т.е. количество измерений звукового давления в секунду (иными словами, количество сэмплов в секунде), было выбрано равным 22050 Гц.

Программой генерировался гармонический несущий сигнал частотой от 10 до 14 кГц с частотой дискретизации 88200 Гц. Таким образом, при сохранении частотных характеристик входного сигнала, на каждый его сэмпл требовалось создать 4 сэмпла выходного сигнала. Для этого необходимо было произвести ту или иную интерполяцию. Были рассмотрены два варианта: линейная интерполяция и интерполяция кубическими квазиэрмитовыми сплайнами [1].

В первой серии экспериментов результат интерполяции накладывался на амплитуду несущего сигнала, который записывался в выходной WAV-файл, прослушиваемый и просматриваемый с помощью программы Sound Forge. Во второй серии экспериментов результат интерполяции накладывался на частоту несущего сигнала, тем самым образуя фазопериодическую функцию [2].

Для работы с оцифрованным звуком был выбран язык программирования C++ за такие достоинства: широкая распространенность, создание эффективных загрузочных модулей, возможность применения операций на низком уровне.

Интерполяции. *Линейная интерполяция* выполнялась так

```
Spl[0] = WorkR[i];
//WorkR[i] соответствует значениям f[i] из исходного файла
float DeltaWork = (WorkR[i+1] - WorkR[i])/4.0;
//интерполяция проходит по трем точкам
for (int k = 1; k < 4; k++) Spl[k] = Spl[k-1] + DeltaWork;
```

Такого типа интерполяция позволяла обрабатывать неограниченную последовательность входных данных (сэмплов). Образующиеся разрывы производных на каждом четвертом узле сглаживались аппаратно.

Для интерполяции гладкими сплайнами потребовался своего рода синтез эрмитовых сплайнов $S_{3,1}$ и кубических $S_{3,2}$. Первые применимы к бесконечным входным данным, но требуют знания производных в узлах, вторые не требуют производных, но разработаны только для конечного набора входных данных [3].

Была применена следующая методика: на участке между нулевым входным и первым узлами интерполяция не производилась (это не принципиально, можно, например, просто добавить фиктивный узел с индексом -1). Для узлов i и $i + 1$ вычислялись разностные производные (шаг берется постоянным, $h_i = h = \text{const}$):

$$m_i = \frac{f_{i+1} - f_{i-1}}{2h},$$

$$m_{i+1} = \frac{f_{i+2} - f_i}{2h}.$$

После чего строился «эрмитов» сплайн между узлами i и $i + 1$.

```

int j = -1; for (int i = 1; i < SIZE; i++)
{float t = 0.25;
float m_i = (WorkR[i+1] - WorkR[i-1])/2;
float m_i1 = (WorkR[i+2] - WorkR[i])/2;
for (int k = 1; k < 4; k++)
{Spl[k] = WorkR[i]*pow(1-t,2)*(1+2*t) + WorkR[i+1]*pow(t,2)
*(3-2*t) + m_i*t*pow(1-t,2) - m_i1*pow(t,2)*(1-t);
t = t + 0.25;
}
}

```

После интерполяции следовала модуляция

```

for (int k = 0; k < 4; k++)
{kk1 = (A0 + eps + Spl[k])*sin(++j*2*M_PI/per);
SinR_spl[j] = UnShort(kk1);
//рабочая часть выходного WAV-файла
}

```

Здесь $A0$ – амплитудный максимум входного сигнала, M_PI — число π , per – число узлов на период, выбором eps добиваемся положительности выходной амплитуды.

Частотная модуляция (рассматривалась только для линейного интерполирования) иницировалась начальной фазой $Phase=0$ и стандартным фазовым смещением $DeltaPhase=2M_PI/per$, и соответствующий блок модуляции выглядел как

```

kk1 = A0 * sin(Phase);
SinR_spl[++j] = UnShort(kk1);
Phase += DeltaPhase*(1-Spl[k]/maxSpl);

```

где в качестве $maxSpl$ выбиралось постоянное число, существенно большее абсолютных значений $Spl[k]$.

Помимо экспериментов над обычной модуляцией, были поставлены эксперименты над отрицательной амплитудной модуляцией, в которой повышение звукового давления входного сигнала преобразовывалось в уменьшение амплитуды выходного. При подобном инвертировании входного сигнала слуховое восприятие результата практически не имело отличий от входа, что вполне естественным образом объяснялось теорией восприятия звука по спектральному составу, но не по форме звуковой волны. При амплитудной модуляции такое объяснение не представляется логичным. Отсюда и возник интерес к отрицательной модуляции.

Результаты. Периоды выходного сигнала выбирались 6, 7, 8, 9, что соответствовало частотам от 14,5 кГц до 9,8 кГц. На всех частотах частотная модуляция не дала различимости речи. При амплитудной модуляции фразы были вполне различимы, причем при интерполяции квазиэрмитовыми кубическими сплайнами звук был ощутимо чище, чем при линейной, поскольку первые гораздо эффективнее сглаживали выходной сигнал. Стоит отметить, что слышимость пропадает при частоте 15 кГц, соответствующей $\text{per}=5$.

Любопытным оказался факт различимости речи при отрицательной модуляции, априори неочевидный.

Согласно теории «бегущей волны» Бекеша функционирования внутреннего уха, начало спиралевидного канала улитки, являющееся рецептором высоких частот, одновременно является и барьером распространения вглубь по каналу высоких частот [4]. При амплитудной модуляции фактически единственной составляющей звукового сигнала является несущая высокая частота, поэтому априори можно было бы предполагать отсутствие слышимости на существенно более низких частотах. В нашем эксперименте несущая частота примерно в 10 раз превосходила высшие частоты модулируемого сигнала. Тем не менее, после модулирования сохранялась различимость сигнала. Это наводит на мысль о целесообразности внесения в теории Бекеша некоторых добавок.

Интересными представляются эксперименты с амплитудной акустической модуляцией на аппаратуре, позволяющей воспроизводить звуковые частоты в диапазоне $14 \div 20$ кГц, а также в ультразвуке. В этом случае несущая частота при воспроизведении не будет слышима, но будет ли восприниматься ухом модуляция?

Литература

1. Михеев С.Е. Нелинейные методы в оптимизации. СПб.: Изд-во СПбГУ, 2001. 276 с.
2. Камачкин А.М., Михеев С.Е., Евстафьева В.В. Модели колебаний в нелинейных системах. СПб.: Изд-во СПбГУ, 2004. 194 с.
3. Завьялов Ю.С., Квасов Б.И., Мирошниченко В.Л. Методы сплайн-функций. М.: Наука, 1980. 352 с.
4. Bekesy G. Experiments in hearing. Mc.Grow-Hill, 1960. 745 p.

Чернэуцану Е.К.

Санкт-Петербургский государственный университет

Строгая h -отделимость двух множеств и линейное программирование¹

Рекомендовано к публикации профессором Поляковой Л.Н.

Пусть в \mathbb{R}^n заданы два конечных множества

$$A = \{a_i\}_{i=1}^m \quad \text{и} \quad B = \{b_j\}_{j=1}^k.$$

Назовём множества A и B строго h -отделимыми, если существует h пар $\{w^s, \gamma_s\}_{s=1}^h$, где $w^s \in \mathbb{R}^n$ и $\gamma_s \in \mathbb{R}$, таких, что при каждом $s \in \overline{1, h}$ выполняются неравенства

$$\langle w^s, a_i \rangle < \gamma_s \quad \text{при всех } i \in \overline{1, m};$$

для каждого $j \in \overline{1, k}$ найдётся индекс $s \in \overline{1, h}$, на котором

$$\langle w^s, b_j \rangle > \gamma_s.$$

Обозначим $G = \{w^s, \gamma_s\}_{s=1}^h$. В работах [1, 2] рассматривалась задача о строгой отделимости множеств A и B с помощью h гиперплоскостей в случае, когда выпуклая оболочка множества A и множество B не пересекаются, т.е. $\text{conv}(A) \cap B = \emptyset$. Было установлено, что эта задача сводится к экстремальной задаче

$$F(G) = \frac{1}{m} \sum_{i=1}^m \varphi_i(G) + \frac{1}{k} \sum_{j=1}^k \psi_j(G) \rightarrow \inf, \quad (1)$$

где

$$\begin{aligned} \varphi_i(G) &= \max_{s \in \overline{1, h}} [\langle w^s, a_i \rangle - \gamma_s + 1]_+, \\ \psi_j(G) &= \min_{s \in \overline{1, h}} [-\langle w^s, b_j \rangle + \gamma_s + 1]_+; \end{aligned}$$

здесь $[u]_+ = \max\{0, u\}$.

¹Работа выполнена при финансовой поддержке РФФИ, грант № 09-01-00360

По определению видно, что $F(G) \geq 0$. Неизвестной является $(h \times (n+1))$ -матрица G со строками (w^s, γ_s) , $s \in \overline{1, h}$.

Обозначим $\Pi = \{S = (s_1, \dots, s_k) \mid s_j \in \overline{1, h}, j \in \overline{1, k}\}$.

Лемма. *Справедливо равенство*

$$\inf_G F(G) = \min_{S \in \Pi} \inf_G \left\{ \frac{1}{m} \sum_{i=1}^m \varphi_i(G) + \frac{1}{k} \sum_{j=1}^k [-\langle w^{s_j}, b_j \rangle + \gamma_{s_j} + 1]_+ \right\}.$$

Лемма показывает, что задача (1) эквивалентна конечному числу экстремальных задач вида

$$\begin{aligned} & \frac{1}{m} \sum_{i=1}^m \max_{s \in \overline{1, h}} [\langle w^s, a_i \rangle - \gamma_s + 1]_+ + \\ & + \frac{1}{k} \sum_{j=1}^k [-\langle w^{s_j}, b_j \rangle + \gamma_{s_j} + 1]_+ \rightarrow \inf_G, \end{aligned} \quad (2)$$

соответствующих различным $S \in \Pi$. В свою очередь, задача (2) эквивалентна задаче линейного программирования

$$\begin{aligned} & \frac{1}{m} \sum_{i=1}^m p_i + \frac{1}{k} \sum_{j=1}^k q_j \rightarrow \inf, \\ & -\langle a_i, w^s \rangle + \gamma_s + p_i \geq 1, \quad i \in \overline{1, m}, s \in \overline{1, h}; \\ & \langle b_j, w^{s_j} \rangle - \gamma_{s_j} + q_j \geq 1, \quad j \in \overline{1, k}; \\ & p_i \geq 0, \quad i \in \overline{1, m}; \quad q_j \geq 0, \quad j \in \overline{1, k}. \end{aligned} \quad (3)$$

Теорема. *Задача (1) эквивалентна конечному числу задач линейного программирования вида (3) в том смысле, что решение задачи (3) при $S \in \Pi$, которому соответствует наименьшее значение целевой функции, является решением задачи (1).*

Рассмотрим пример. Пусть на плоскости заданы множества A и B , состоящие соответственно из точек

$$\begin{aligned} a_1 = (-2, 0), \quad a_2 = (2, 0), \quad a_3 = (0, 2), \quad a_4 = (0, 1); \\ b_1 = (0, 3), \quad b_2 = (3, 0), \quad b_3 = (-3, 0). \end{aligned}$$

Очевидно, что $\text{conv}(A) \cap B = \emptyset$ (см. рис. 1).

Решением данной задачи является вектор

$$z = \begin{pmatrix} 111, 2210 \\ 112, 0012 \\ 272, 9097 \\ -78, 1474 \\ 27, 3511 \\ 192, 1601 \\ 0, 0000 \\ 0, 0000 \\ 0, 0000 \\ 0, 0000 \\ 0, 0000 \\ 0, 0000 \\ 0, 0000 \end{pmatrix}.$$

Минимальное значение целевой функции равно нулю. Строгое двух-отделение множеств A и B при $S = (1, 1, 2)$ выглядит как показано на рис. 2.

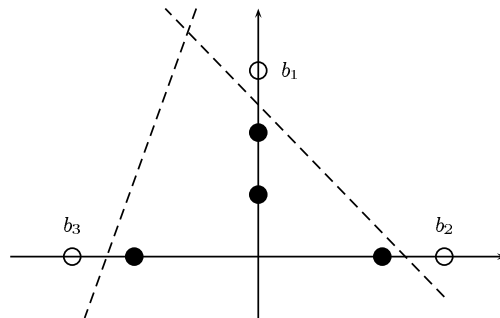


Рис. 2. Строгое двух-отделение при $S = (1, 1, 2)$

Отметим, что задание вектора индексов $S = (1, 1, 2)$ соответствует разбиению множества B на два подмножества $\{b_1, b_2\} \cup \{b_3\}$. Эти два подмножества согласовано отделяются от множества A с помощью двух прямых

$$\langle w^1, x \rangle = \gamma_1 \quad \text{и} \quad \langle w^2, x \rangle = \gamma_2.$$

Существуют ещё два разбиения множества B на два подмножества:

$$\{b_1, b_3\} \cup \{b_2\} \quad \text{и} \quad \{b_2, b_3\} \cup \{b_1\}.$$

Им соответствуют векторы $S = (1, 2, 1)$ и $S = (2, 1, 1)$.

Результат строгого двух-отделения при $S = (1, 2, 1)$ показан на рис. 3. Этот случай симметричен случаю $S = (1, 1, 2)$. При $S =$

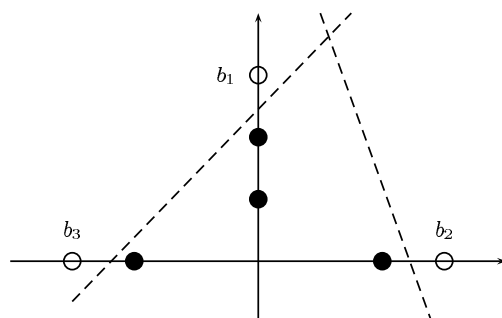


Рис. 3. Строгое двух-отделение при $S = (1, 2, 1)$

$(2, 1, 1)$ строгой двух-отделимости нет (см. рис. 4).

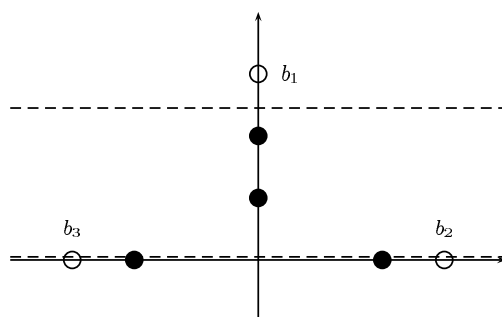


Рис. 4. При $S = (2, 1, 1)$ строгой двух-отделимости нет

Решением задачи, аналогичной (4), является вектор

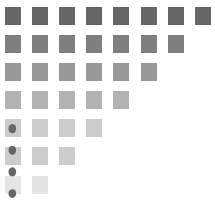
$$z = \begin{pmatrix} 0,0000 \\ -112,0230 \\ -1,0000 \\ 0,0000 \\ 111,8673 \\ 273,7824 \\ 2,0000 \\ 2,0000 \\ 0,0000 \\ 0,0000 \\ 0,0000 \\ 0,0000 \\ 0,0000 \\ 0,0000 \end{pmatrix}.$$

Минимальное значение целевой функции равно единице.

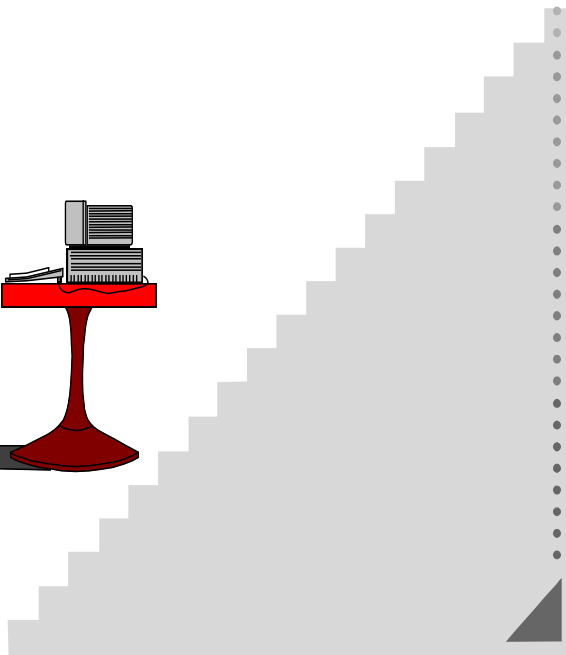
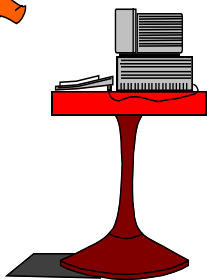
В общем случае задание вектора $S \in \Pi$ соответствует разбиению множества B , состоящего из k векторов, на h подмножеств. Число таких разбиений и определяет количество задач линейного программирования вида (3), к решению которых сводится решение задачи (1).

Литература

1. Astorino A., Gaudio M. Polyhedral separability throught successive LP // JOTA, 2002. Vol. 112, № 2 P. 265–293.
2. Малозёмов В.Н., Чернэуцану Е.К. Строгая h -отделимость двух множеств // Семинар «ДНА & CAGD», 2010. <http://dha.spb.ru/refs10.shtml.1218>



*4. Информационные
и компьютерные
технологии*



Антипин И.А., Козынченко В.А.
Санкт-Петербургский государственный университет

Об одной модификации нейронной сети Хемминга

Введение. В настоящее время искусственные нейронные сети представляют собой быстро развивающуюся область знаний с большой сферой практического применения. Искусственные нейронные сети широко применяются в задачах прогнозирования, распознавания образов, классификации и кластеризации, аппроксимации, сжатия и архивации данных и других практических задачах. Важное прикладное значение имеет задача распознавания образов. Для решения этой задачи используются такие искусственные нейронные системы как сеть Хопфилда, сеть Хемминга, сети адаптивной резонансной теории (АРТ), сети двунаправленной ассоциативной памяти, когнитрон и неокогнитрон и другие. В настоящее время сети адаптивного резонанса считаются наиболее эффективными нейронными структурами для решения задач распознавания образов [1]. Память сетей АРТ является пластичной, способной к запоминанию новых образов, и в то же время сохраняет относительную стабильность ранее запомненных образов. Сети адаптивного резонанса способны сравнивать входящий сигнал с хранящимися в памяти сети образами и принимать решение о наличии или отсутствии существенного различия между ними. В то же время, несмотря на некоторое биоподобие и свойство пластичности-стабильности памяти, сети адаптивного резонанса обладают некоторыми недостатками. Так, в сетях АРТ, выбор из памяти наиболее подходящего входному сигналу образа осуществляется по критерию максимума скалярного произведения входного вектора и вектора весовых коэффициентов запоминающих нейронов. Такой выбор, с учетом алгоритма работы модулятора сброса сетей АРТ, приводит к необходимости последовательного перебора многих запоминающих нейронов, т.е. к фазе поиска при функционировании сетей адаптивного резонанса. Необходимость поиска при подаче на вход сетей АРТ зашумленной версии запомненного сетью образа, а также необходимость перебора всех связанных (ранее участвовавших в запоминании) нейронов при подаче на вход сети нового образа может существенно замедлять работу нейронной сети. В данной работе предлагается нейронная сеть, являющаяся модификацией нейронной сети Хемминга, свободная от

указанного недостатка.

Как известно (см. [2]), предложенная Р. Липшманом сеть Хемминга, является рекуррентной структурой ассоциативной памяти. Сеть Хемминга работает с биполярными векторами. Нейроны первого слоя сети Хемминга хранят запомненные образы в векторах своих весовых коэффициентов и вычисляют расстояние Хемминга между вектором своих весовых коэффициентов и поступившим на вход сети вектором. Второй слой сети Хемминга, содержащий обратные связи, функционирует в режиме WTA и определяет нейрон с наименьшим расстоянием Хемминга между входным вектором и запомненными векторами. Нейроны третьего слоя формируют выходной вектор, соответствующий ближайшему из запомненных образов к предъявленному на вход вектору. Таким образом, сеть Хемминга не обладает свойством пластичности памяти и не способна к запоминанию новых образов после окончания обучения. Также сеть Хемминга не способна определить, что поданный на вход образ существенно отличается от всех хранящихся в памяти сети образов.

В данной работе предлагается модификация нейронной сети Хемминга, обладающая свойством пластичности-стабильности памяти. В отличие от сети Хемминга предлагаемая модифицированная сеть работает с бинарными векторами и состоит из пяти слоев. Так же как и сеть Хемминга, она является гетероассоциативной рекуррентной структурой, но в отличие от сети Хемминга обладает свойством стабильности-пластичности памяти. В данной работе будем называть рассматриваемую нейронную сеть модифицированной сетью Хемминга.

Архитектура модифицированной сети Хемминга. На рис. 1 представлена обобщенная схема модифицированной сети Хемминга. Сеть состоит из пяти нейронных слоев, модулятора сигнала и модулятора сброса. Первый слой является слоем формирования сигнала. Он состоит из n нейронов с пороговой функцией активации

$$f(y) = 0,5(\operatorname{sgn}(x) + |\operatorname{sgn}(x)|)$$

Количество нейронов n совпадает с размерностью вектора, поданного на вход сети. Смещения всех нейронов первого слоя полагаем равными нулю. Помимо входа для соответствующей компоненты входного вектора, каждый нейрон первого слоя имеет также вход от модулятора сброса.

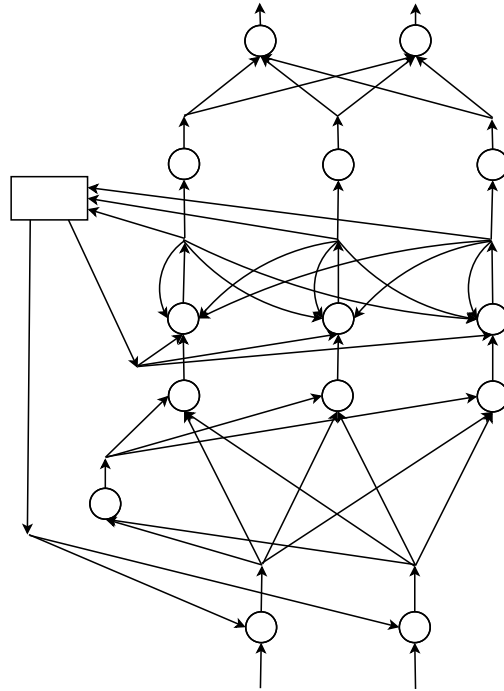


Рис. 1. Структура модифицированной сети Хемминга

Второй слой также имеет однонаправленное распространение сигналов. Количество нейронов второго слоя совпадает с количеством запоминаемых сетью векторов и определяет емкость памяти нейронной сети. Каждый нейрон второго слоя связан входными связями со всеми нейронами первого слоя, а также с модулятором сигнала. Нейроны второго слоя вычисляют функцию от расстояния Хемминга между входным вектором и вектором весов связей нейрона второго слоя с нейронами первого слоя. Смещения нейронов второго слоя принимаются равными сумме весов его связей с нейронами первого слоя. Функция активации нейронов второго слоя имеет вид

$$f(s) = 1 - \frac{s}{n}. \quad (1)$$

Третий слой состоит из нейронов с нулевым смещением и полулинейной функцией активации

$$f(y) = 0,5(y + |y|).$$

Третий слой является слоем конкуренции. Каждый нейрон третьего слоя связан обратными связями с остальными нейронами третьего слоя. Третий слой нейронов и по структуре, и функционированию совпадает со слоем MAXNET сети Хемминга. Количество нейронов в третьем слое то же, что и во втором слое. Нейроны третьего слоя в отличие от нейронов слоя MAXNET сети Хемминга имеют тормозящие входы от модулятора сброса, а также выходные связи с модулятором сброса.

Выход каждого нейрона третьего слоя имеет связь с единичным весом с одним нейроном четвертого слоя. Четвертый слой является вспомогательным и состоит из нейронов с нулевым смещением и пороговой функцией активации. Нейроны данного слоя предназначены для того, чтобы обеспечить единичный выход нейрону-победителю из третьего слоя. Выход каждого нейрона четвертого слоя связан со всеми нейронами пятого (выходного) слоя.

Пятый (выходной слой) состоит из того же количества нейронов, что и первый слой. Нейроны пятого слоя имеют нулевое смещение и пороговую функцию активации и предназначены для формирования выходного вектора, соответствующего ближайшему к входному вектору запомненному образу.

Помимо пяти нейронных слоев в рассматриваемой нейронной сети присутствуют два модулятора. Модулятор сигнала представляет собой один нейрон с линейной функцией активации и нулевым смещением. Он имеет входные соединения с единичными весами со всеми нейронами первого слоя и связан возбуждающими связями с единичными весами с нейронами третьего слоя. Модулятор сигнала обеспечивает вычисление меры Хемминга нейронами второго слоя.

Модулятор сброса представляет собой управляющую схему, предназначенную для сравнения расстояния Хемминга между входным сигналом и ближайшим (по мере Хемминга) запомненным сетью образом с заданным критерием сходства. В случае, если условие сходства не выполняется, модулятор сброса вырабатывает управляющий сигнал для определения нейрона второго слоя, не участвовавшего в запоминании образов, с целью его последующего обучения. Также модулятор сброса определяет момент окончания конкуренции в третьем слое и выдает нейронам третьего слоя управляющий сигнал на прекращение конкуренции. Модулятор сброса связан с выходами всех нейронов третьего слоя связями с единичными весами.

Функционирование и обучение сети. В процессе функционирования сети можно выделить три фазы. В первой фазе на вход сети подается входной вектор $x = (x_1, \dots, x_n)$. Так как сигнал модулятора сброса нулевой, входной вектор проходит через первый слой без изменений и подается во второй слой и модулятор сигнала. Модулятор сигнала вычисляет сумму компонент входного сигнала и формирует сигнал $m_1 = \sum x_i$. Все нейроны второго слоя являются либо связанными (хранящими в памяти некоторый образ), либо несвязанными. Веса связей несвязанного нейрона с нейронами первого слоя принимаются равными -2 . Вектор весов связей w_j связанного нейрона с номером j второго слоя с нейронами первого слоя принимается равным $w_j = -2d$, где d – запомненный вектор. Здесь и далее полагается $w = w_{ij}$, где w_{ij} – вес связи от нейрона с номером i к нейрону с номером j . Таким образом, сумматор каждого нейрона второго слоя вычисляет следующее значение

$$s = \sum d_i + \sum x_i + \sum x_i d_i .$$

Нетрудно убедиться, что вычисленное значение представляет собой не что иное, как расстояние Хемминга между входным вектором x и запомненным вектором d , равное числу отличающихся битов. После применения к значению сумматора функции активации (1) на вход нейронов третьего слоя подаются инвертированные расстояния Хемминга между входным вектором и всеми запомненными образами, лежащие в диапазоне $[0, 1]$.

Вторая фаза является фазой конкуренции нейронов третьего слоя, в процессе которой определяется нейрон, соответствующий образу с наименьшим расстоянием Хемминга до входного вектора x . Процесс конкуренции подробно изложен в [2]. В начале конкуренции сигналы, инициировавшие процесс конкуренции, удаляются путем обнуления входного вектора.

На каждом шаге итеративного конкурентного процесса на выходах нейронов третьего слоя формируется сигнал, поступающий в соответствующие нейроны четвертого слоя. Нейроны четвертого слоя превращают каждый положительный выходной сигнал нейронов третьего слоя в единичный и передают в модулятор сброса. На каждой итерации конкурентного процесса модулятор сброса вычисляет сумму модулей сигналов нейронов третьего слоя, и когда она станет равной единице, что свидетельствует об окончании конкуренции, посылает тормозящий сигнал (отрицательное значение, превы-

шающее по модулю число n) всем нейронам третьего слоя. Этот сигнал останавливает конкуренцию за счет обнуления выходных сигналов нейронов третьего слоя. Одновременно модулятор сброса сравнивает инвертированное расстояние Хемминга между запомненным образом, соответствующим нейрону-победителю, и входным вектором, с заранее заданным числом $\alpha = 1 - \varepsilon/n$. Здесь ε – максимально допустимое количество отличающихся битов во входном векторе и ближайшем по мере Хемминга запомненном векторе. Если ближайший запомненный вектор отличается от входного вектора на большее число битов, модулятор сброса вырабатывает единичный сигнал, который передается на все нейроны первого слоя, устанавливая их выходы в единицу.

В третьей фазе единичный вектор из первого слоя поступает во второй слой. Несложно убедиться, что все несвязанные нейроны будут иметь одинаковые выходы, равные единице. В результате последующей конкуренции нейронов третьего слоя один из них (например, нейрон с номером j) будет признан победителем. После чего соответствующий нейрон второго слоя обучается по правилу: вектор весов его связей с нейронами первого слоя устанавливается равным $w_j = -2x$, где x – входной вектор. Те же значения устанавливаются для вектора весов связей соответствующего нейрона четвертого слоя со всеми нейронами пятого слоя.

Заключение. Предложенная в данной работе модифицированная сеть Хемминга реализована в виде компьютерной программы на языке С. Тестирование программы показало эффективность предложенных архитектуры и алгоритма обучения сети.

Литература

1. Уоссермен Ф. Нейрокомпьютерная техника. М.: Мир, 1992. 240 с.
2. Осовский С. Нейронные сети для обработки информации. М.: Финансы и статистика, 2002. 344 с.

Борисов Е.А.

Санкт-Петербургский государственный университет

**Применение оптического преобразования Фурье
к задаче определения ориентации случайно
расположенного в кадре текстового блока
относительно апертуры объектива**

Рекомендовано к публикации доцентом Гришкиным В.М.

Введение. Важными параметрами для оптимизации в системах видеоконтроля являются сведение к минимуму необходимости управления оператором и упрощение алгоритмов и соответственно временных и вычислительных затрат при обработке информации. В работе исследуется возможность повышения быстродействия системы распознавания случайно ориентированного текста путем введения оптического элемента, позволяющего определить ориентацию текста до выполнения его распознавания. Общие алгоритмы распознавания хорошо известны [1]. Сначала локализуется непосредственно сама область нахождения символов на изображении, затем она сегментируется на участки, содержащие по одному символу, а эти участки уже сравниваются с некоторой базой шаблонов символов. При этом нужно учитывать необходимость очистки исходного изображения от шумов и различия в масштабе. Когда же возможное положение и особенно сама ориентация символов на изображении неизвестны, перед началом работы алгоритмов распознавания необходимо определить возможный угол смещения относительно некоторого нормального положения символов.

Остановимся на последнем вопросе – определении угла поворота. Предложенное далее решение хорошо своей универсальностью и применимо для различных видов символов и, вообще, для определения ориентации практически любых объектов с высокой степенью пространственной периодичности в информации. Оптическая обработка в определении пространственной ориентации информации позволяет значительно уменьшить требования к ресурсам ЭВМ.

1. Оптическое преобразование Фурье. Основывается предлагаемый метод на свойстве положительной двояковыпуклой линзы: в когерентном свете распределение амплитуды излучения в зад-

ней фокальной плоскости линзы может быть представлено как двумерное комплексное преобразование Фурье от функции распределения амплитуды света в передней фокальной плоскости линзы. Дифракционный интеграл, описывающий взаимосвязь распределений поля во входной и выходной плоскостях такой системы, имеет вид [2]

$$F(u, \nu) = \frac{1}{j\lambda a} F[f(x, y)] = \frac{1}{j\lambda a} \iint f(x, y) e^{-j2\pi(xu+y\nu)} dx dy, \quad (1)$$

где $u = \frac{x}{\lambda a}$, $\nu = \frac{y}{\lambda a}$, λ – длина волны, a – фокусное расстояние линзы, $f(x, y)$ – комплексная амплитуда световой волны в передней (входной) фокальной плоскости линзы, $F(u, \nu)$ – комплексная амплитуда световой волны в задней (выходной, спектральной) фокальной плоскости линзы, j – мнимая единица.

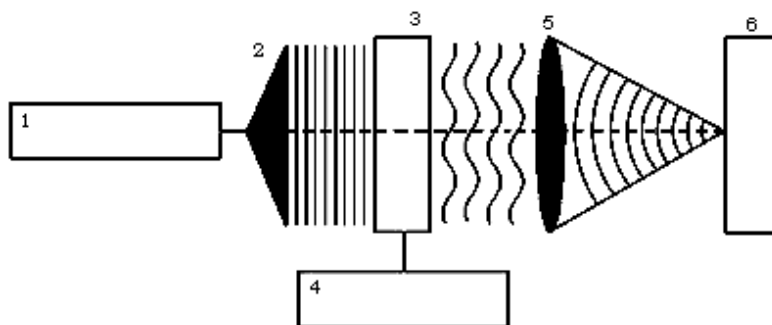


Рис. 1. Схема аналогового оптического вычислительного устройства: 1 – Лазер, 2 – расширитель пучка, 3 – устройство ввода, 4 – блок контроля и подачи сигнала на устройство ввода, 5 – аналоговый оптический вычислитель, 6 – устройство вывода

Лазер генерирует монохроматический свет с высокой степенью когерентности, расширитель пучка формирует из него плоскую волну, а устройство ввода модулирует её входным сигналом по амплитуде и фазе. Аналоговый оптический вычислитель (АОВ) осуществляет преобразование модулированного когерентного сигнала, реализуя требуемую математическую операцию, и представляет результат вычислений в виде распределения интенсивности света в фокальной плоскости Фурье-линзы (АОВ). Устройство вывода информации преобразует результат вычислений в электрический сигнал для последующей обработки цифровой системой [3].

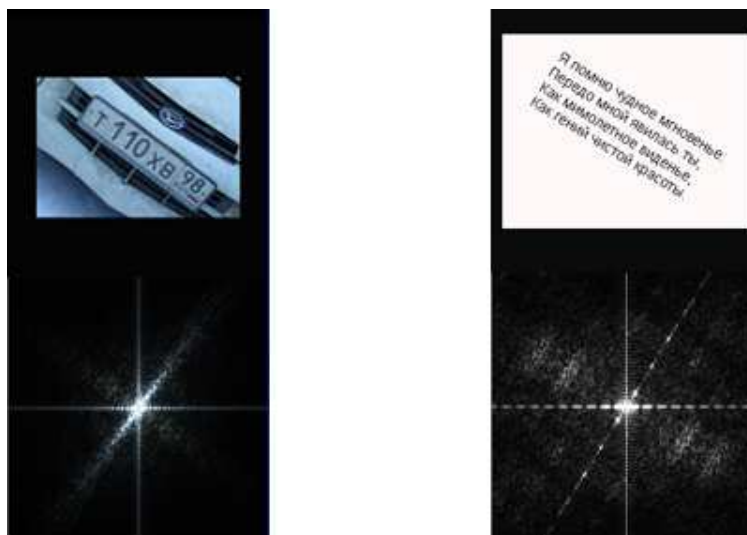


Рис. 2. спектры автомобильного номера (слева) и текста на белом фоне (справа)

В работе описывается математическое моделирование результатов аналоговых преобразований, выполняемых в оптическом диапазоне. Математическая модель преобразования Фурье реализуется с помощью алгоритмов быстрого преобразования Фурье (БПФ). Анализ полученных результатов справедлив и для случая оптического преобразования. Рассмотрим результаты БПФ для двух классических примеров задачи определения ориентации текста: случайно повернутый в кадре автомобильный номер (рис. 2), и блок текста на белом фоне (рис. 2). Разберёмся подробнее в представленных иллюстрациях. Здесь под чёрной рамкой вокруг изображения подразумевается прямоугольная входная апертура объектива. Сравнивая ее Фурье-образ с образом входной информации, определяем взаимную ориентацию объектива и исследуемого объекта.

На рис. 3 представлен образ одной прямоугольной апертуры без изображения. Подробное изложение дифракции Фраунгофера на прямоугольной апертуре представлено в работе [4]. Спектр одной лишь апертуры представляет собой две взаимно перпендикулярные центрально-симметричные линии. Таким образом, различия спек-

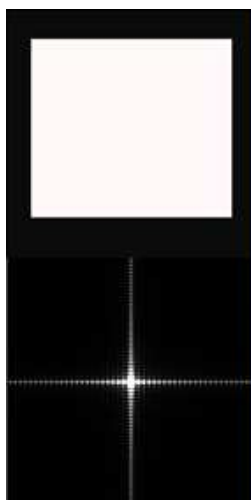


Рис. 3. Спектр одной апертуры

тров на рис. 3 и рис. 2 – есть спектры изображения с автомобильным номером и для картинка с текстом.

$$f(x, y) = \iint F(u, \nu) e^{j2\pi(xu+y\nu)} du d\nu. \quad (2)$$

Преобразование Фурье (2) можно рассматривать как представление функции $f(x, y)$ в виде линейной комбинации элементарных функций типа $e^{-j2\pi(xu+y\nu)}$, которые обладают следующим свойством: для любой пары фиксированных частот (u, ν) фаза соответствующей элементарной функции равна нулю вдоль линии $y = -\frac{u}{\nu}x + \frac{n}{\nu}$, $n \in Z$, т.е. в плоскости Oxy (передней фокальной) она направлена под углом $\theta = \arctg \frac{\nu}{u}$ к горизонтальной оси Ox [5]. При модулировании входным сигналом когерентного излучения текст на изображении выступает в роли набора дифракционных решёток, каждой из которых будет соответствовать своя элементарная функция. Ввиду регулярности и однородности расположения текстовых символов фазовое изменение у соответствующих им элементарных функций будет одинаковым, их Фурье-образы выстроятся на одной прямой.

Преобразование Хафа. Для определения взаимного пространственного положения объекта и входной апертуры оптической системы анализируем полученные спектры, имеющие (как показа-

но выше) в Фурье-плоскости форму прямых линий. Для этого будет удобным использовать преобразование Хафа [6], которое имеет невысокую вычислительную сложность. Перебором локальных максимумов в пространстве Хафа получаем значения параметров соответствующих прямых на изображении. Далее приводятся результаты, полученные в среде MATLAB. Чтобы получить максимально достоверное направление для случая более зашумлённых изображений, осуществляем поиск большего числа локальных максимумов и усредняем параметры для близко расположенных.

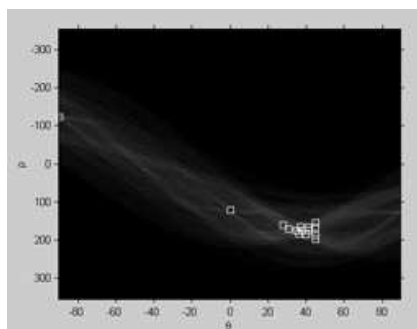


Рис. 4. Выборка 15 максимумов в пространстве Хафа для случая автомобильного номера

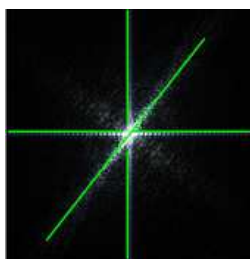


Рис. 5. Построенные по заданным параметрам прямые

Заключение. Полученные прямые образуются однозначным образом для обоих рассмотренных примеров. Тогда, находя угол наклона прямой в системе координат, оси которой образованы двумя взаимно перпендикулярными прямыми от образа апертуры объектива, решаем поставленную задачу и тем самым даём положитель-

ную оценку проверки работоспособности предложенного алгоритма. В дальнейшем остаётся лишь сравнить практические результаты с теоретическими и сделать выводы о целесообразности его использования. Предложенная математическая модель комбинированной системы распознавания информации позволяет решить одну из наиболее ресурсоемких вычислительных задач – распознавание текста с неизвестной ориентацией в реальном времени, разделив задачи ориентации регулярных структур в информационном массиве и непосредственно само распознавание символов.

Литература

1. Распознавание регистрационных номеров автомобилей с помощью MATLAB
http://habrahabr.ru/blogs/image_processing/114519/
2. Дмитриев А.Л. Оптические методы обработки информации. Учебное пособие. СПб.: СПбГУИТМО, 2005. 46 с.
3. Оптическая вычислительная техника (Лекции по ФООЭ)
<http://schembl.net/?p=1287>
4. Каров Д.Д. Дифракция света. Описание к лабораторной работе. СПб.: СПбГПУ Кафедра прикладной физики и оптики твёрдого тела, 2001. 24с.
5. Гудмен Дж. Введение в Фурье-оптику. М.: Мир, 1970. 364с.
6. Дегтярева А., Вежнев В. Сетевой журнал «Компьютерная Графика и Мультимедиа». – Преобразование Хафа (Hough transform)
<http://cgm.computergraphics.ru/content/view/36 Heading19>

Быков К.Н.

Санкт-Петербургский государственный университет

**Инвариантность калибровочных соотношений
для сплайнов эрмитова типа
при различном порядке выбрасывания узлов**

Рекомендовано к публикации профессором Демьяновичем Ю.К.

Введение. Основная проблема сжатия информационных потоков состоит в создании алгоритмов их расщепления на основную и вейвлетную составляющие. Использование производных или разностных отношений улучшает аппроксимационные характеристики таких разложений (см. [1]). Для получения сплайн-вейвлетных разложений существенны калибровочные соотношения.

Цель данной работы – показать, что калибровочные соотношения не зависят от порядка выбрасывания сеточных узлов.

1. Сплайны эрмитова типа. Рассмотрим четырех-компонентную вектор-функцию $\varphi(t)$ класса $C^1(\alpha, \beta)$, удовлетворяющую условию

$$W(x, y; \varphi) \stackrel{\text{def}}{=} \det(\varphi(x), \varphi'(x), \varphi(y), \varphi'(y)) \neq 0 \quad (1)$$
$$\forall x, y \in (\alpha, \beta), x \neq y.$$

Пусть X – сетка вида

$$X : \dots < x_{-1} < x_0 < x_1 < \dots; \quad \alpha \stackrel{\text{def}}{=} \lim_{j \rightarrow -\infty} x_j, \quad \beta \stackrel{\text{def}}{=} \lim_{j \rightarrow +\infty} x_j. \quad (2)$$

Введем обозначения

$$G \stackrel{\text{def}}{=} \bigcup_{j \in \mathbb{Z}} (x_j, x_{j+1}), \quad \varphi_j \stackrel{\text{def}}{=} \varphi(x_j), \quad \varphi'_j \stackrel{\text{def}}{=} \varphi'(x_j).$$

Рассмотрим функции $\omega_j(t)$, $t \in G$, $j \in \mathbb{Z}$, удовлетворяющие аппроксимационным соотношениям

$$\sum_j (\varphi'_{j+1} \omega_{2j-1}(t) + \varphi_{j+1} \omega_{2j}(t)) = \varphi(t), \quad (3)$$

предполагая, что

$$\text{supp} \omega_{2j-1} \subset [x_j, x_{j+2}], \quad \text{supp} \omega_{2j} \subset [x_j, x_{j+2}] \quad \forall j \in \mathbb{Z}. \quad (4)$$

При фиксированном $k \in \mathbb{Z}$ из (3), (4) для $t \in (x_k, x_{k+1})$ получаем

$$\varphi'_k \omega_{2k-3}(t) + \varphi_k \omega_{2k-2}(t) + \varphi'_{k+1} \omega_{2k-1}(t) + \varphi_{k+1} \omega_{2k}(t) = \varphi(t). \quad (5)$$

Благодаря свойству (1) система (5) однозначно разрешима. При $t \in (x_k, x_{k+1})$ из (5), последовательно полагая $k = q$, $k = q + 1$, находим

$$\omega_{2q-1}(t) = \frac{\det(\varphi'_q, \varphi_q, \varphi(t), \varphi_{q+1})}{\det(\varphi'_q, \varphi_q, \varphi'_{q+1}, \varphi_{q+1})} \quad \text{при } t \in (x_q, x_{q+1}), \quad (6)$$

$$\omega_{2q-1}(t) = \frac{\det(\varphi(t), \varphi_{q+1}, \varphi'_{q+2}, \varphi_{q+2})}{\det(\varphi'_{q+1}, \varphi_{q+1}, \varphi'_{q+2}, \varphi_{q+2})} \quad \text{при } t \in (x_{q+1}, x_{q+2}), \quad (7)$$

$$\omega_{2q}(t) = \frac{\det(\varphi'_q, \varphi_q, \varphi'_{q+1}, \varphi(t))}{\det(\varphi'_q, \varphi_q, \varphi'_{q+1}, \varphi_{q+1})} \quad \text{при } t \in (x_q, x_{q+1}), \quad (8)$$

$$\omega_{2q}(t) = \frac{\det(\varphi'_{q+1}, \varphi(t), \varphi'_{q+2}, \varphi_{q+2})}{\det(\varphi'_{q+1}, \varphi_{q+1}, \varphi'_{q+2}, \varphi_{q+2})} \quad \text{при } t \in (x_{q+1}, x_{q+2}) \quad (9)$$

для любого $q \in \mathbb{Z}$.

Пространство

$$\mathbb{S}_\varphi^1(X) \stackrel{\text{def}}{=} \left\{ u \mid u = \sum_j c_j \omega_j \quad \forall c_j \in \mathbb{R}^1, j \in \mathbb{Z} \right\} \quad (10)$$

называется пространством сплайнов эрмитова типа (первой высоты), а множество $\{\omega_j\}_{j \in \mathbb{Z}}$ называется главным базисом пространства $\mathbb{S}_\varphi^1(X)$.

2. Калибровочные соотношения для сплайнов эрмитова типа. Пусть \tilde{X} – сетка, полученная из сетки (2) выбрасыванием узла x_k . Обозначим через \tilde{x}_j узлы вновь полученной сетки:

$$\tilde{x}_j \stackrel{\text{def}}{=} x_j \quad \text{при } j \leq k-1, \quad \tilde{x}_j \stackrel{\text{def}}{=} x_{j+1} \quad \text{при } j \geq k. \quad (11)$$

По только что введенной новой сетке \tilde{X} построим функции $\tilde{\omega}_j$ по формулам, аналогичным формулам (6)–(9):

$$\tilde{\omega}_{2q-1}(t) = \frac{\det(\tilde{\varphi}'_q, \tilde{\varphi}_q, \varphi(t), \tilde{\varphi}_{q+1})}{\det(\tilde{\varphi}'_q, \tilde{\varphi}_q, \tilde{\varphi}'_{q+1}, \tilde{\varphi}_{q+1})} \quad \text{при } t \in (\tilde{x}_q, \tilde{x}_{q+1}), \quad (12)$$

$$\tilde{\omega}_{2q-1}(t) = \frac{\det(\varphi(t), \tilde{\varphi}_{q+1}, \tilde{\varphi}'_{q+2}, \tilde{\varphi}_{q+2})}{\det(\tilde{\varphi}'_{q+1}, \tilde{\varphi}_{q+1}, \tilde{\varphi}'_{q+2}, \tilde{\varphi}_{q+2})} \quad \text{при } t \in (\tilde{x}_{q+1}, \tilde{x}_{q+2}), \quad (13)$$

$$\tilde{\omega}_{2q}(t) = \frac{\det(\tilde{\varphi}'_q, \tilde{\varphi}_q, \tilde{\varphi}'_{q+1}, \varphi(t))}{\det(\tilde{\varphi}'_q, \tilde{\varphi}_q, \tilde{\varphi}'_{q+1}, \tilde{\varphi}_{q+1})} \quad \text{при } t \in (\tilde{x}_q, \tilde{x}_{q+1}), \quad (14)$$

$$\tilde{\omega}_{2q}(t) = \frac{\det(\tilde{\varphi}'_{q+1}, \varphi(t), \tilde{\varphi}'_{q+2}, \tilde{\varphi}_{q+2})}{\det(\tilde{\varphi}'_{q+1}, \tilde{\varphi}_{q+1}, \tilde{\varphi}'_{q+2}, \tilde{\varphi}_{q+2})} \quad \text{при } t \in (\tilde{x}_{q+1}, \tilde{x}_{q+2}), \quad (15)$$

где $\tilde{\varphi}_j = \varphi(\tilde{x}_j)$, $\tilde{\varphi}'_j = \varphi'(\tilde{x}_j)$. Учитывая, что в принятых обозначениях

$$\tilde{\varphi}_q = \varphi_q \quad \text{при } q \leq k-1, \quad \tilde{\varphi}_{q-1} = \varphi_q \quad \text{при } q \geq k+1,$$

нетрудно видеть, что

$$\begin{aligned} \tilde{\omega}_i(t) &\equiv \omega_i(t) & \text{при } i &\leq 2k-6, \\ \tilde{\omega}_i(t) &\equiv \omega_{i+2}(t) & \text{при } i &\geq 2k-1. \end{aligned} \quad (16)$$

Получим представления для функций $\tilde{\omega}_{2k-5}$, $\tilde{\omega}_{2k-4}$, $\tilde{\omega}_{2k-3}$ и $\tilde{\omega}_{2k-2}$.

Теорема 1. Если выполнено условие (1), то при $t \in (\alpha, \beta)$ справедливы тождества

$$\begin{aligned} \tilde{\omega}_{2k-5}(t) &\equiv \omega_{2k-5}(t) + \\ &+ \tilde{\omega}'_{2k-5}(x_k) \omega_{2k-3}(t) + \tilde{\omega}_{2k-5}(x_k) \omega_{2k-2}(t), \end{aligned} \quad (17)$$

$$\begin{aligned} \tilde{\omega}_{2k-4}(t) &\equiv \omega_{2k-4}(t) + \\ &+ \tilde{\omega}'_{2k-4}(x_k) \omega_{2k-3}(t) + \tilde{\omega}_{2k-4}(x_k) \omega_{2k-2}(t), \end{aligned} \quad (18)$$

$$\begin{aligned} \tilde{\omega}_{2k-3}(t) &\equiv \tilde{\omega}'_{2k-3}(x_k) \omega_{2k-3}(t) + \\ &+ \tilde{\omega}_{2k-3}(x_k) \omega_{2k-2}(t) + \omega_{2k-1}(t), \end{aligned} \quad (19)$$

$$\tilde{\omega}_{2k-2}(t) \equiv \tilde{\omega}'_{2k-3}(x_k) \omega_{2k-3}(t) + \tilde{\omega}_{2k-2}(x_k) \omega_{2k-2}(t) + \omega_{2k}(t). \quad (20)$$

Доказательство. Покажем справедливость тождества (17) во всех точках $t = \tilde{x}_j$, $j \in \mathbb{Z}$.

Пусть $j < k-2$. В этом случае по формулам (11) имеем $x_j = \tilde{x}_j$, левая часть и все слагаемые правой части равны нулю в силу задания функций ω_j и $\tilde{\omega}_j$.

Пусть $j = k-2$. В [1] показано, что введенные функции (6)–(9) и (12)–(15) можно продолжить по непрерывности в точках x_j и \tilde{x}_j следующим образом:

$$\begin{aligned}
\omega_{2q-1}(x_q) &= 0, & \omega_{2q-1}(x_{q+1}) &= 0, & \omega_{2q-1}(x_{q+2}) &= 0, \\
\omega'_{2q-1}(x_q) &= 0, & \omega'_{2q-1}(x_{q+1}) &= 1, & \omega'_{2q-1}(x_{q+2}) &= 0, \\
\omega_{2q}(x_q) &= 0, & \omega_{2q}(x_{q+1}) &= 1, & \omega_{2q}(x_{q+2}) &= 0, \\
\omega'_{2q}(x_q) &= 0, & \omega'_{2q}(x_{q+1}) &= 0, & \omega'_{2q}(x_{q+2}) &= 0, \\
\tilde{\omega}_{2q}(\tilde{x}_q) &= 0, & \tilde{\omega}_{2q}(\tilde{x}_{q+1}) &= 1, & \tilde{\omega}_{2q}(\tilde{x}_{q+2}) &= 0, \\
\tilde{\omega}'_{2q}(\tilde{x}_q) &= 0, & \tilde{\omega}'_{2q}(\tilde{x}_{q+1}) &= 0, & \tilde{\omega}'_{2q}(\tilde{x}_{q+2}) &= 0, \\
\tilde{\omega}_{2q-1}(\tilde{x}_q) &= 0, & \tilde{\omega}_{2q-1}(\tilde{x}_{q+1}) &= 0, & \tilde{\omega}_{2q-1}(\tilde{x}_{q+2}) &= 0, \\
\tilde{\omega}'_{2q-1}(\tilde{x}_q) &= 0, & \tilde{\omega}'_{2q-1}(\tilde{x}_{q+1}) &= 1, & \tilde{\omega}'_{2q-1}(\tilde{x}_{q+2}) &= 0.
\end{aligned} \tag{21}$$

Кроме того, $x_{k-2} = \tilde{x}_{k-2}$. Очевидно что в поставленных условиях обе части (17) обращаются в ноль. Аналогичное тождество получается и при $j = k - 1$.

Пусть $j = k$. В правой части единственное ненулевое слагаемое равно $\tilde{\omega}_{2k-5}(x_k)$, что соответствует значению правой части в точке x_k .

Пусть $j \geq k + 1$. В этом случае обе части (17) обращаются в ноль при $j = k + 1$ в силу (21) и равенства $x_{k+1} = \tilde{x}_k$, а при $j > k + 1$ – в силу задания.

Теперь необходимо установить справедливость тождества

$$\begin{aligned}
&\tilde{\omega}'_{2k-5}(t) \equiv \omega'_{2k-5}(t) + \\
&+ \tilde{\omega}'_{2k-5}(x_k) \omega'_{2k-3}(t) + \tilde{\omega}_{2k-5}(x_k) \omega'_{2k-2}(t).
\end{aligned} \tag{22}$$

Это нетрудно сделать, проводя аналогичные рассуждения для всех точек $t = \tilde{x}_j$. Обе части равенства (22) обращаются в ноль за исключением точки x_k , дающей тождество $\tilde{\omega}'_{2k-5}(x_k) = \tilde{\omega}'_{2k-5}(x_k)$.

Аналогичным образом могут быть доказаны тождества (18)–(20). Теорема доказана.

Следствие 1. *В условиях теоремы 1 справедливы соотношения*

$$\tilde{\omega}_i(t) \equiv \sum_{j \in \mathbb{Z}} \mathbf{p}_{ij} \omega_j(t),$$

где

$$\begin{aligned}
\mathfrak{p}_{ij} &= \delta_{i,j} \quad \text{при } i \leq 2k - 6 \quad \forall j \in \mathbb{Z}, \\
\mathfrak{p}_{ij} &= \delta_{i,j+2} \quad \text{при } i \geq 2k - 1 \quad \forall j \in \mathbb{Z}, \\
\mathfrak{p}_{2k-5,j} &= \delta_{2k-5,j} \quad \text{при } j \leq 2k - 4 \vee j \geq 2k - 1, \\
\mathfrak{p}_{2k-4,j} &= \delta_{2k-4,j} \quad \text{при } j \leq 2k - 3 \vee j \geq 2k - 1, \\
\mathfrak{p}_{2k-3,j} &= \delta_{2k-3,j-2} \quad \text{при } j \leq 2k - 4 \vee j \geq 2k - 1, \\
\mathfrak{p}_{2k-2,j} &= \delta_{2k-3,j-2} \quad \text{при } j \leq 2k - 4 \vee j \geq 2k, \\
\mathfrak{p}_{2k-5,2k-3} &= \tilde{\omega}'_{2k-5}(x_k), \quad \mathfrak{p}_{2k-5,2k-2} = \tilde{\omega}_{2k-5}(x_k), \\
\mathfrak{p}_{2k-4,2k-3} &= \tilde{\omega}'_{2k-4}(x_k), \quad \mathfrak{p}_{2k-4,2k-2} = \tilde{\omega}_{2k-4}(x_k), \\
\mathfrak{p}_{2k-3,2k-3} &= \tilde{\omega}'_{2k-3}(x_k), \quad \mathfrak{p}_{2k-3,2k-2} = \tilde{\omega}_{2k-3}(x_k), \\
\mathfrak{p}_{2k-2,2k-3} &= \tilde{\omega}'_{2k-3}(x_k), \quad \mathfrak{p}_{2k-2,2k-2} = \tilde{\omega}_{2k-2}(x_k).
\end{aligned} \tag{23}$$

Формулы (23) называются калибровочными соотношениями.

Следствие 2. В условиях теоремы 1 пространство

$$\mathbb{S}_\varphi^1(\tilde{X}) \stackrel{\text{def}}{=} \left\{ u \mid u = \sum_j c_j \tilde{\omega}_j \quad \forall c_j \in \mathbb{R}^1, j \in \mathbb{Z} \right\}$$

содержится в пространстве $\mathbb{S}_\varphi^1(X)$.

3. Инвариантность калибровочных соотношений и биортогональная система функционалов. Выбросим теперь из сетки (2) узел x_s , полагая для определенности $s < k$. Обозначим через \hat{x}_j узлы вновь полученной сетки \hat{X} :

$$\begin{aligned}
\hat{x}_j &\stackrel{\text{def}}{=} x_j \quad \text{при } j \leq s - 1, \quad \hat{x}_j \stackrel{\text{def}}{=} x_{j+2} \quad \text{при } j \geq k - 1, \\
\hat{x}_j &\stackrel{\text{def}}{=} x_{j+1} \quad \text{при } s \leq j \leq k - 2.
\end{aligned} \tag{24}$$

Функции $\hat{\omega}_j$ построим по узлам сетки \hat{x}_j аналогично формулам (6)–(9) и (12)–(15). Эти функции связаны с функциями $\{\omega_j\}_{j \in \mathbb{Z}}$ калибровочными соотношениями вида

$$\hat{\omega}_i(t) \equiv \sum_{j \in \mathbb{Z}} \hat{\mathfrak{p}}_{ij} \omega_j(t), \tag{25}$$

где коэффициенты $\hat{\mathfrak{p}}_{ij}$ вычисляются аналогичным описанному в теореме способом. Покажем, что вид этих соотношений не зависит от порядка выброса точек.

Теорема 2. *Калибровочные соотношения (25) не зависят от порядка выброса точек x_k и x_s .*

Доказательство. Над пространством $C^1(\alpha, \beta)$ рассмотрим систему линейных функционалов $\{g^{(i)}\}_{i \in \mathbb{Z}}$, определяемых соотношениями

$$\langle g^{(2q-1)}, u \rangle \stackrel{\text{def}}{=} u'(x_{q+1}), \quad \langle g^{(2q)}, u \rangle \stackrel{\text{def}}{=} u(x_{q+1}) \quad \forall q \in \mathbb{Z}.$$

Ввиду формул (21) имеем $\langle g^{(i)}, \omega_j \rangle = \delta_{i,j}$. Подействуем на (25) функционалом g_k :

$$\langle g_k, \hat{\omega}_i \rangle = \sum_{j \in \mathbb{Z}} \hat{\mathfrak{p}}_{ij} \langle g_k, \omega_j \rangle = \hat{\mathfrak{p}}_{ik}.$$

Таким образом, калибровочные соотношения (25) полностью определяются функциями $\{\hat{\omega}_j\}_{j \in \mathbb{Z}}$, вид которых зависит только от номеров, но не от порядка выбрасываемых точек.

Теорема доказана.

Литература

1. Зимин А.В., Демьянович Ю.К. О всплесковом разложении сплайнов эрмитова типа // Проблемы математического анализа, 2007. Вып. 35. С. 33–45.

Гузницева А.П.

Санкт-Петербургский государственный университет

Программная реализация позиционного управления нелинейным объектом

Рекомендовано к публикации профессором Смирновым Н.В.

1. Постановка задачи оптимального управления. Задача оптимального управления нелинейным объектом является одной из актуальных задач теории оптимального управления. В данной статье предлагается обзор и реализация одного подхода к управлению нелинейными объектами. Этот подход был предложен Р. Габасовым и Ф.М. Кирилловой [1] и активно разрабатывался ими и их учениками.

Рассматривается следующая задача оптимального управления

$$\begin{cases} \dot{x} = F(t, x) + bu, & t \in [t_*, t^*], & x \in \mathbb{R}^n \\ x(t_*) = x_0, \\ Hx(t^*) = g, & \text{rank} H = m. \end{cases} \quad (1)$$

В качестве допустимых управлений рассмотрим дискретные управления с фиксированным количеством N точек переключения

$$\begin{cases} u(t) = u(t_0 + kh), & t \in [t_0 + kh, t_0 + (k+1)h), \\ k = \overline{0, N-1}. \end{cases}$$

Функционал качества зададим следующим образом

$$\begin{aligned} J(u) = \mu \rightarrow \min, \\ \mu d_* + d_0 \leq u(t) \leq d_0 + \mu d^*, \quad \mu \geq 0. \end{aligned}$$

Такой функционал позволяет нам не только находить в определенном смысле оптимальное управление, но и позволяет оценивать конструктивные возможности решения задачи управления нелинейным объектом при выбранной нами линеаризации. Таким образом, сформулирована задача построения управления нелинейной системой минимальной интенсивности. Введем понятие оптимального управления типа обратной связи (позиционного решения) задачи (1). Для

этого погрузим ее в множество задач

$$\begin{cases} \dot{x} = F(t, x) + bu, & t \in T(\tau) = [\tau, t^*], & x \in \mathbb{R}^n, \\ x(\tau) = z, \\ Hx(t^*) = g, & \text{rank} H = m, \\ J(u) = \mu \rightarrow \min, \\ \mu d_* + d_0 \leq u(t) \leq d_0 + \mu d^*, & \mu \geq 0. \end{cases} \quad (2)$$

Пусть $u^0(t|\tau, z)$ – оптимальное программное управление задачи (2) для положения (τ, z) . Тогда функцию

$$u^0(\tau, z) = u^0(\tau|\tau, z) \quad (3)$$

будем называть оптимальным управлением типа обратной связи или позиционным решением задачи (1).

Предположим, что функция (3) построена. Замкнем ей систему (1) и рассмотрим поведение полученной системы в условиях заранее неизвестных непрерывно действующих ограниченных возмущений $w(t)$:

$$\begin{cases} \dot{x} = F(t, x) + bu^0(t, x) + w(t), & t \in [t_*, t^*], \\ x(t_*) = x_0. \end{cases} \quad (4)$$

Под траекторией (4) будем понимать решение системы

$$\begin{cases} \dot{x} = F(t, x) + bu(t) + w(t), & t \in [t_*, t^*], \\ x(t_*) = x_0, \\ u(t) = u^0(\tau, x(\tau)), & t \in [\tau, \tau + h), \\ \tau = t_* + kh, & k = 0, \dots, N-1. \end{cases}$$

В процессе управления системой (4) реализуется только обратная связь вдоль конкретной траектории, причем не нужно знать ее заранее, достаточно уметь вычислять ее в конкретные моменты времени. Будем говорить, что построение $u^*(t)$ (реализация оптимальной обратной связи) осуществляется в режиме реального времени, если в каждой текущей позиции $(\tau, x^*(\tau))$ время $s(\tau)$ вычисления $u^*(\tau) = u^0(\tau, x^*(\tau))$ не превосходит h . Устройство, способное выполнить такую работу, называется оптимальным регулятором [2, 3].

Построение позиционного управления осуществляется в два этапа: 1) решение вспомогательных задач оптимального программного управления; 2) коррекция построенных программных движений с учетом поступающей информации в режиме реального времени.

2. Синтез оптимальной обратной связи в линейном случае. Рассмотрим задачу оптимального управления линейной системой (ОУ) при наличии ограниченных возмущений:

$$\begin{cases} \dot{x} = A(t)x + b(t)u + w(t), & x(t_*) = x_0, \\ Hx(t^*) = g, \\ J(u) = \mu \rightarrow \min, \\ \mu d_* + d_0 \leq u(t) \leq d_0 + \mu d^*, \quad \mu \geq 0, \end{cases}$$

где $x \in \mathbb{R}^n$, $u \in \mathbb{R}$, $|u(t)| \leq 1$, $g \in \mathbb{R}^m$, $\text{rank} H = m < n$, $t \in [t_*, t^*]$. Матрицы $A(t)$ и $b(t)$ будем полагать кусочно-непрерывными на $[t_*, t^*]$.

Поставим вспомогательную задачу оптимального программного управления:

$$\begin{cases} \dot{x} = A(t)x + b(t)u, & x(t_*) = x_0, \\ Hx(t^*) = g, \\ J(u) = \mu \rightarrow \min, \\ \mu d_* + d_0 \leq u(t) \leq d_0 + \mu d^*, \quad \mu \geq 0. \end{cases} \quad (5)$$

Воспользовавшись формулой Коши, получим функциональную форму этой задачи

$$\begin{cases} \sum_{t \in T_u} d_h(t)u(t) = g_0, \\ J(u) = \mu \rightarrow \min, \\ \mu d_* + d_0 \leq u(t) \leq d_0 + \mu d^*, \quad \mu \geq 0. \end{cases}$$

Здесь

$$d_h(t) = \int_t^{t+h} G(\xi)b(\xi)d\xi, \\ g_0 = g - Hx_1(t^*),$$

$x_1(t)$ – траектория системы (5) при $u(t) = 0$, $t \in T$, $G(t)$ – $(m \times n)$ -матричная функция, удовлетворяющая системе

$$\dot{G} = -GA(t), \quad G(t^*) = H.$$

Вместо переменных u_i и μ введем новые переменные

$$\xi_i = \frac{u_i - d_0}{\mu}, \quad i = \overline{1, N},$$

$$\xi_{N+1} = \frac{1}{\mu},$$

и получим задачу линейного программирования

$$J(\Xi) = \xi_{N+1} \rightarrow \max,$$

$$\tilde{D}\Xi = 0,$$

$$d_* \leq \xi_i \leq d^*, \quad i = \overline{1, N}, \quad \xi_{N+1} \geq 0,$$

где $\Xi = (\xi_1, \dots, \xi_{N+1})^T$, \tilde{D} составлена из $d_h(t)$, $t \in T_u$ и столбца $-g_0$.

Итак, получена задача линейного программирования, которую предлагается решать адаптивным методом [3, 4]. Предположим, что оптимальный регулятор уже отработал на k промежутках. Чтобы получить значение обратной связи на следующем промежутке, необходимо воспользоваться процедурой двойственной коррекции опоры, полученной на предыдущем шаге [3].

3. Оптимальное управление нелинейной системой. Разделим область непрерывности Ω функции $F(t, x)$ на области Ω_i , $i = 1, \dots, k$, в каждой из которых зададим линейное приближение функции $F(t, x)$.

Решать задачу предлагается следующим образом. Предварительно анализируя начальные условия, выбираем и фиксируем области Ω_i , $i \in I$, по которым пройдет траектория, и устанавливаем количество l переходов между областями. Затем решаем следующую задачу оптимального программного управления кусочно-линейной системой

$$\begin{cases} \dot{x} = A_i(t)x + bu, & t \in [t_*, t^*], \quad x \in \Omega_i, \quad i \in I, \\ Hx(t^*) = g, \\ x(\theta_i) = x_i, \quad i = 1, \dots, l, \\ J(u) = \mu \rightarrow \min, \\ \mu d_* + d_0 \leq u(t) \leq d_0 + \mu d^*, \quad \mu \geq 0. \end{cases}$$

Сведение задачи к задаче линейного программирования осуществляется аналогично линейному случаю. Оптимизация по моментам

перехода θ_i осуществляется любым доступным способом, например, перебором. Затем регулятор начинает работу. На следующем шаге регулятор получает данные о реальном состоянии нелинейной системы и решает новую задачу оптимального программного управления кусочно-линейной системой, используя полученную на предыдущем шаге опору и, возможно, корректируя моменты перехода.

Численные эксперименты показали, что коррекция происходит достаточно быстро. Предварительный расчет траектории и оптимизация моментов перехода достаточно трудоемка, однако эти расчеты являются предварительными и могут быть выполнены до начала процесса управления.

4. Программная реализация. Примеры расчетов. Программная реализация предложенного метода оптимизации нелинейных систем была реализована в среде MATLAB. Пользователю необходимо задавать кусочно-линейную систему, аппроксимирующую нелинейную, и области фазового пространства, по которым пройдет траектория. Были получены результаты расчетов для маятника с большими начальными отклонениями.

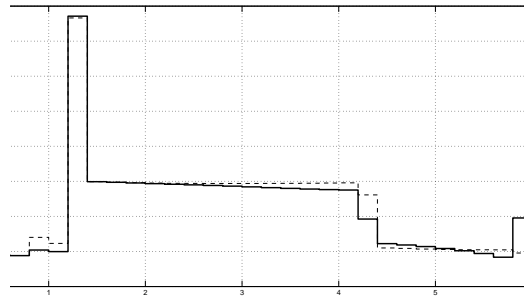


Рис. 1. Значение управления

Математическая модель маятника с приложенным к его оси управляющим моментом u имеет вид

$$\ddot{x} + \sin x = u.$$

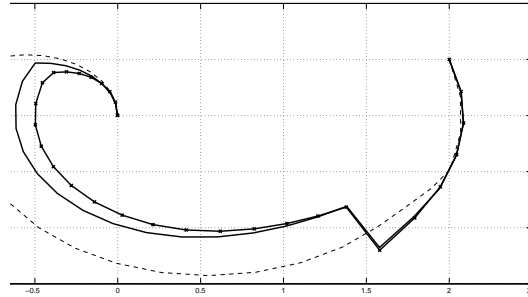


Рис. 2. Фазовая плоскость на периодическую кусочно-линейную функцию

$$f(x) = \begin{cases} x, & |x| \leq \pi/2, \\ -\pi - x, & -3\pi/2 \leq x \leq -\pi/2, \\ \pi - x, & \pi/2 \leq x \leq 3\pi/2. \end{cases}$$

Значение управления представлено на рис. 1, фазовый портрет кусочно-линейной (пунктирная линия), нелинейной с возмущением (сплошная) и нелинейной без возмущения (сплошная с метками) систем приведен на рис. 2. Начальное состояние $x = 2$, $\dot{x} = 0,5$.

Литература

1. Габасов Р., Кириллова Ф.М., Ружицкая Е.А. Демпфирование и стабилизация маятника при больших начальных возмущениях // Изв. АН, Теория и системы управления, 2001. №1. С. 29–38
2. Балашевич Н.В., Габасов Р., Кириллова Ф.М. Численные методы программной и позиционной оптимизации кусочно-линейных систем // Журн. вычисл. математики и мат. физики, 2001. Т. 41. №11. С. 1658–1674.
3. Балашевич Н.В., Габасов Р., Кириллова Ф.М. Численные методы программной и позиционной оптимизации линейных систем // Журн. вычисл. математики и мат. физики, 2000. Т. 40. №6. С. 838–859.
4. Альсевич В.В., Габасов Р., Глушенко В.С. Оптимизация линейных экономических моделей. Статические задачи. Минск: БГУ, 2000. 210 с.

Дрокин И.С.

Санкт-Петербургский государственный университет

Компьютерный анализ робастных свойств сверхустойчивых замкнутых систем

Рекомендовано к публикации профессором Веремеем Е.И.

1. Введение. Целью работы является формирование эффективной компьютерной поддержки для реализации идей обеспечения сверхустойчивости, предложенной в [1], с учетом ее средств и возможностей. Будем рассматривать ЛТИ-систему

$$\dot{x} = Ax + Bu, \quad y = Cx, \quad u = Ky, \quad (1)$$

где $x \in E^n$ – вектор состояния системы, $u \in E^1$ – управляемый вход, $y \in E^r$ – измеряемый выход.

Будем говорить, что матрица $A + BKC$ замкнутой системы (1) является сверхустойчивой, если

$$\sigma = \sigma(A + BKC) = \min_i \left(-a_{ii} - \sum_{j \neq i} |a_{ij}| \right) > 0.$$

Систему со сверхустойчивой матрицей также будем называть сверхустойчивой. Заметим, что сверхустойчивая матрица является устойчивой в обычном смысле, что следует из теоремы о кругах Гершгорина [2].

Задача о сверхстабилизации ставится следующим образом. Необходимо найти пару матриц $(K_{1 \times r}, T_{n \times n})$ таких, что:

- 1) $\det T \neq 0$;
- 2) в новых переменных состояния $z = Tx$ матрица замкнутой системы $T(A + BKC)T^{-1}$ является сверхустойчивой.

2. Робастное управление с параметрической неопределенностью. Рассмотрим линейную стационарную систему с математической моделью

$$\dot{x} = \tilde{A}x + Bu, \quad y = Cx, \quad u = Ky, \quad (2)$$

где $x \in E^n$, $u \in E^m$, $y \in E^r$, $\tilde{A} = A + \gamma\Delta$, $\Delta = \{\Delta_{ij}\}_{i,j=1}^n$, $|\Delta_{ij}| \leq m_{ij}$, $M = \{m_{ij}\}_{i,j=1}^n$.

Поставим вопрос о нахождении радиуса устойчивости семейства в предположении, что $A + BKC$ гурвицева, т.е. нахождения такого γ^* , что устойчивость сохранялась бы при любом $\gamma \leq \gamma^*$.

Теорема 1. Пусть пара матриц K, T представляет решение задачи о сверхстабилизации системы

$$\dot{x} = Ax + Bu, \quad y = Cx, \quad u = Ky. \quad (3)$$

Тогда радиус устойчивости семейства $(A + BKC) + \gamma\Delta$ определяется следующим соотношением

$$\gamma = \min_i \frac{a_{ii} - \sum_{j \neq i} |a_{ij}|}{\sum_j \tilde{m}_{ij}}, \quad (4)$$

где a_{ij} – элементы матрицы $T(A + BKC)T^{-1}$, $|(T\Delta T^{-1})| = \{\tilde{m}_{ij}\}$.

Доказательство. Сделаем замену $x = T^{-1}z$ в системе (3) и замкнем ее обратной связью. Получим

$$\dot{z} = (T(A + BKC)T^{-1} + \gamma T\Delta T^{-1})z. \quad (5)$$

Обозначим $\tilde{\Delta} = T\Delta T^{-1}$. Поскольку $|\Delta_{ij}| \leq m_{ij}$, то $|\tilde{\Delta}_{ij}| \leq \tilde{m}_{ij}$, где \tilde{m}_{ij} из условия теоремы.

По условию теоремы матрица $T(A + BKC)T^{-1}$ гурвицева. Тогда из соотношения, приведенного в [1], получаем утверждение теоремы.

Задача нахождения радиуса устойчивости семейства может быть решена при помощи следующего алгоритма.

1. Построить начальные значения матриц T и K , сверхстабилизирующую систему (3).

2. Используя полученные начальные данные, максимизировать значение

$$\gamma = \min_i \frac{a_{ii} - \sum_{j \neq i} |a_{ij}|}{\sum_j \tilde{m}_{ij}} \rightarrow \max,$$

$$-A_{i,i}^* - \sum_{j \neq i} |A_{i,j}^*| \geq \sigma, \quad i = \overline{1, n}.$$

где a_{ij} – элементы матрицы $A^* = T(A + BKC)T^{-1}$, $\{\tilde{m}_{ij}\}$ взяты из условия теоремы, переменными являются элементы матрицы T .

3. Если требуемый результат не достигнут, получить новые начальные значения матрицы T и повторить п. 2.

3. Робастное управление с неструктурированной неопределенностью. Рассмотрим систему (1), сверхстабилизированную парой (K, T) . Пусть задана весовая функция $W_d(s)$, $P_n(s)$ – передаточная функция системы (1). Пусть математическая модель объекта подвергается возмущению $\Delta(s)$. Рассмотрим любой объект с возмущенной моделью, имеющей передаточную функцию

$$P(s) = P_n(s)(1 + \Delta(s)W_d(s)),$$

и введём допустимое множество передаточных функций этих моделей

$$\Phi = \{P : P = P_n(1 + \Delta W_d), \quad |\Delta(j\omega)| \leq 1 \quad \forall \omega \in R^1\}.$$

Дополнительно потребуем, чтобы при любых возмущениях Δ количество полюсов передаточной функции $P = P_n(1 + \Delta W_d)$, находящихся в правой полуплоскости, совпадало с их количеством для номинальной передаточной функции P_n .

Согласно результату, приведенному в [3], регулятор $u = Ky$ будет стабилизировать возмущенный объект при любом допустимом возмущении, если

$$\alpha = \max_{\omega \in R^1} m(\omega) = \|M(s)\|_\infty < 1, \quad (6)$$

где $M(s) = W_d(s)K[1 - P_n(s)K]^{-1}P_n(s)$.

Матрица $H(s) = K[1 - P_n(s)K]^{-1}P_n(s)$ соответствует передаточной матрице замкнутой системы номинального объекта с матрицей C , слева умноженной на K .

Оценим $\|M(s)\|_\infty \leq \|W_d(s)\|_\infty \left\| K[1 - P_n(s)K]^{-1}P_n(s) \right\|_\infty$.

$$H_{ij}(i\omega) = \int_0^\infty C_i e^{A_c t} B_j e^{-i\omega t} dt,$$

где C_i – i -ая строка матрицы KCT^{-1} , B_j – j -ый столбец матрицы

TB . Оценим по абсолютной величине, используя результат, приведенный в [1]:

$$\begin{aligned} |H_{ij}(i\omega)| &\leq \int_0^{\infty} |C_i e^{A_c t} B_j e^{-i\omega t}| dt \leq \{ |e^{A_c t}| \leq e^{-\sigma t}, \quad \sigma = \sigma(A_c) \} \leq \\ &\leq \max_k |(B)_{kj}| \sum_{l=1}^n |(C)_{il}| \int_0^{\infty} e^{-\sigma t} dt \Rightarrow \\ \Rightarrow |H_{ij}(i\omega)| &\leq \frac{1}{\sigma} \max_k |(TB)_{kj}| \sum_{l=1}^n |(KCT^{-1})_{il}| \quad \forall \omega \in R^1. \end{aligned}$$

Составим матрицу

$$Q = \left\{ \max_k |(TB)_{kj}| \sum_{l=1}^n |(KCT^{-1})_{il}| \right\}, \quad i = \overline{1, r}, j = \overline{1, m}.$$

Согласно оценке, приведенной в [4],

$$|\lambda_{\max}(H^T(i\omega)H(i\omega))| \leq \min(|H^T(i\omega)H(i\omega)|_1, |H^T(i\omega)H(i\omega)|_{\infty}),$$

$$|H^T(i\omega)H(i\omega)|_1 \leq |H^T(i\omega)|_1 |H(i\omega)|_1 \leq |Q^T|_1 |Q|_1 = \frac{1}{\sigma^2} |Q|_{\infty} |Q|_1,$$

$$\begin{aligned} |H^T(i\omega)H(i\omega)|_{\infty} &\leq |H^T(i\omega)|_{\infty} |H(i\omega)|_{\infty} \leq |Q^T|_{\infty} |Q|_{\infty} = \\ &= \frac{1}{\sigma^2} |Q|_1 |Q|_{\infty}, \end{aligned}$$

$$|\lambda_{\max}(H^T(i\omega)H(i\omega))| \leq \frac{1}{\sigma^2} |Q|_1 |Q|_{\infty}.$$

Следовательно, $\left\| [1 - P_n(s)K(s)]^{-1} P_n(s) \right\|_{\infty} \leq \frac{1}{\sigma} \sqrt{|Q|_1 |Q|_{\infty}}$ и

$$\alpha \leq \frac{1}{\sigma} \sqrt{|Q|_1 |Q|_{\infty}} \|W_d(s)\|_{\infty}. \quad (7)$$

Пусть система $\dot{w} = A_w w + B_w u_w$, $y_w = C_w w$, есть ss-форма для системы $u_w = W_d(s) y_w$. Пусть матрицы A_w , B_w , C_w размерности $\psi \times \psi$, $\psi \times \xi$, $\varsigma \times \psi$ соответственно.

Система с передаточной матрицей

$$H(s) = K [1 - P_n(s) K(s)]^{-1} P_n(s)$$

в ss-форме имеет вид

$$\dot{x} = T(A + BKC)T^{-1}x + TBu, \quad y = KCT^{-1}x.$$

Тогда для объекта с передаточной матрицей

$$M(s) = W_d(s) K [1 - P_n(s) K]^{-1} P_n(s)$$

ss-форма имеет вид

$$\begin{aligned} \begin{pmatrix} \dot{x} \\ \dot{w} \end{pmatrix} &= \begin{pmatrix} T(A + BKC)T^{-1} & 0 \\ B_w KCT^{-1} & A_w \end{pmatrix} \begin{pmatrix} x \\ w \end{pmatrix} + \begin{pmatrix} TB \\ 0 \end{pmatrix}, \\ y_w &= (0 \quad C_w) \begin{pmatrix} x \\ w \end{pmatrix}. \end{aligned}$$

При дополнительном предположении, что $\sigma(A_w) > \|B_w KCT^{-1}\|_\infty$, аналогично рассуждениям, изложенным выше, получаем результат

$$\begin{aligned} \alpha &\leq \frac{1}{\sigma} \sqrt{|Q|_1 |Q|_\infty}, \\ Q &= \left\{ \max_k \left| \begin{pmatrix} TD \\ 0 \end{pmatrix}_{kj} \right| \sum_{l=1}^n |(0 \quad C_w)_{il}| \right\}, \quad i = \overline{1, \zeta}, \quad j = \overline{1, m}, \\ \sigma &= \sigma \begin{pmatrix} T(A + BKC)T^{-1} & 0 \\ B_w KCT^{-1} & A_w \end{pmatrix}. \end{aligned}$$

Заметим, что достаточным условием для $\sigma(A_w) > 0$ является расположение собственных чисел λ_i знаменателей $W_d(s)$ в области $\{\lambda : \operatorname{Re} \lambda < 0, -\operatorname{Re} \lambda > |\operatorname{Im} \lambda|\}$.

Для нахождения границ робастной устойчивости может быть использован следующий алгоритм.

1. Построить начальные значения матриц T и K , сверхстабилизирующую систему (1).

2. Используя полученные начальные данные и элементы матрицы T в качестве переменных, минимизировать

$$\alpha = \frac{1}{\sigma} \sqrt{|Q|_1 |Q|_\infty} \|W_d(s)\|_\infty \rightarrow \min$$

$$Q = \left\{ \max_k |(TB)_{kj}| \left| \sum_{l=1}^n |(KCT^{-1})_{il}| \right. \right\}, \quad i = \overline{1, r}, \quad j = \overline{1, m},$$

$$-A_{i,i}^* - \sum_{j \neq i} |A_{i,j}^*| \geq 0, \quad i = \overline{1, n},$$

где $A^* = T(A + BKC)T^{-1}$, либо

$$\alpha = \frac{1}{\sigma} \sqrt{|Q|_1 |Q|_\infty} \rightarrow \min,$$

$$Q = \left\{ \max_k \left| \begin{pmatrix} TD \\ 0 \end{pmatrix}_{kj} \right| \left| \sum_{l=1}^n |(0 \ C_w)_{il}| \right. \right\}, \quad i = \overline{1, \varsigma}, \quad j = \overline{1, m},$$

$$\sigma = \sigma \begin{pmatrix} T(A + BKC)T^{-1} & 0 \\ B_w KCT^{-1} & A_w \end{pmatrix} > 0.$$

3. Если требуемый результат не достигнут, получить новые начальные значения матрицы T и повторить п. 2.

4. Заключение. Рассмотрены некоторые методы сверхустойчивости линейных стационарных систем. Предложены алгоритмы построения обратной связи по выходу для линейного стационарного объекта, обеспечивающей сверхустойчивость замкнутой системы, с дополнительными робастными свойствами. Все алгоритмы легко реализуются в среде MATLAB.

Литература

1. Поляк Б.Т., Щербаков П.С. Робастная устойчивость и управление. М.: Наука, 2002. 303 с.
2. Утешев А.Ю. Высшая алгебра, Раздел II. СПб.: Золотое Сечение, 2007. 162 с.
3. www.matlab.exponenta.ru/optimrobast/book2/index.php
4. Корн Г., Корн Т. Справочник по математике для научных сотрудников и инженеров. М.: Наука, 1974. 831 с.

Дурновцева С.А.

Санкт-Петербургский государственный университет

Моделирование исходных сейсмических колебаний грунта для формирования проектных основ

Рекомендовано к публикации профессором Веремеем Е.И.

Введение. В понятие «проектные основы» входят исходные данные и постулируемые события для проектирования объекта использования атомной энергии (ОИАЭ), изготовления его оборудования, систем и устройств, их монтажа и наладки, строительства, обеспечения его нормального функционирования на протяжении установленного срока эксплуатации [1]. Определение сейсмических колебаний грунта является неотъемлемой частью анализа, антисейсмического проектирования и тестирования на вибрационную надежность как новых, так и действующих энергетических объектов. Задание динамических нагрузок на объект в виде акселерограмм (записей во времени однокомпонентного процесса изменения ускорения) необходимо при расчете ответных перемещений, ускорений, опорных реакций, внутренних усилий и напряжений в трубопроводах объекта.

В статье приведен математический метод моделирования акселерограмм, удовлетворяющих исходным данным. Реализация и тестирование предложенного метода выполнены в среде MATLAB.

Содержательная постановка задачи. Введем понятие спектра ответа.

Спектр ответа (СО) – совокупность значений (график) абсолютных максимальных ответных ускорений одномассовой колебательной системы (сейсмоосциллятора) при воздействии акселерограммы, определенных в зависимости от собственной частоты и уровней демпфирования осциллятора.

Идеальным представлением сейсмических характеристик площадки строительства ОИАЭ явился бы достаточно полный ансамбль записей интенсивных сейсмических движений грунта, зарегистрированных непосредственно в данном месте. По ним мог бы быть вычислен спектр ответа. При таком подходе автоматически обеспечивается соответствие («совместимость») спектра и акселерограммы.

Но сильные землетрясения происходят редко, и получить такой ансамбль записей в обозримые сроки не представляется возможным. Поэтому обычно поступают наоборот: в качестве исходной сейсмической характеристики площадки задают СО, а уже по нему генерируют акселерограмму. При этом СО определяют путем статистической обработки реальных данных сильных землетрясений в интересующей зоне, а также данных для других регионов, имеющих сходные сейсмо-тектонические, геологические и грунтовые условия. После статистической обработки итоговый спектр обычно сглаживают, подвергают процедуре расширения пиков и уже в таком виде используют в качестве расчетного.

Необходимо отметить, что построение акселерограммы, соответствующей заданному СО, не является однозначной задачей [2], так как может быть построено бесконечное множество временных зависимостей (различных как по амплитуде, так и по продолжительности воздействия), имеющих нужный СО.

Математическая постановка задачи. Уравнение колебаний осциллятора имеет вид:

$$\ddot{u} + 2\tilde{\omega}\zeta\dot{u} + \tilde{\omega}^2u = -a(t), \quad (1)$$

где u – относительное перемещение, ζ – коэффициент демпфирования, $\tilde{\omega}$ – частота осциллятора. Сейсмическое воздействие описывается функцией $a(t)$. С математической точки зрения СО – это график, отображающий зависимость величины

$$S_a = \max_t |u(t)| \quad (2)$$

от собственной частоты осциллятора при заданном ζ . Задача состоит в синтезе функции $a(t)$. Коэффициент демпфирования, СО и диапазон частот осциллятора считаются известными. Время t распространения сейсмического воздействия не задано.

Метод синтеза акселерограмм. Исходя из аналоговых записей реальных акселерограмм, логично предположить, что сейсмическое воздействие можно представить в виде суперпозиции синусоид с различными амплитудами и частотами:

$$a(t) = F(t) \sum A_i \sin(\omega_i t + \varphi_i), \quad (3)$$

где $F(t)$ – огибающая колебаний, а фазовые углы φ_i есть равномерно распределенные в интервале от 0 до 2π случайные величины. Длина акселерограммы принимается равной 15 секундам. Частоты ω_i совпадают с диапазоном частот $\tilde{\omega}$ на графике СО. Начальные амплитуды A_i выбраны как $A_i = S_{target}(\tilde{\omega}_i)/n$, где S_{target} – целевой СО.

После задания нулевого приближения к акселерограмме в виде (3) выполняется интегрирование уравнений (1) для всех $\tilde{\omega}$ (используемый численный метод интегрирования приведен в источнике [3, с. 153–156]) и построение СО. Вычисляется вектор расхождения d между целевым СО (S_{target}) и вычисленным на текущей итерации ($S_{current}$) в каждой «контрольной точке» $\tilde{\omega}$ (для сокращения количества вычислений на каждом шаге целесообразно рассматривать частоты осциллятора $\tilde{\omega}$ «через одну»):

$$d = \frac{S_{target}}{S_{current}}. \quad (4)$$

Амплитуды корректируются по следующему закону:

$$A := A\sqrt{d}. \quad (5)$$

Результаты. На рис. 1, 2 приведены графики, иллюстрирующие результаты работы алгоритма.

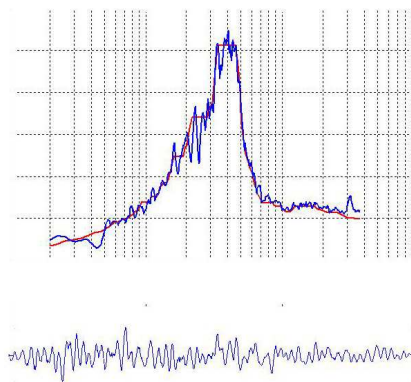


Рис. 1. Исходный СО № 1, синтезированная по нему акселерограмма (нижний график) и соответствующий ей СО

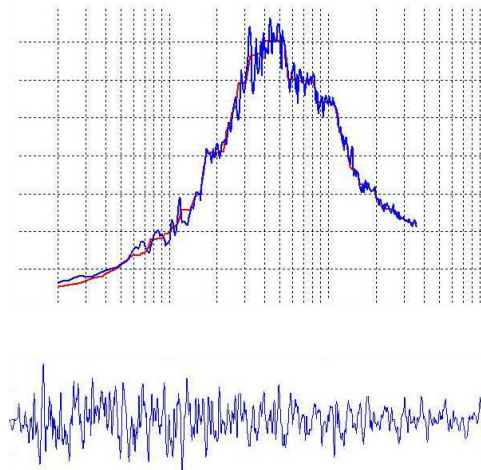


Рис. 2. Исходный СО № 2, синтезированная по нему акселерограмма (нижний график) и соответствующий ей СО

Ясно, что указанный метод позволяет синтезировать искусственную запись колебания грунта, предоставляя, таким образом, совместимые друг с другом спектр ответа и акселерограмму для проектных основ ОИАЭ.

Литература

1. Руководства по ядерной и радиационной безопасности: «Определение исходных сейсмических колебаний грунта для проектных основ» (РБ-006-98). Нормативный документ. М.: НТЦ ЯРБ 2000. 76 с.
2. Бирбраер А.Н. Расчет конструкций на сейсмостойкость. СПб.: Наука, 1998. 255 с.
3. Gupta A.K. Response spectrum method in seismic analysis and design of structures / Ed. by Blackwell Scientific Publications, 1990. P. 364.

Журавлев М.М.

Санкт-Петербургский государственный университет

Применение принципа минимальной длины описания к задаче идентификации механических повреждений по изображению

Рекомендовано к публикации профессором Андриановым С.Н.

Введение. За последние 10 лет в России было утрачено около 2000 памятников культуры, на грани разрушения находятся 2663 объекта [1]. Для планирования работ по сохранению памятников культурного наследия необходимо проводить мониторинг как их текущего состояния, так и факторов, влияющих на скорость разрушения.

В работе Гришкина В.М. [2] предложена схема системы мониторинга состояния объектов культурного наследия, базирующаяся на методах цифровой обработки изображений, в частности, на распознавании образов. Информация об областях деструкции может быть получена при обработке цветных изображений объектов. При этом механические области деструкции (трещины) проявляются на изображении в виде относительно темных участков и могут быть выявлены при анализе яркостной составляющей изображения. Области биологической деструкции связаны с цветовыми характеристиками в зоне биологического поражения.

В работе [3] показано, что разрабатываемый метод идентификации механических повреждений по изображению позволяет сократить временные затраты специалистов, проводящих мониторинг. Точность и стоимость применения метода удовлетворительные.

Схема метода идентификации трещин на изображении приведена на рис. 1. Исходный снимок предварительно очищается от шумов, являющихся следствием наличия в воздухе между камерой и поверхностью пыли, фильтром 2-D median filtering [4].

Информация о цвете пикселя снимка сохраняется с использованием координат цвета в пространстве RGB (Red, Green, Blue). В цветовом пространстве LAB цвет определяется однозначно, поэтому цветовые координаты пикселей снимка должны быть преобразованы в LAB. Разместив около объекта шаблон серого цвета, возможно

учесть условия освещения. Цвет шаблона на снимке будет отличаться от известного, а цвета интересующей поверхности фрагмента также будут отличаться от их истинного значения. Алгоритм цветовой коррекции определяет среднее значение цвета шаблона на исходном снимке и на основании его истинного значения вычисляет коэффициенты k_r, k_g, k_b коррекции по каналам цвета R, G, B. Коррекция цвета исходного изображения производится простым умножением значений цветовых составляющих каждого пикселя на соответствующий коэффициент.



Рис. 1. Схема идентификации трещин на снимке объекта

При выполнении некоторых ограничений на освещение поверхности объекта, которые будут формализованы впоследствии, можно считать, что образ трещины на снимке имеет меньшую яркость, чем окружающая стена. Яркость пикселя является отдельной координатой цветового пространства LAB. Для упрощения вычисления меры близости пикселей исходный снимок преобразуется в данное цветовое пространство.

Кластерный анализ. Одним из методов цифровой обработки изображений является применение кластерного анализа. В процессе кластеризации осуществляется разбиение заданного множества объектов на непересекающиеся подмножества, называемые кластерами, таким образом, чтобы каждый кластер состоял из схожих объектов, а объекты разных кластеров существенно отличались [5].

Анализ проводится в предположении о наличии одной выраженной трещины на анализируемом изображении. Выделение производится вручную. При наличии нескольких трещин на поверхности необходимо применять алгоритм несколько раз к разным частям изображения. Разработка алгоритма автоматической обработки изображения, содержащего несколько трещин – одно из направлений развития метода.

С целью идентификации трещины снимок разбивается на кластеры по яркости пикселей. Связное множество пикселей наибольшей площади, принадлежащее кластеру наименьшей яркости – оценка трещины. Получив множество точек образа трещины на снимке, зная условия съемки, можно получить числовые характеристики трещины и оценку погрешности результата. В качестве математической модели съемки рассматривается перспективная проекция [6]. Оценка относительной и абсолютной погрешности метода дана в работе [7]. Важным достоинством метода является постоянная относительная погрешность результатов вычислений.

В разрабатываемой программе-прототипе используется алгоритм кластеризации « k -средних». Алгоритм разбивает множество на заранее заданное число кластеров. В результате вычислительных экспериментов получен эмпирический критерий необходимости повторного проведения расчетов с увеличенным числом кластеров. Критерий основан на сравнении площадей областей (количество пикселей в области изображения), идентифицированных как трещины. Исходные данные, позволившие вывести данный критерий опубликованы в работе [7].

Для мониторинга состояния объекта необходимо сопоставлять информацию, полученную путем обработки снимков, сделанных в разные моменты времени. Каждый снимок порождает набор описаний трещин (множество пикселей, вычисленные параметры). В настоящее время пользователь осуществляет анализ динамики вручную.

Алгоритм « k -средних» не использует информацию, полученную при анализе предыдущих снимков, для проведения кластеризации. Таким образом, задача автоматизированного сопоставления результатов для проведения анализа динамики максимальной ширины раскрытия трещины должна решаться отдельно. Алгоритм кластеризации, использующий ранее полученную информацию, сможет решить одновременно и задачу сопоставления множеств трещин на снимках.

Анализ литературы по распознаванию образов показал, что при наличии априорной информации об анализируемом множестве, может быть использован подход, основанный на принципе минимальной длины описания [5, 8, 9].

Прототип программы, реализующей метод идентификации механических повреждений по изображению, разработан с использованием системы MATLAB. Это позволило сократить время разработки,

увеличив время расчетов, за счет использования реализованного в системе автоматического управления памятью. В MATLAB имеются встроенные реализации необходимых для решения поставленной задачи методов, например, 2-D median filtering. Данный метод используется для предварительной очистки подаваемого на вход основному алгоритму изображения. Изображение очищается от случайных помех, вызванных неравномерным распределением пыли в воздухе между объектом и камерой.

В работе [7] анализируются результаты проведенных вычислительных экспериментов. Эксперименты, в частности, показали следующее:

1. При увеличении числа кластеров свыше некоторого предела оценка максимальной ширины раскрытия трещины практически перестаёт изменяться.
2. Определить ординату максимального раскрытия трещины методом с удовлетворительной точностью не позволяет, так как отнесение или не отнесение одной граничной точки к кластеру может значительно изменить значение вычисленной ординаты, особенно ярко это выражается в случае почти вертикальных границ трещины.

Для применения принципа минимальной длины описания к решению задачи идентификации механических повреждений по изображению планируется реализовать алгоритм кластеризации, подобный алгоритмам в [5, 9]. Критерий кластеризации будет учитывать информацию о предыдущем состоянии поверхности объекта и модель распространения трещины. Представляется, что это позволит решить проблему, упомянутую в пункте 2, а именно, повысить точность определения ординаты максимальной ширины раскрытия трещины.

Заключение. Проведен анализ текущего состояния метода идентификации механических повреждений по изображению. Сформулированы задачи, которые необходимо решить для дальнейшего развития метода и повышения удобства использования программного обеспечения, реализующего данный метод. Анализ литературы показал, что применение принципа минимальной длины описания (minimum description length) к задаче кластеризации изображения поверхности с механическими повреждениями (трещинами) позволит решить поставленные задачи.

Литература

1. Россия теряет по 200 памятников в год
http://news.km.ru/rossiya_teryaet_po_200_pamyatnik
2. Гришкин В.М. Компьютерная система мониторинга состояния объектов культурного наследия // Вестник СПбГУ. Сер. 10, 2009. Вып. 3. С. 182–189.
3. Журавлев М.М. Идентификация механических повреждений по изображению // Процессы управления и устойчивость: Труды 40-й международной научной конференции аспирантов и студентов / Под ред. Н.В. Смирнова, Г.Ш. Тамасяна. СПб.: Издат. Дом С.-Петербур. гос. ун-та, 2010. С. 424–428.
4. MATLAB documentation
<http://www.mathworks.com/help/toolbox/images/ref/medfilt2.html>
5. Navarro J. Daniel An application of minimum description length clustering to partitioning learning curves // Proceedings of the 2005 IEEE International Symposium on Information Theory, 2005. P. 587–591.
6. Luong Q.T., Fuuergas O.D. The fundamental matrix: Theory, algorithms, and stability analysis // Intern. J. of Computer Vision, 1996. Vol. 17, № 1. P. 43–75.
7. Журавлев М.М. Выбор числа кластеров в задаче идентификации механических повреждений по изображению // Процессы управления и устойчивость: Труды 41-й международной научной конференции аспирантов и студентов / Под ред. Н.В. Смирнова, Г.Ш. Тамасяна. СПб.: Издат. Дом С.-Петербур. гос. ун-та, 2010. С. 412–417.
8. Потапов А.С. Распознавание образов и машинное восприятие: Общий подход на основе принципа минимальной длины описания. СПб.: Политехника, 2007. 548 с.
9. Kontkanen P., Myllymaki P., Buntine W., Rissanen J., Tirri H. An MDL framework for data clustering // Grunwald P., Myung M. Pitt A. (Eds.) Advances in Minimum Description Length: Theory and Applications. Cambridge, MA: MITPress, 2005. P. 323–354.

Иванова О.А., Пешехонов К.А.
 Санкт-Петербургский государственный университет

Программная реализация метода вычисления эллиптических функций

В работе реализуется алгоритм построения аппроксимации Паде для вычисления эллиптических функций на компьютере и приводятся численные результаты.

Напомним, что аппроксимация Паде – это рациональная функция вида [1]:

$$[L/M] = \frac{d_0 + d_1z + d_2z^2 + \dots + d_Lz^L}{b_0 + b_1z + b_2z^2 + \dots + b_Mz^M}, \quad (1)$$

где коэффициенты числителя d_i , $i = 0, \dots, L$, и знаменателя b_j , $j = 0, \dots, M$, $b_0 = 1$, определяются через коэффициенты разложения в ряд исходной функции, т.е. через коэффициенты Тейлора.

Используя ранее выведенные формулы в работах [2, 3], выпишем коэффициенты Тейлора

$$a_i = \left. \frac{\wp^{(i)}(z)}{i!} \right|_{z=z_0}, \quad i = 0, 1, \dots, \quad (2)$$

где $\wp^{(i)}(z)$ – функция Вейерштрасса. Воспользовавшись тем, что функция Вейерштрасса удовлетворяет дифференциальному уравнению, распишем производные $\wp(z)$ -функции в точке z_0 [2-6]:

$$\begin{aligned} \wp^{(1)}(z)|_{z=z_0} &= \sqrt{4\wp_0^3 - g_2\wp_0 - g_3}, \\ \wp^{(2)}(z)|_{z=z_0} &= 6\wp_0^2 - \frac{1}{2}g_2, \\ \wp^{(3)}(z)|_{z=z_0} &= 12\wp_0\sqrt{4\wp_0^3 - g_2\wp_0 - g_3}, \\ \dots\dots\dots & \dots\dots\dots \end{aligned} \quad (3)$$

В системе компьютерной алгебры Maple 11 коэффициенты Тейлора сначала находятся в символьном виде, а затем вычисляются непосредственно в требуемой точке в среде реализации (C#).

```

> a[1] := (4*pz^3-g2*pz-g3)^(1/2);
> a[2] := 3*pz^2-1/4*g2;
> a[3] := 2*pz*(4*pz^3-g2*pz-g3)^(1/2);
> a[4] := 5*pz^3-3/4*g2*pz-1/2*g3;
> a[5] := -3/20*(4*pz^3-g2*pz-g3)^(1/2)*(-20*pz^2+g2);
> a[6] := 7*pz^4-7/5*pz^2*g2+1/80*g2^2-pz*g3;
> a[7] := -1/35*(4*pz^3-g2*pz-g3)^(1/2)*(-140*pz^3+14*g2*
    pz+5*g3);
> a[8] := 9*pz^5-9/4*pz^3*g2+3/40*g2^2*pz-45/28*g3*pz^2+
    33/560*g3*g2;

```

Функция вида (1) имеет $L+1$ коэффициентов в числителе и $M+1$ коэффициентов в знаменателе. Строим диагональные аппроксимации Паде, т.е. $L = M$. Задаем M и сразу же проверяем существование в задаваемой точке аппроксимации Паде – строим определитель Ганкеля [1–4]:

$$Q^{[M/M]}(0) = \begin{vmatrix} a_1 & a_2 & \dots & a_M \\ a_2 & a_3 & \dots & a_{M+1} \\ \vdots & \vdots & \ddots & \vdots \\ a_{M-1} & a_M & \dots & a_{2M-2} \\ a_M & a_{M+1} & \dots & a_{2M-1} \end{vmatrix}, \quad (4)$$

где $a_j, j = 0, \dots, 2M+1$, определяются по формулам (2)–(3) из начальных данных.

Проверяем, не равен ли нулю построенный определитель Ганкеля (4). Если он не равен нулю, то приступаем непосредственно к вычислению аппроксимации Паде. Если же он равен нулю, то речь может идти о том, что аппроксимации Паде может не существовать в заданной точке, т.е. заданная точка может быть особой точкой. Чтобы убедиться, что мы попали в особую точку, надо попробовать пересчитать определитель Ганкеля при других L и M . Если же это ничего не даст, и вновь построенный определитель Ганкеля также будет равен нулю, то тогда можно попытаться обойти эту точку, задав δ , и посчитать определитель Ганкеля в точке $z_0 + \delta$, либо в точке $z_0 - \delta$.

Итак, если определитель не равен нулю, то переходим к постро-

ению знаменателя аппроксимации Паде

$$Q^{[M/M]}(z) \Big|_{z, z=z_0} = \begin{vmatrix} a_1 & a_2 & \dots & a_{M+1} \\ a_2 & a_3 & \dots & a_{M+2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{M-1} & a_M & \dots & a_{2M-1} \\ a_M & a_{M+1} & \dots & a_{2M} \\ z^M & z^{M-1} & \dots & 1 \end{vmatrix}_{z, z=z_0},$$

где $a_j, j = 0, \dots, 2M + 1$, определяются также по формулам (2)–(3) и начальным данным.

Теперь распишем числитель аппроксимации Паде

$$P^{[M/M]}(z) \Big|_{z, z=z_0} = \begin{vmatrix} a_1 & a_2 & \dots & a_{M+1} \\ a_2 & a_3 & \dots & a_{M+2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{M-1} & a_M & \dots & a_{2M-1} \\ a_M & a_{M+1} & \dots & a_{2M} \\ a_0 z^M & \sum_{i=0}^1 a_i z^{M+i-1} & \dots & \sum_{i=0}^M a_i z^i \end{vmatrix}_{z, z=z_0},$$

где $a_j, j = 0, \dots, 2M + 1$, также определяются по формулам (2)–(3) и начальным данным.

Теперь считаем значение аппроксимации Паде в точке z :

$$[M/M]_{z, z=z_0} = \frac{P^{[M/M]}(z) \Big|_{z, z=z_0}}{Q^{[M/M]}(z) \Big|_{z, z=z_0}}. \quad (5)$$

Полученное выражение (5) дает приближенное значение исходной $\varphi(z)$ -функции в точке z с начальными данными.

Вышеописанный алгоритм реализован на платформе .NET с использованием языка C# и компонентной модели. Программная реализация состоит из двух компонентов: ComplexMath, реализующего класс для работы с комплексными числами и PFunctionAppr, реализующего метод аппроксимации. Эти компоненты реализованы в виде библиотек классов ComplexMath.dll и PFunctionAppr.dll соответственно. Работа метода аппроксимации разделена на две основные части.

Первая часть, функция `Init`, выполняемая один раз для каждого набора начальных данных (начальной точки z_0 и вещественного и комплексного полупериодов ω_1 и ω_3), отвечает за инициализацию начальных данных метода, в ней производятся все объемные по времени расчеты, не требующие повторения при изменении точки, в которой вычисляется значение функции, а именно, вычисление значения функции в начальной точке, инвариантов g_2 и g_3 и коэффициентов ряда Тейлора a_i . Здесь же проверяется существование аппроксимации Паде, то есть строится определитель Ганкеля и проверяется, не равен ли он нулю. В случае, если он равен нулю, то программа предлагает использовать другое M . Если же ситуация не меняется, и определитель Ганкеля по-прежнему равен нулю, то расчет останавливается и пользователю возвращается соответствующая ошибка, сообщающая о необходимости изменения начальных данных. Функция `Init` возвращает объект класса `ApproxData`, а на вход принимает начальную точку z_0 как аргумент `Complex z0`, полупериоды ω_1 и ω_3 как аргументы `Complex w1` и `Complex w3` и погрешность как аргумент `R`, с которой вычисляются $\wp(z_0)$, g_2 и g_3 .

Заголовок функции `Init`:

```
public static ApproxData Init(Complex z0, Complex w1
                             Complex w3, double R)
```

Класс `ApproxData` хранит все рассчитанные данные и позволяет их использовать во второй части реализации – функции `ApproxP`. Эта функция отвечает за реализацию самого метода: построение матриц числителя и знаменателя аппроксимации Паде P и Q , их деление, вычисление определителя и возврат результата. Для операций с матрицами служит класс `Matrix`, реализующий стандартные алгоритмы работы с матрицами, которые составлены из элементов, представленных объектами класса `Complex`. На вход функция `ApproxP` принимает точку z как аргумент `Complex z`, в которой производится вычисление значения функции Вейерштрасса [5, 6], и объект класса `ApproxData`, хранящий начальные данные для алгоритма как аргумент `ApproxData approxData`.

Заголовок функции `ApproxP`:

```
public static Complex ApproxP(Complex z,
                              ApproxData approxData)
```

Внутренняя реализация этих библиотек такова: библиотека классов ComplexMath.dll содержит класс Complex, реализующий работу с комплексными числами, и класс Matrix, реализующий работу с матрицами, состоящими из элементов, представленных объектами класса Complex. Библиотека RFunctionAppr содержит реализации классов Apprrox, реализующего аппроксимацию, и Init, к которому обращается класс Apprrox для вычисления начальных значений.

В качестве исходных данных для проверки работы реализации алгоритма были взяты начальные данные $\omega_1 = 1$, $\omega_3 = 1, 2i$ [7]. В качестве начальной точки была выбрана $z_0 = 1, 1$. Значение функции Вейерштрасса вычислено в точке $z = 0, 8$, $\wp(0, 8) = 1, 924285443$, погрешность вычисления составила 10^{-3} .

Литература

1. Бейкер Дж., Грейвс-Моррис П. Аппроксимация Паде. М.: Мир, 1986. 502 с.
2. Иванов А.А., Иванова О.А. Построение аппроксимаций Паде для решения уравнения Дуффинга // Процессы управления и устойчивость: Труды 38-й научной конференции аспирантов и студентов / Под ред. А.В. Платонова, Н.В. Смирнова. СПб.: Изд-во СПбГУ, 2007. С. 28–32.
3. Иванов А.А., Иванова О.А. Новый алгоритм вычисления \wp -функции Вейерштрасса // Процессы управления и устойчивость: Труды 39-й научной конференции аспирантов и студентов / Под ред. Н.В. Смирнова, Г.Ш. Тамасяна. СПб.: Изд-во СПбГУ, 2008. С. 33–38.
4. Вишневецкий В.Э., Иванова О.А. Функционально-аналитические представления решений в нелинейных задачах теории управления. Учебное пособие. СПб.: Изд-во СПбГУ, 2003. 112 с.
5. Маркушевич А.И. Теория аналитических функций. Т. 2. М.: Наука, 1968. 624 с.
6. Справочник по специальным функциям с формулами, графиками и таблицами / Под ред. М. Абрамовица, И. Стиган. М.: Наука, 1979. 832 с.

Лобов А.Л.

Санкт-Петербургский государственный университет

Компьютерные алгоритмы идентификации моделей морских судов

Рекомендовано к публикации профессором Веремеем Е.И.

1. Введение. Идентификация представляет собой определение наилучшей в определенном смысле модели из заданного класса на основании экспериментальных данных. В качестве данных для анализа используются записи входов и выходов системы или некоторые их характеристики. Для линейных стационарных математических моделей идентификация главным образом представляет собой процесс определения передаточной функции на основе экспериментальных данных.

Вычислительно дешевым и простым является подход к идентификации в частотной области. Этот подход подразумевает оценку частотной характеристики системы по известным оценкам спектров входного и выходного сигналов [1]. Поскольку методы оценки спектральной плотности мощности хорошо проработаны и основываются преимущественно на вычислении преобразования Фурье [2], частотный подход позволяет реализовывать эффективные процедуры идентификации, в том числе в реальном времени. Если структура искомой модели известна, и модель может быть параметризована относительно некоторых компонентов передаточной функции, то в таком случае задача идентификации превращается в задачу подгонки кривой [3].

Данная работа посвящена нахождению наилучшего способа оценки параметров линейной модели морского подвижного объекта с использованием идентификации в частотной области.

2. Структура модели и процесса идентификации. Модель морского подвижного объекта представляет собой нелинейную модель, замкнутую обратной связью. Схема модели представлена на рис. 1.

В целях идентификации в канал управления аддитивно подается специальным образом сформированный сигнал, обозначенный на рисунке как u_0 . Сигнал представляет собой белый шум, пропущенный

через формирующий фильтр, и является дробно-рациональной моделью спектра морского волнения. Такой подход к построению возбуждающего сигнала позволяет иметь три основных параметра: среднюю частоту, дисперсию и размытость, управляя которыми можно выбрать сигнал, дающий наилучшую в своем классе оценку частотной характеристики.

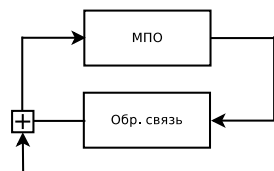


Рис. 1. Схема модели

Структура искомой передаточной функции линейной системы известна и может быть параметризована относительно шести компонентов матриц линейных уравнений морского объекта: $H(z, \gamma)$, где γ – вектор параметров. В связи с этим суть процесса идентификации сводится к нахождению оценки $\hat{H}(j\omega)$ частотной характеристики системы $H(j\omega)$ на некотором наборе частот $\omega_1, \dots, \omega_n$ и последующей минимизации функционала

$$J(\gamma) = \sum_{k=1}^n \left| \hat{H}(j\omega_k) - H(j\omega_k, \gamma) \right|^2. \quad (1)$$

3. Оценки частотной характеристики обычно основываются на использовании соотношения вход-выход в частотной области:

$$Y(\omega) = H(j\omega)X(\omega), \quad (2)$$

где $Y(\omega)$ и $X(\omega)$ – это, соответственно, преобразования Фурье выхода и входа системы, а $H(j\omega)$ – это комплексная частотная характеристика системы.

Поскольку на практике в качестве спектральной характеристики сигнала рассматривается спектральная плотность мощности (которая для сигналов, интегрируемых с квадратом, равна квадрату модуля преобразования Фурье), обычно используется следующий вид соотношения (2): $S_{yy}(\omega) = |H(j\omega)|^2 S_{xx}(\omega)$. Здесь S_{xx} и S_{yy} – спектральные плотности мощности входа и выхода системы, соответственно.

Из этой формулы следует наиболее очевидная оценка

$$|H(j\omega)|^2 = \frac{S_{yy}(\omega)}{S_{xx}(\omega)}, \quad (3)$$

которая именуется эмпирической оценкой передаточной функции (Empirical Transfer Function Estimate, ETFE). Было показано [4], что в случае случайного возмущения оценка является асимптотически несмещенной, но ее дисперсия не убывает при возрастании числа отсчетов, а определяется соотношением сигнал-шум на частоте интереса. Таким образом, в случае случайных сигналов эмпирическая оценка передаточной функции является плохой в статистическом смысле оценкой.

Тем не менее, реальная частотная характеристика системы обычно является гладкой функцией частоты [4]. Это дает возможность использовать осреднение для получения более качественных оценок. Осреднение выполняется по набору записей входов-выходов, которые могут быть получены: а) либо в результате соответствующего количества независимых экспериментов, в которых случайные возмущения являются стационарными; б) либо в результате разбиения на сегменты результата одного эксперимента, в котором возмущения могут быть представлены белым шумом, пропущенным через фильтр [1].

Наиболее очевидный путь к осреднению эмпирической оценки передаточной функции лежит в прямом осреднении вычисленных ETFE по всем блокам данных:

$$\widehat{H}(j\omega) = \sum_{k=1}^N \frac{S_{yy}^{[k]}(\omega)}{S_{xx}^{[k]}(\omega)}. \quad (4)$$

Однако, было показано, что прямое вычисление ETFE по формуле (4) имеет бесконечную дисперсию [1]. В первую очередь, это вызвано нелинейностью деления, вследствие чего возникают ошибки наложения. Это означает, что необходимо выполнять осреднение до выполнения деления. Если брать во внимание приведенный выше случай б), то мы имеем не что иное, как частное периодограмм Уэлча [2], вычисленных для входного и выходного сигналов. Такой вариант осреднения ETFE был также предложен Льюнгом [4]. Оценка имеет вид

$$\widehat{H}(j\omega) = \frac{\sum_{k=1}^N S_{yy}^{[k]}(\omega)}{\sum_{k=1}^N S_{xx}^{[k]}(\omega)}. \quad (5)$$

Тем не менее, качество оценки (5) страдает за счет того, что фаза входа является случайной величиной с нулевым средним значени-

ем [1]. Для того, чтобы повысить качество оценки, обе части соотношения (2) обычно умножают на величину, сопряженную спектру входного сигнала. Это позволяет получить векторы с фиксированной фазой. В результате имеем следующую оценку частотной характеристики

$$\widehat{H}(j\omega) = \frac{\sum_{k=1}^N \widehat{Y}^{[k]}(\omega) \overline{\widehat{X}^{[k]}(\omega)}}{\sum_{k=1}^N |\widehat{X}^{[k]}(\omega)|^2}, \quad (6)$$

где $\widehat{Y}^{[k]}$ и $\widehat{X}^{[k]}$ – преобразования Фурье соответствующих сегментов выхода и входа, а верхнее подчеркивание означает комплексное сопряжение. Эта оценка представляет собой дискретный аналог уравнения Винера – Хопфа и также упоминается под названием оценки H_1 [1].

Оценка (6) является более качественной оценкой частотной характеристики системы, чем рассмотренные ранее (3), (4) и (5). Тем не менее, растекание спектра, присущее дискретному преобразованию Фурье, может оказывать существенное влияние на качество такой оценки. Временные параметрические модели лишены этого недостатка, однако простота частотного подхода и возможность наглядно отслеживать влияние входного сигнала на полученную оценку непосредственно по частотной характеристике системы дают ему полное право быть хорошим практическим инструментом в задачах идентификации. Далее мы увидим, что в нашей задаче частотные методы приводят к приемлемым оценкам частотных характеристик.

4. Реализация процесса идентификации. Для выбора метода оценки частотной характеристики системы рассмотрим приведенные выше оценки для одной реализации входа и выхода системы. На рис. 2 представлены графики несглаженной ЕТФЕ (3), оценки на основе периодограмм Уэлча (5) и оценки H_1 (6) (штриховая линия) в сравнении с реальной частотной характеристикой (сплошная линия).

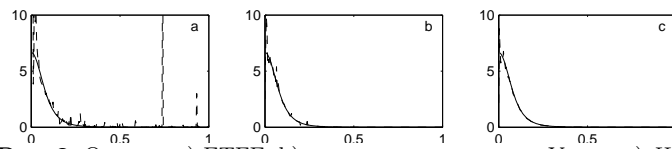


Рис. 2. Оценки: а) ЕТФЕ, б) частное периодограмм Уэлча, в) H_1

На приведенном рис. 2 видно, что несглаженная ЕТФЕ достаточ-

но сильно флуктуирует относительно реальной частотной характеристики. Отношение периодограмм Уэлча входа и выхода дает лучший результат, однако визуально заметны всплески на некоторых частотах, которые могут иметь негативный эффект при дальнейшей минимизации функционала (1). Оценка H_1 имеет меньше всплесков, и все они приходятся на область частот ниже 0,1 рад/с.

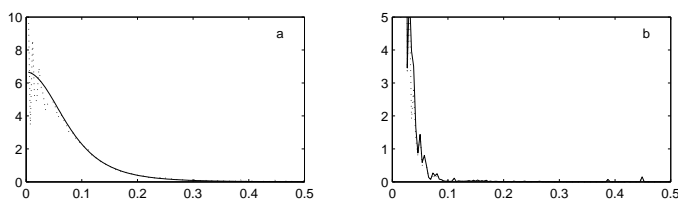


Рис. 3. Оценка H_1 : а) среднее, б) СКО

Статистические свойства H_1 были оценены по 10 реализациям входа-выхода системы. На рис. 3 приведены графики среднего значения (точками) в сравнении с реальной характеристикой системы (сплошная линия) и среднеквадратического отклонения (СКО) в сравнении с квадратом отклонения от реальной характеристики системы. Заметно, что H_1 в среднем дает оценку, мало отличающуюся от реальной характеристики системы, для частот выше 0,1 рад/с. Выберем эту частоту в качестве нижней границы рабочего диапазона частот, на котором будет минимизироваться функционал (1).

В качестве верхней границы выберем частоту, выше которой заметны всплески СКО – 0,4 рад/с. Таким образом, функционал (1) будет минимизирован на интервале частот $[0,1; 0,4]$ рад/с.

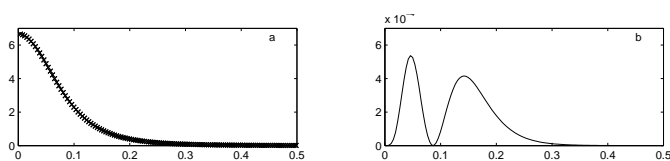


Рис. 4. Результаты: а) идентифицированная кривая, б) квадратичное отклонение от реальной кривой

Минимизация функционала (1) производилась методом Нелдера – Мида с использованием функции `fminsearch` из Optimization Toolbox пакета MATLAB. В качестве оценки частотной характеристики использовалась оценка H_1 . Результат оптимизации можно

видеть на рис. 4. Здесь приведены полученная кривая (крестиками) и реальная частотная характеристика (сплошная линия), а также график отклонения полученной кривой от реальной частотной характеристики.

При этом получены следующие значения коэффициентов линейной модели и их расхождения с коэффициентами линейного приближения исходной модели.

Таблица. Полученные значения коэффициентов

	a_{11}	a_{12}	a_{21}	a_{22}	b_1	b_2
оценки	-0,0937	0,398	0,0122	-0,4119	-0,0241	0,0151
расхож.	11,7%	5,9%	27%	2,9%	31%	0,7%

Сейчас максимальное расхождение составляет 31% при максимальном квадратичном отклонении оценки от реальной характеристики $5,5 \cdot 10^{-4}$. Дальнейшие усилия будут направлены на поиск оптимальных параметров (средняя частота, размытость, дисперсия) входного сигнала с точки зрения минимального отклонения оценок коэффициентов относительно их реальных значений.

Отметим, что в случае более простого вида передаточной функции модели можно аналитически найти градиент функционала (1), что позволит применить более совершенные процедуры оптимизации. В некоторых случаях возможно полностью аналитическое решение задачи (1), что даст весьма эффективную с точки зрения вычислений процедуру идентификации.

Литература

1. Pintelon R., Schoukens J. System identification: a frequency domain approach. New York: IEEE Press, 2001. 648 p.
2. Сергиенко А.Б. Цифровая обработка сигналов. СПб.: Питер, 2002. 608 с.
3. Ljung L. Linear system identification as curve fitting // Springer Lecture Notes on Control and Information, 2003. Vol. 286. P. 203–215.
4. Ljung L. System identification. Theory for the user. Second Edition. Upper Saddle River, New Jersey: Prentice Hall PTR, 1999. 672 p.

Михеев П.А.

Томский государственный университет

Индивидуальное быстродействие абонента беспроводной ЛВС стандарта 802.11

Рекомендовано к публикации профессором Сущенко С.П.

Введение. Рассмотрим беспроводную локальную вычислительную сеть (ЛВС), основывающуюся на стандарте 802.11 [1]. Протокол, обеспечивающий работоспособность беспроводных ЛВС в распределенном режиме доступа (DCF) [1, 2], использует механизм доступа к среде, который называется «множественный доступ с контролем несущей и предотвращением коллизий» (*carrier sense multiple access with collision avoidance, CSMA/CA*). Данный механизм основан на том, что передающая станция проверяет, присутствует ли в среде сигнал несущей, и, прежде чем начать отправку кадра, ожидает освобождения среды передачи данных. Беспроводные станции стандарта 802.11, в отличие от проводных Ethernet, не способны обнаруживать коллизии в среде передачи данных [3, 4]. В силу этого обнаружение коллизий и бесконфликтных передач протокольных блоков данных основано на механизме тайм-аутов и алгоритме положительной решающей обратной связи.

Вкратце цикл передачи кадра данных от станции-отправителя к станции-получателю выглядит следующим образом. Прежде всего, станция-отправитель прослушивает среду для определения ее занятости. Далее по истечении межкадрового интервала запускается алгоритм случайной задержки для выбора номера слота, в котором можно начать передачу данных. Номер слота равновероятно выбирается из промежутка $[0, 2^{N_0+N} - 1]$, где N_0 – значение, определяющее начальный размер конкурентного окна, а N – номер повторной передачи. Номер выбранного слота присваивается значению таймера отсрочки t_o , после чего начинают отсчитываться слотовые интервалы. В конце каждого слотового интервала таймер отсрочки уменьшается на единицу, при этом прослушивается среда. Как только фиксируется занятость среды, таймер отсрочки замораживается (останавливается) до тех пор, пока не освободится среда передачи данных. После освобождения среды таймер запускается со значения, зафиксированного непосредственно перед замораживанием. По истечении

таймера отсрочки ($t_o = 0$) станция-отправитель начинает передачу кадра данных. По окончании передачи отправитель ждет квитанции в течении времени t_{out} , по истечении которого считается, что была коллизия и станции, попавшие в коллизию, увеличивают значение N на единицу, а действия, направленные на передачу данных, повторяются.

Математическое моделирование. Рассмотрим функционирование беспроводной локальной сети до первой безошибочной передачи кадра и получения квитанции об успешной доставке данных рассматриваемым абонентом. Исходя из этого, найдем основные операционные характеристики системы. Предположим, что в беспроводной ЛВС имеется K станций – источников данных. Считаем, что все источники независимы, равноправны, всегда имеют кадры данных для отправки, а все интервальные промежутки выражены в слотовых интервалах t_c .

Пусть все станции обмениваются кадрами одинакового размера. Тогда согласно последовательности протокольных действий элементарный цикл отправки кадра получателю определится размером межкадрового промежутка t_m , периодом случайной отсрочки t_o , длительностью «заморозки» таймера случайной отсрочки t_z , временем передачи информационного кадра t_k , а также величиной тайм-аута ожидания положительной квитанции t_{out} , которая складывается из короткого межкадрового промежутка и времени передачи положительной квитанции. Следует отметить, что среди указанных компонент цикла времена t_o и t_z являются функциями номера повторной передачи.

Среднее время передачи кадра $T(K)$ складывается из взвешенной суммы средних времен ожидания неудачных отправок и времени, затраченного на успешную передачу. Поскольку при неудачных передачах кадра получатель не отправляет подтверждений, то среднее время определится соотношением

$$T(K) = t_m + t_k + t_{out} + \sum_{N=0}^{\infty} \left\{ N(t_m + t_k + t_{out}) + \sum_{n=0}^{N-1} t(n, K) + \tau(N, K) \right\} f(N, K). \quad (1)$$

Здесь $t(N, K)$ и $\tau(N, K)$ – средние условные времена до неудачной

и успешной N -ой повторной попытки отправить кадр рассматриваемым абонентом, а $f(N, K)$ – функция вероятностей продолжительности конкуренции между абонентами за эфир, которая определяется вероятностью успешной передачи кадра на N -ом повторном шаге после $(N - 1)$ -ой неудачной повторной попытки послать отправленное: $f(N, K) = P(N, K) \prod_{n=0}^{N-1} \pi(n, K)$, где $P(N, K)$ – условная вероятность успешной передачи кадра данных, рассматриваемым абонентом после $(N - 1)$ -ой конфликтной передачи в сети из K станций, а $\pi(N, K)$ – условная вероятность неудачи на N -ом повторном шаге. Для этих вероятностей выполняется следующее соотношение $\pi(N, K) = 1 - P(N, K)$.

Беспроводная ЛВС с $K = 2$ станциями. Рассмотрим случай с $K = 2$ беспроводными станциями локальной вычислительной сети. Вероятность выбора i -го слота для передачи данных на n -ой повторной передаче рассматриваемым абонентом обозначим $p_n(i)$, а вероятность выбора i -го слота для передачи данных на n -ой повторной передаче соперничающей станцией обозначим через $f_n(i)$. Тогда условная вероятность возникновения конфликта на n -ой повторной передаче для рассматриваемого абонента будет

$$\pi(0, 2) = p_0(0)f_0(0) + \sum_{i=1}^{2^{N_0}-1} p_0(i)Q_i; \quad (2)$$

$$\pi(n, 2) = \sum_{k=1}^n \frac{E_k(n)}{\pi(n-1, 2)} \left\{ \sum_{i=0}^{2^{N_0+k}-1} p_n(i)f_k(i) + \sum_{i=1}^{2^{N_0+k}-1} p_n(i) \sum_{j=0}^{i-1} f_k(j)L_{i-j} + \sum_{i=2^{N_0+n}-1}^{2^{N_0+n}-1} p_n(i) \sum_{j=0}^{2^{N_0+k}-1} f_k(j)L_{i-j} \right\}, \quad n > 0,$$

где $L_k = Q_k = \sum_{j=0}^{\infty} f_0^j(0) \sum_{i=0}^{k-1} f_0(k-i)Q_i$, для $k = \overline{1, 2^{N_0}-1}$, $Q_0 = 1$, а при $k = \overline{2^{N_0}, 2^{N_0+n}-1}$ имеет следующий рекуррентный вид:

$$L_k = \sum_{j=0}^{\infty} f_0^j(0) \sum_{i=1}^{2^{N_0}-1} f_0(i)L_{k-i}. \quad (3)$$

Элементы L_k это по сути различные наборы таких вариантов попадания в конфликт рассматриваемой и соперничающей станций,

при которых соперничающая станция успевает один или несколько раз бесконфликтно передать данные, пока рассматриваемый абонент дожидается своей n -ой повторной передачи. Коэффициенты $E_m(n)$ являются вероятностями того, что на n -ой повторной передаче рассматриваемого абонента у соперника это будет m -ая повторная передача, и они имеют следующий вид:

$$\begin{aligned}
E_1(1) &= p_0(0)f_0(0) + \sum_{i=1}^{2^{N_0}-1} p_0(i)Q_i; \\
E_1(n) &= \sum_{k=1}^{n-1} \left\{ \frac{E_k(n-1)}{\sum_{i=1}^{n-1} E_i(n-1)} \left[\sum_{i=1}^{2^{N_0+k}-1} p_{n-1}(i) \sum_{j=0}^{i-1} f_k(j)L_{i-j} + \right. \right. \\
&\quad \left. \left. + \sum_{i=2^{N_0+k}}^{2^{N_0+n-1}-1} p_{n-1}(i) \sum_{j=0}^{2^{N_0+k}-1} f_k(j)L_{i-j} \right] \right\}, \quad n > 1; \quad (4) \\
E_m(n) &= \sum_{i=1}^{2^{N_0+m-1}-1} p_{n-1}(i)f_{m-1}(i), \quad n > 1, \quad m = \overline{2, n}.
\end{aligned}$$

Среднее условное время до неудачной n -ой попытки передачи данных $t(n, 2)$ складывается из среднего количества слотов до начала передачи и среднего количества заморонок из-за передачи данных соперничающей станцией $N_s(n)$ и $Z_t(n)$ соответственно.

$$t(n, 2) = N_s(n) + Z_t(n)[t_k + t_{out} + t_m], \quad (5)$$

где $N_s(n) = \sum_{i=0}^{2^{N_0+n}-1} ip_n(i)$, а $Z_t(n) = \sum_{i=1}^{2^{N_0}-1} p_0(i)M_i$, при $n = 0$ и

$$\begin{aligned}
Z_t(n) &= \sum_{k=1}^n \frac{E_k(n)}{\pi(n-1, 2)} \left\{ \sum_{i=1}^{2^{N_0+k}-1} p_n(i) \sum_{j=0}^{i-1} f_k(j)M_{i-j} + \right. \\
&\quad \left. + \sum_{i=2^{N_0+k}}^{2^{N_0+n}-1} p_n(i) \sum_{j=0}^{2^{N_0+k}-1} f_k(j)M_{i-j} \right\}, \quad \text{при } n > 0. \quad (6)
\end{aligned}$$

Элементы M_k здесь собственно и показывают, сколько в среднем заморонок вынужден будет сделать рассматриваемый абонент, если на

n -ой повторной передаче он выберет i -ый слот, когда как соперничающая станция выберет j -ый слот (предшествующий i -му).

$$M_k = f_0(k) \sum_{i=0}^{\infty} i f_0^i(0) + \sum_{m=1}^{k-1} f_0(m) \sum_{i=0}^{\infty} (i+1 + M_{k-m}) f_0^i(0), \quad n = 0;$$

$$M_k = f_0(k) \sum_{i=0}^{\infty} (i+1) f_0^i(0) + \sum_{m=1}^{k-1} f_0(m) \sum_{i=0}^{\infty} (i+1 + M_{k-m}) f_0^i(0),$$

$$n > 0, \quad k = \overline{1, 2^{N_0} - 1};$$

$$M_k = \sum_{m=1}^{2^{N_0}-1} f_0(m) \sum_{i=0}^{\infty} (i+1 + M_{k-m}) f_0^i(0), \quad n > 0, \quad k = \overline{2^{N_0}, 2^{N_0+n} - 1}.$$

Среднее условное время до успешной передачи на n -ой повторной попытке передачи данных $\tau(n, 2)$ получается по аналогии с (5) и также складывается из среднего количества слотов до передачи и среднего количества заморозок из-за передачи данных соперничающей станцией $N_s(n)$ и $Z_\tau(n)$:

$$\tau(n, 2) = N_s(n) + Z_\tau(n)[t_k + t_{out} + t_m].$$

Среднее количество заморозок $Z_\tau(n)$ имеет схожий вид с $Z_t(n)$ (6), в запись которого входят элементы V_k , схожие с элементами M_k .

Обсуждение полученных результатов. Уже на этапе формализации этой задачи стало понятно, что получить аналитическое решение будет крайне непросто. Рекуррентные зависимости, такие как (3), (4), а также имеющиеся в записях элементов M_k и V_k , не позволяют с легкостью получить аналитический вид.

Пока что абсолютно четкий аналитический вид удалось получить для вероятности возникновения конфликта на самой первой передаче (нулевая повторная) (2):

$$\pi(0, 2) = \frac{2^{N_0} 2^{2^{N_0}} - (2^{N_0} - 1)^{2^{N_0}}}{2^{2^{N_0}} (2^{N_0} - 1)^{2^{N_0}-1}}.$$

Численные исследования среднего времени передачи кадра абонентом показывают, что функция (1) имеет ярко выраженный минимум по координате N_0 (рис. 1), определяющей начальный размер

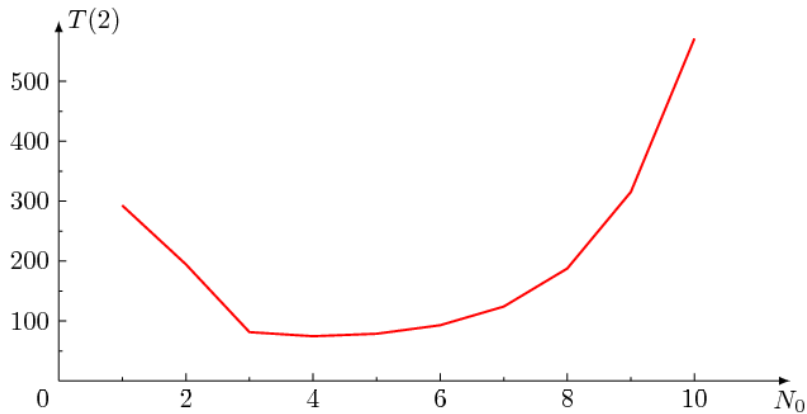


Рис. 1. Зависимость среднего времени передачи кадра $T(K)$ (1) от степени начальной ширины окна N_0 для двух абонентов сети ($K = 2$)

конкурентного окна и, как следствие, степень рассеяния станций по длительностям отсрочки перед началом процедуры соперничества.

Кроме того, уже на этапе формализации задачи стал очевиден эффект захвата среды передачи данных одним из абонентов, о котором упоминается в [4]. Особенно сильно этот эффект проявляется при малых значениях N_0 .

Литература

1. IEEE Std 802.11 – 2007, Revision of IEEE Std 802.11 – 1999. Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications. IEEE Computer Society, 2007. 1184 p.
2. Олифер В.Г., Олифер Н.А. Компьютерные сети. СПб.: Питер, 2006. 958 с.
3. Новиков Ю.В., Кондратенко С.В. Локальные сети: Архитектура, алгоритмы, проектирование. М.: ЭКОМ, 2000. 312 с.
4. Вишневикий В.М., Ляхов А.И., Портной С.Л., Шахнович И.В. Широкополосные беспроводные сети передачи информации. М.: Техносфера, 2005. 592 с.

Нвохири А.М.

Санкт-Петербургский государственный университет

Ранжирование веб-сайтов нигерийских университетов

Рекомендовано к публикации доцентом Печниковым А.А.

1. Введение. Термин «вебометрика» определяет научное направление, в рамках которого исследуются количественные аспекты конструирования и использования информационных ресурсов, структур и технологий применительно к World Wide Web (далее – Веб), опирающееся на библиометрические и информметрические подходы [1].

Проект испанской исследовательской группы Cybermetrics Lab «Webometrics Ranking of World Universities» [2] – один из наиболее известных вебометрических проектов. Он посвящен вебометрическим исследованиям сайтов вузов и научно-исследовательских институтов мира.

Проведенное автором самостоятельное исследование сайтов университетов Нигерии, использующее подходы Cybermetrics Lab, позволило выявить ряд интересных моментов. Предложен модифицированный подход к ранжированию сайтов и с его использованием определен рейтинг сайтов нигерийских университетов.

2. Целевое множество исследования. Выбор официальных сайтов университетов Нигерии в качестве объекта исследований позволяет получить достаточно обширную выборку. Основным способом выявления сайтов для целевого множества являлся анализ информации, содержащейся на сайте National Universities Commission (NUC) [3]. NUC – государственный орган, который дает разрешение на учреждение высших учебных заведений, предлагающих образовательные программы, и проводящий аккредитацию всех учебных программ университетов Нигерии.

В случае отсутствия на сайте NUC данных о доменном имени сайта университета выполнялся поиск в поисковых системах Google, Yahoo и Bing по его названию. Каждый найденный веб-адрес несколько раз проверялся на работоспособность. Из 104 университетов, найденных на сайте [3], 20 не рассматриваются в рамках данного исследования, поскольку либо доменные имена их официальных

сайтов отсутствовали на сайте NUC, либо поиск с помощью Google, Yahoo и Bing не дал результатов, либо найденные доменные имена оказались неработающими.

В случае наличия у университета нескольких сайтов в целевое множество включался сайт с наибольшими значениями веб-индикаторов, описанных в следующем разделе.

В итоге было сформировано целевое множество исследования, содержащее 84 официальных сайта университетов, из которых 23 являются сайтами федеральных университетов, 25 – университетов нигерийских штатов и 36 – частных университетов.

Каждый сайт целевого множества однозначно определяется своим уникальным доменным именем. Большинство университетских сайтов имеют доменные имена третьего уровня в зоне ng, как например сайт University of Ilorin с именем www.unilorin.edu.ng. Меньшая часть сайтов имеют доменные имена в зонах com, org и net. Все сайты разработаны на английском языке.

3. Индикаторы и методики измерений поисковыми машинами. В результате проведенного предварительного анализа из известных поисковых машин были выбраны Google, Yahoo и Bing. Например, известная поисковая машина Яндекс не была выбрана, так как она в основном индексирует русскоязычную часть Веба [4]. Для измерения так называемой «научности» сайтов используется Google Scholar, не имеющая работоспособных аналогов.

В соответствии с [2] определяются следующие четыре вебометрических индикатора: S (size) – размер сайта, т.е. количество страниц, обнаруживаемых на сайте поисковыми машинами Google, Yahoo, и Bing; V (visibility) – видимость сайта, т.е. количество уникальных гипертекстовых ссылок с других веб-ресурсов, обнаруживаемых Google, Yahoo, и Bing; R (rich files) – количество полнотекстовых файлов, под которыми понимаются файлы с расширениями *.pdf, *.ps, *.doc, *.xls, *.ppt и *.rtf; Sc (scholar) – количество ссылок на научные статьи, размещенные на сайте, обнаруживаемые Google Scholar.

3.1. Измерения в Google. Размер сайта S измеряется на странице <http://www.google.com>, в строке поиска вводится «site:доменное имя сайта». Результаты поиска выводятся под строкой поиска в виде «About S results (T seconds)», где T – время поиска. Видимость сайта V измеряется на странице <http://www.google.com>,

в строке поиска вводится «link:доменное имя сайта». Результаты поиска выводятся под строкой поиска в виде «About V results (T seconds)». Количество полнотекстовых файлов R измеряется на странице <http://www.google.com>, в строке поиска вводится «filetype:тип site:доменное имя сайта». В позиции «тип» поочередно выбираются pdf, ps, doc, rtf, ppt, xls. Для каждого типа файла результаты выводятся под строкой поиска в виде «About R results (T seconds)», результаты по всем типам суммируются. Количество ссылок на научные статьи Sc измеряется на странице <http://scholar.google.com>, в строке поиска вводится «site:доменное имя сайта», в выпадающем меню выбирается опция «at least summaries». Результаты поиска выводятся под строкой поиска, а их количество – слева в виде «Results 1– N of about Sc . (T sec)», где N – целое число.

3.2. Измерения в Yahoo. Размер сайта S измеряется на странице <https://siteexplorer.search.yahoo.com>, в строке поиска вводится доменное имя сайта. Промежуточные результаты поиска выводятся под строкой поиска в виде «Pages (S)». Далее выбирается опция «Only this domain», а окончательные результаты выводятся под строкой поиска в виде «Pages (S)». Видимость сайта V измеряется на странице <https://siteexplorer.search.yahoo.com>, в строке поиска вводится доменное имя сайта. Промежуточные результаты поиска выводятся под строкой поиска в виде «Inlinks (V)». Выбирается кнопка «Inlinks (V)», в разделе выбора опций «Show Inlinks:» устанавливаются опции «Except from this domain» и «Entire Site». Окончательные результаты выводятся под строкой поиска в виде «Inlinks (V)». Количество полнотекстовых файлов R измеряется на странице <https://siteexplorer.search.yahoo.com>, в строке поиска вводится «filetype:тип site:доменное имя сайта». Последующие действия – как и в случае Google. Для каждого типа файла результаты выводятся во фразе « R results for filetype:тип site:доменное имя сайта». Результаты по всем типам суммируются.

3.3. Измерения в Bing. Размер сайта S измеряется на странице <http://www.bing.com>, в строке поиска вводится «site:доменное имя сайта». Результаты поиска выводятся под строкой поиска в виде «1– N of S results», где N – целое число. Видимость сайта V измеряется на странице <http://www.bing.com>, в строке поиска вводится «inbody:доменное имя сайта». Результаты поиска выводятся под строкой поиска в виде «1– N of V results». Количество полнотексто-

вых файлов R измеряется на странице <http://www.bing.com>, в строке поиска вводится «filetype:тип site:доменное имя сайта». Последующие действия – как у Google. Для каждого типа файла результаты выводятся под строкой поиска в виде «1– N of R results», результаты по всем типам суммируются.

Так как содержание Веба очень часто меняется, а также учитывая постоянную деятельность роботов поисковых машин, следует отметить, что измеренное значение индикатора – это значение, полученное данной поисковой машиной в конкретный момент времени. Измерения индикаторов S , V , R и Sc проводились в январе – феврале 2011 г.

4. Ранжирование сайтов. Для ранжирования сайтов целевого множества использовалась модифицированная формула из [2] (раздел Methodology):

$$Rt(i) = 5RankV_a(i) + 2RankS_a(i) + 1,5RankR_a(i) + 1,5RankSc(i)$$

где V_a – среднее количество уникальных гипертекстовых ссылок с других веб-ресурсов, обнаруживаемых Google, Yahoo, и Bing; S_a – среднее количество страниц, обнаруживаемых на сайте этими поисковыми машинами; R_a – среднее количество полнотекстовых файлов; $RankV_a(i)$ – порядковый номер i -го сайта в упорядоченном по убыванию векторе V_a . Значения $RankS_a(i)$, $RankR_a(i)$, $RankSc(i)$ определяются аналогично. Тот сайт, у которого значение $Rt(i)$ будет самым маленьким, получает значение W (Webometrics Rating) равное 1, следующий за ним – 2, и т.д. Первые пять официальных сайтов университетов, имеющие наивысший ранг, приведены в таблице.

Таблица. Рейтинг сайтов нигерийских университетов

Название	тип	URL	W	S_a	V_a	R_a	Sc
Uni. of Ilorin	фед	www.unilorin.edu.ng	1	8524	1298	2057	1280
Uni. of Lagos	фед	www.unilag.edu.ng	2	9392	950	399	663
National Open Uni. of Nig.	част	www.nou.edu.ng	3	1688	1206	2031	70
Uni. of Nigeria Nsukka.	фед	www.unn.edu.ng	4	32094	493	172	57
Ahm. Bel. Uni.	фед	www.abu.edu.ng	5	12008	502	100	28

5. Выводы. Выбор тех или иных поисковых машин остается спорным вопросом, поскольку во многих работах, включая и данную, есть примеры, которые заставляют задуматься о надежности

использования поисковых машин. Это вопрос, который должен быть тщательно исследован специалистами. Конечно, при измерении индикаторов сайтов конкретной страны необходимо использовать в первую очередь поисковые машины, наиболее распространенные в этой стране.

Можно предположить, что методики ранжирования и выбор поисковых машин не оказывают значительного влияния на положение явных лидеров. Результаты ранжирования, полученные в данной работе и результаты ранжирования, приведенные, например, в [5] для одних и тех же единиц анализа (в нашем случае это доменные имена официальных сайтов), совпадают на 90 процентов.

Так как механизмы индексации страниц являются коммерческой тайной, измерение количества страниц на сайте с помощью поисковой машины не обязательно является корректным, иначе было бы надежней измерять его с помощью программ-краулеров (например, [6]), которые дают более точные оценки. В случае использования замеров поисковыми машинами наверняка лучше взять среднее значение V , S , R , S_c , найденных разными поисковыми машинами, что и было сделано в этой работе.

Остановимся очень кратко на результатах ранжирования сайтов нигерийских университетов. В первую десятку входят 8 федеральных университетов, 1 частный и 1 университет штата. Эти 8 университетов действительно являются одними из самых «крутых» в Нигерии. University of Ibadan – первый университет Нигерии. University of Nigeria – один из самых престижных нигерийских университетов со старейшим юридическим факультетом. Ahmadu Bello University – самый большой университет в Нигерии, который был изначально создан для граждан северной Нигерии.

Частный National Open University of Nigeria является единственным университетом, специализирующемся в области открытого и дистанционного высшего обучения [7]. Единственный университет нигерийского штата, который попал в первую десятку – Lagos State University, находится в крупнейшей метрополии Лагос.

Поэтому можно сказать, что традиции, современные методы обучения, исторические и географические аспекты повлияли на рейтинг сайтов вышеуказанных университетов.

Из 36 частных университетов 22 заняли места ниже среднего. Только 4 федеральных университета имеют такой же показатель, но 3 из этих 4 федеральных университетов – специализированные (в на-

шем случае это такие направления, как оборона, сельское хозяйство, нефть). Это означает, что количество учебных программ университета влияет на рейтинг его сайта.

6. Заключение. В целом, сайты федеральных университетов имеют значения рейтингов больше, чем сайты университетов нигерийских у штатов и частных университетов. Федеральные университеты получают большее финансирование, считается, что у них есть более квалифицированные сотрудники, в них сложнее поступить, что и увеличивает их престижность. Можно считать, что указанные факторы нашли свое отражение в Вебе.

Вебометрические исследования рейтингов сайтов смогут послужить основой для более точного количественного сравнительного анализа некоторых характеристик уже созданных веб-ресурсов, а значит, позволит определить направления их дальнейшего развития.

Литература

1. Bjerneborn L., Ingwersen P. Toward a basic framework for webometrics // Journal of the American Society for Information Science and Technology, 2004. Vol. 14, № 55. P. 1216–1227.
2. Ranking Web of World Universities
URL: <http://www.webometrics.info>
3. National Universities Commission (NUC) of Nigeria.
URL: www.nuc.edu.ng
4. Шокин Ю.И., Клименко О.А., Рычкова Е.В., Шабальников И.В. Рейтинг сайтов научных организаций СО РАН // Вычислительные технологии, 2008. Т. 13, № 3. С. 128–135.
5. Top Colleges and Universities in Nigeria.
URL: <http://www.4icu.org/ng>
6. SocSciBot 3. Link crawler for the social sciences. 2007.
URL: <http://socscibot.wlv.ac.uk>
7. National Open University of Nigeria.
URL: <http://www.nou.edu.ng>

Погожев С.В., Шарлай А.С.
Санкт-Петербургский государственный университет

Об определении выделенных формул в изображениях печатного текста

В настоящее время широко распространяется применение электронных методов работы с информацией, но, несмотря на использование электронных документов, объем применяемых бумажных документов на практике не только не уменьшается, но и возрастает. Происходит это вследствие роста объемов обрабатываемых данных, а также обусловлено удобством использования и давно укоренившейся традицией чтения и архивирования бумажных документов. В свою очередь, электронный документ позволяет осуществлять редактирование текста, поиск слов или фраз, анализировать информацию, а также применять к тексту электронный перевод, форматирование, преобразование в речь, позволяя оказать помощь лицам с нарушениями зрения в чтении документов. И если преобразование электронного документа в бумажный не вызывает затруднения, то автоматизированная обработка бумажных документов и преобразование их в электронный вид является хорошо известной, но непростой задачей. Одним из подходов к решению является использование систем оптического распознавания символов.

Существующие системы распознавания показывают высокую точность обработки фрагментов текста, позволяя восстанавливать исходное форматирование текста, включая колонки, таблицы, изображения. Однако существуют серьезные ограничения при преобразовании научных работ, технических документов, которые, как правило, содержат большое количество математических выражений (формул), с которыми системы распознавания не могут справиться. На рис. 1 и рис. 2 продемонстрирован один из примеров, полученный из популярной системы оптического распознавания символов АBBYY FineReader OCR Professional 8.0. Документ был распознан с ошибками как в структуре документа (выделенная формула (8.72) была разбита на несколько блоков), так и в распознавании математических символов. Это происходит по той причине, что существующие системы не в состоянии правильно распознать математические выражения в тексте, а также отчасти ошибки вызваны тем, что математические выражения являются сложными объектами для распо-

знания, так как включают большой набор символов из различных шрифтов и размеров и, кроме того, используют сложные пространственные связи между символами. В связи с этим преобразование научных работ от печатной к электронной форме остается трудной задачей. Среди большого количества программных средств, предназначенных для распознавания, было выявлено только одно программное средство – InftyReader, которое изначально ориентировано на распознавание научных документов с математическими формулами. Однако данное программное средство, разрабатываемое организацией InftyProject, в состав которой входят исследователи японских университетов, является коммерческим и выдвигает жесткие ограничения на качество исходного отсканированного изображения.

Moreover, in the case without constraints, *i.e.*, $\mathcal{K}_1 = L^2(\Omega)$ and $\mathcal{K}_2 = L^2(\Sigma)$ the optimality conditions (8.71) become

$$\frac{\tilde{u}^*}{\gamma_1} = -\theta^*, \quad \frac{\tilde{u}^*|_{\Sigma}}{\gamma_2} = \xi^*,$$

so, we can obtain the optimal control by the resolution of the coupled system:

$$\begin{cases} \frac{\partial u^*}{\partial t} + A(t)u^* + a_0 u^* = \varphi & \text{on } \mathcal{Q}, \\ \frac{\partial \tilde{u}^*}{\partial t} + A^*(t)\tilde{u}^* + a_0 \tilde{u}^* = 0 & \text{on } \mathcal{Q}, \\ \alpha \frac{\partial u^*}{\partial \eta_A} + \beta u^* = \frac{\alpha \tilde{u}^*|_{\Sigma}}{\gamma_2} + \alpha \psi_0 & \text{on } \Sigma, \\ \alpha \frac{\partial \tilde{u}^*}{\partial \eta_{A^*}} + \beta \tilde{u}^* = \mathcal{D}^*(\mathcal{D}(u^*|_{\Sigma}) - u_{\text{obs}}) & \text{on } \Sigma, \end{cases} \quad (8.72)$$

$$u^*(0) = -\frac{\tilde{u}^*}{\gamma_1} + v_0, \quad \tilde{u}^*(T) = 0 \quad \text{on } \Omega. \quad \diamond$$

In the two next sections we study two classes of non-linear robust control problems, more precisely a class of bilinear problems (the primal problem is linear on the state variable when the control is fixed, and conversely) and

Рис. 1. Пример автоматического выделения блоков в тексте

Весь процесс распознавания формул в научных документах можно представить в виде следующих задач: идентификация блоков, содержащих математические выражения, распознавание символов, интерпретация структуры математического выражения. Для решения второй и третьей задач предложено много различных методов (обзор публикаций можно найти, например в [1]). Большинство предложенных методов предполагают, что математические выражения доступ-

Moreover, in the case without constraints, i.e., $K = \emptyset$ and $IC2 = \emptyset$, the optimality conditions (8.71) become

$$\begin{aligned} -\dot{p} &= -Q^* p + f^* \\ p(T) &= p_0 \end{aligned}$$

so, we can obtain the optimal control by the resolution of the coupled system:

$$\begin{aligned} \dot{u}^* &= A(t)u^* + a_0 u^* - ip \text{ on } Q, \\ \dot{p}^* &+ A^*(t)p^* + a_0^* p^* = 0 \text{ on } Q, \\ A^* + \lambda u^* &= \lambda^* + a_0^* \text{ on } \mathcal{L}, \\ \text{or } A^* &= \lambda^* + a_0^* \\ \dot{u}^* &+ f_3 u^* = V^* \{V(u^*) - u_{obs}\} \text{ on } \mathcal{L}, \\ u^*(0) &= u_0 + v_0, \quad u^*(T) = 0 \text{ on } \mathcal{L}. \end{aligned} \tag{8.72}$$

In the two next sections we study two classes of non-linear robust control problems, more precisely a class of bilinear problems (the primal problem is linear on the state variable when the control is fixed, and conversely) and

Рис. 2. Результат распознавания

ны в изолированном от текста виде, являются уже «вырезанными» из текста, т.е. первая задача предполагается решенной.

Математические выражения обычно содержатся в документах либо в виде выделенных формул (displayed), либо как встроенные (embedded), т.е. формулы в тексте. Выделенные формулы проще идентифицируются, так как занимают отдельные строки в тексте, высота которых зачастую отличается от высоты строк текста, также данные строки зачастую имеют интервалы перед строкой и после. Сегментация строк текста на предмет выделения в них встроенных формул является более сложной задачей. В то же самое время, встроенные формулы, как правило, являются достаточно простыми, небольшими выражениями, которые проще распознать, нежели сложные математические выражения в выделенных формулах.

Данная работа посвящена разработке метода для идентификации выделенных формул на изображениях отсканированных документов.

Среди существующих подходов к идентификации блоков математических выражений особого внимания заслуживают методы, предложенные Гарайном [2], Ли и Вонгом [3], Фейтменом [4], Иноэ и др. [5], Туми [6]. Однако все эти методы для классификации бло-

ков текста требуют распознавания символов. В статье [7] предложен двухпроходный алгоритм, который не делает упор на распознавание символов. Первоначально выражения выделяются в тексте с использованием первичной маркировки, которая использует нечеткую логику для определения некоторых математических символов. Далее, вторичная маркировка использует некоторые эвристики и находит верхние и нижние индексы в тексте. Позже Chowdhury и др. [8] предложили подход без распознавания символов, который использует обычное пространственное распределение черных пикселей в математических формулах. Экспериментальные результаты показывают, что метод хорошо работает для сегментации выделенных формул, а для формул в тексте дает невысокую точность.

Будем предполагать, что на изображении тем или иным способом уже выделены горизонтальные блоки со строками и наша задача состоит в классификации этих строк и определении тех, что содержат выделенные формулы.

Обычно строки с выделенными формулами имеют некоторые особенности, позволяющие отличать их от обычных строк текста. Например, строки, содержащие выделенные формулы, обычно имеют высоту больше, чем текстовые строки, а также, перед и после строки с выделенной формулой чаще всего добавляется дополнительный отступ. Это подтверждается и статистическим исследованием [2], в котором получено, что среднее значение расстояния между двумя строками текста составляет около 0,4 от средней высоты строк текста документа в то время, как среднее значение отступов выше и ниже для строк с выделенными формулами почти в 1,8 раза больше высоты текста. Следующей особенностью, на которую обратим внимание, является выравнивание формул в строках. Чаще всего блок с формулой располагается по центру строки, а номер формулы, если он есть, выровнен по правому краю строки. Другая важная характеристика строк с выделенными формулами связана с пространственным расположением символов. Для обычных текстовых строк координаты центров блоков (по вертикали), ограничивающих символы, преимущественно лежат на одной или двух прямых линиях, а для строк с математическими формулами, как правило, наблюдается существенный разброс, так как в формулах зачастую присутствуют верхние и/или нижние индексы, дроби и т.д. Кроме того, наличие дробей, знаков сумм, интегралов, матриц, систем уравнений и т.п., которые зачастую и являются причиной для создания выделенной

формулы, приводит к следующему: если рассмотреть строку с выделенной формулой и построить для нее ограничивающий блок, удалив отступы слева и справа строки, то отношение количества черных пикселей в этом блоке к общей сумме пикселей в нем, будет меньше, чем в строках текста не содержащих выделенные формулы.

Основываясь на этих наблюдениях, вычислим для каждого горизонтального блока (строки) на странице следующие величины:

1. $f_1 = \max\left(\frac{h-\bar{h}}{h_m-\bar{h}}, 0\right)$, где h – высота строки, \bar{h} – средняя высота строки на странице, h_m – максимальная высота строки.
2. $f_2 = \max\left(\frac{s-\bar{s}}{s_m-\bar{s}}, 0\right)$, где s – среднее арифметическое отступов перед и после строки, \bar{s} – среднее расстояние между строками, s_m – максимальное расстояние между строками.
3. $f_3 = \max\left(\frac{m-\bar{m}}{m_m-\bar{m}}, 0\right)$, где m – отступ слева строки, \bar{m} – среднее значение отступов строк на странице, m_m – максимальный отступ.
4. $f_4 = \max\left(\frac{\sigma_y-\bar{\sigma}_y}{\sigma_{ym}-\bar{\sigma}_y}, 0\right)$, где σ_y – стандартное отклонение координат (по вертикали) центров блоков, ограничивающих символы в строке от середины высоты строки, $\bar{\sigma}_y$, σ_{ym} – среднее и максимальное значения стандартных отклонений координат центров блоков на странице.
5. $f_5 = \max\left(\frac{b-\bar{b}}{b_m-\bar{b}}, 0\right)$, где b – отношение количества черных пикселей к общему количеству пикселей в блоке, ограничивающем строку, \bar{b} и b_m – среднее и минимальное значения для строк на странице.

Каждая из вычисленных величин принимает значения на отрезке $[0, 1]$. Далее вычислим среднеарифметическое полученных величин $f = (f_1 + f_2 + f_3 + f_4 + f_5)/5$. Если полученная величина принимает значения больше некоторого порогового значения p , то данную строку классифицируем как строку, содержащую выделенную формулу, иначе классифицируем как текстовую. В ходе проведенных экспериментов было получено, что в качестве порогового значения можно принять величину $p = 0,4$.

Литература

1. Xue-Dong Tian, Hai-Yan Li, Xin-Fu Li, Li-Ping Zhang. Research on Symbol Recognition for Mathematical Expressions // ICICIC, 2006. Vol. 3. P. 357–360.
2. Garain U. Identification of mathematical expressions in document images // Proc. of the 10th Int. Conf. on Document Analysis and Recognition (ICDAR), 2009. P. 1340–1344.
3. Lee H.J., Wang J.-S. Design of a mathematical expression understanding system // Pattern Recognition Letters, 1997. Vol. 18, No 3. P. 289–298.
4. Fateman R.J. How to find mathematics on a scanned page // Proc. of the SPIE, 1999. Vol. 3967. P. 98–109.
5. Inoue K., Miyazaki R., Suzuki M. Optical recognition of printed mathematical documents // Proc. of Asian Technology Conference in Mathematics (ATCM). New York: Springer, 1998. P. 280–289.
6. Toumit J.-Y., Garcia-Salicetti S., Emptoz H. A hierarchical and recursive model of mathematical expressions for automatic reading of mathematical documents // Proc. of the 5th ICDAR, 1999. P. 119–122.
7. Kacem A., Belaid A., Ben Ahmed M. Automatic extracton of printed mathematical formulas using fuzzy logic and propagation of context // Int. Journal on Document Analysis and Recognition (IJ DAR), 2001. Vol. 4, No 2. P. 97–108.
8. Chowdhury S.P., Mandal S., Das A.K., Chanda B. Automated segmentation of math-zones from document images // Proc. of the 7th ICDAR, 2003. P. 755–759.

Проскура Е.С.

Санкт-Петербургский государственный университет,
ЗАО «Диджитал Дизайн»

Автоматизация получения пользовательского сертификата для мобильного клиента MS Exchange

*Рекомендовано к публикации старшим преподавателем
Севрюковым С.Ю.*

Ведение. Исследование посвящено реализации криптографического стандарта S/Mime в почтовом мобильном клиенте Mail для MS Exchange. Для данной реализации нужны открытый и закрытый ключ, выданные Центром Сертификации. Пара ключей обычно хранится в PKCS#12 файле [1]. Современные мобильные клиенты предполагают импортировать данный файл на мобильное устройство одним из следующих способов:

- используя приложение MS Outlook;
- запросить у системного администратора MS Exchange.

Упомянутые выше механизмы, предлагаемые платформой, не технологичны и не применимы для использования в автоматизированных системах.

Постановка задачи. Требуется автоматизировать процесс получения файла ключей PKCS#12 в рамках реализации следующих прикладных задач:

1. Использование только электронной цифровой подписи (ЭЦП).
2. Использование S/Mime шифрование при участии MS Outlook на стороне отправителя.

Известны следующие начальные данные:

- доменная учетная запись, обеспеченная доступом к Центру сертификации;
- обеспечен доступ к каталогу LDAP с настройками по умолчанию;

В рамках исследования не рассматривается ситуация, при которой все необходимые данные хранятся в каталоге LDAP (в атрибуте UserPKCS12Data) и доменная учетная запись обеспечена соответствующими правами на чтение. Данное допущение обусловлено небезопасностью такого подхода, который в промышленных средах не применяется [2].

Решение задачи 1. Чтобы пользоваться только функциями ЭЦП, необходимо реализовать традиционную схему, которую используют системные администраторы:

1. Сгенерировать новый закрытый ключ и запрос к Удостоверяющему центру (УЦ).
2. Сымитировать запрос через браузер (отправляется последовательность запросов из программы; информация об этом недоступна пользователю).
3. В ответ от УЦ получить X509 сертификат – новый сертификат пользователя (УЦ также автоматически пропишет его в каталоге LDAP в поле userCertificate для пользователя).
4. Получить сертификат компании, которым подписываются все пользовательские сертификаты (иначе нельзя доверять полученному пользовательскому сертификату). Он расположен по постоянной ссылке на сервере Удостоверяющего центра.

По окончании описанного выше процесса, будут получены 3 основные составляющие PKCS#12 файла, что позволит его создать и использовать в дальнейшем. На рис. 1 представлена схема решения.

Решение задачи 2. Полученный результат задачи 1 не справится с реализацией полноценного S/Mime шифрования. Если кто-нибудь решит отправить пользователю зашифрованное сообщение с помощью стандартного Outlook, пользователь сможет его расшифровать только со своего Outlook, но не с мобильного клиента. Для начала нужно рассмотреть работу Outlook при отправке зашифрованного сообщения. В LDAP каталоге у каждого пользователя есть два типа атрибута: UserCertificate и UserSmimeCertificate [2, 3]. Было бы логично, если бы они хранились по одному и тому же значению в единственном экземпляре. Но после запроса за новым сертификатом УЦ просто добавил еще одно значение в атрибут UserCertificate. В то же время Outlook, в первую очередь, ищет действительный

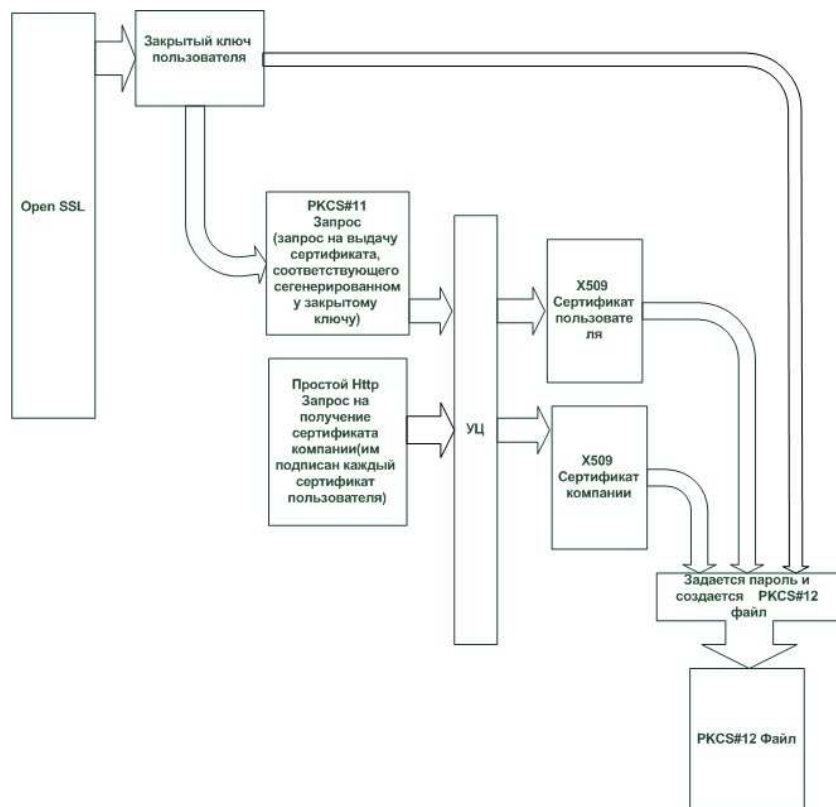


Рис. 1. Получение PKCS#12 файла

сертификат среди значений атрибута UserSmimeCertificate и если он его найдет, им и воспользуется [4]. Но закрытый ключ будет храниться не на мобильном устройстве, а на компьютере пользователя. Проблему легко решить. В LDAP каталоге пользователю по умолчанию разрешено менять значения любых своих атрибутов, на которые есть доступ на чтение. Поэтому достаточно просто заменить значение атрибута UserSmimeCertificate полученным сертификатом из задачи 1. Следует также учесть одну тонкость: от УЦ получен просто X509 сертификат, а в значение атрибута нужно будет записать PKCS#7 представление [2]. Чтобы его получить, необходимо подписать сертификат пользователя его же закрытым ключом. Далее отправить полученное представление в запросе на изменение. Те-

перь, когда пользователю придет зашифрованное письмо, он сможет его расшифровать. На рис. 2 отображен алгоритм решения задачи 2.



Рис. 2. Замена сертификата в LDAP

Выявленные недочеты. У полученного решения есть ряд минусов. Очевидно, письма, которые шифровались до работы с LDAP каталогом согласно задаче 2, не будут расшифрованы с мобильного устройства (как и в ситуации с задачей 1). Из того, что на мобильном устройстве и на постоянном компьютере лежат разные закрытые ключи, становится также очевидно, что все зашифрованные после взаимодействия с LDAP письма будут читаться только на мобильном клиенте, но не на основном компьютере. Возникшая проблема легко устраняется, но только до следующего использования функции шифрования на мобильном клиенте. В настройках безопасности Outlook необходимо выбрать «опубликовать сертификат в GAL». При тестировании решения задачи 2 была выявлена еще одна особенность. Письма, зашифрованные и отправленные со стандартных Exchange клиентов, основанных на Outlook Web Access (OWA), не будут расшифровываться на мобильном клиенте. OWA при шифровании действует противоположно Outlook [5]. Первым делом он ищет действительный сертификат среди значений UserCertificate LDAP каталога, и только если не найдет там, переходит к поиску среди значений UserSmimeCertificate. Значений у атрибута UserCertificate много, среди них значение, содержащее самый новый сертификат. Однако OWA использует не тот сертификат, который опубликовал УЦ, а другой. Почему так и по какому принципу выбирается именно этот сертификат, выяснить не удалось. Был проведен небольшой эксперимент. В распоряжении автора был Outlook и клиент Exchange для Mac, основанный на OWA, и по одной учетной записи на каждого клиента. Из стандартного Outlook выполнена команда «Опубликовать сертификат в GAL» (команда заменяет значение

UserSmimeCertificate). Далее из OWA клиента было отправлено зашифрованное письмо на стандартный Outlook клиент, которое не удалось расшифровать. Вероятно, OWA клиенты редко используют функции S/Mime шифрования, поэтому устранение выявленной ошибки не критично для создателей. У выявленной проблемы существует несколько возможных решений, но все они не оптимальны: можно открыть доступ пользователю на удаление значений любых своих атрибутов и убирать лишнее в UserCertificate. Такой подход небезопасен. Можно при помощи изменения значений атрибутов переписать все значения атрибута UserCertificate, кроме нужного. Но тут необходимо вспомнить, что количество значений у атрибута UserCertificate будет увеличиваться с каждым обращением к УЦ. Таким образом, может получиться, что изменения каждого неподходящего сертификата займет слишком много времени.

Литература

1. Управление ключами шифрования и безопасность сети. <http://www.intuit.ru/department/security/manencryptk/6/10.html>
2. RFC 2798 – Definition of the inetOrgPerson LDAP Object Class. <http://www.faqs.org/rfcs/rfc2798.html>
3. Википедия, LDAP. <http://ru.wikipedia.org/wiki/LDAP>
4. Outlook S/MIME certificate selection. <http://blogs.technet.com/b/pki/archive/2008/12/17/outlook-smime-certificate-selection.aspx>
5. Поддержка Microsoft. <http://support.microsoft.com/kb/822504>

Романовский Л.М.

Санкт-Петербургский государственный университет

Локальное укрупнение триангуляции и калибровочные соотношения¹

Рекомендовано к публикации профессором Демьяновичем Ю.К.

Введение. Построение адаптивных сеток необходимо при решении задач аппроксимации, при численном решении задач математической физики и при вейвлетной обработке числовых потоков информации (см. [1–4]). Построение адаптивных сплайн-вейвлетных разложений базируется на аппроксимационных соотношениях, вложенности пространств и на их проектировании с использованием биортогональных систем функционалов.

В случае двумерных потоков числовой информации построение адаптивных вейвлетов наталкивается на трудности: оказывается, не всякая триангуляция допускает локальное укрупнение с сохранением правильности триангуляции. Например, наиболее часто используемая триангуляция в методе конечных элементов не допускает упомянутого укрупнения. В данной работе делаются определенные шаги в направлении преодоления этой трудности.

Цель статьи состоит в том, чтобы в ситуации регулярной сетки рассмотреть такую триангуляцию плоскости, которая допускает локальное укрупнение с сохранением правильности триангуляции, и вывести калибровочные соотношения для координатных функций Куранта.

1. Локальное укрупнение триангуляции. Здесь дается описание исходной триангуляции и ее локального укрупнения.

В дальнейшем рассматривается плоскость, правильно подразделенная на треугольники (правильность подразделения означает, что вершина треугольника не может лежать внутри стороны другого треугольника). Полученное подразделение назовем *правильной триангуляцией*. Нас интересуют локальные укрупнения в исходной правильной триангуляции (т.е. объединения конечного числа треугольников), приводящие снова к правильной триангуляции.

¹Работа частично поддержана РФФИ, гранты № 10-01-00245, 10-01-00297

Для описания триангуляции достаточно указать таблицу инцидентий, каждая строка которой описывает треугольник перечнем инцидентных ему вершин.

Введем обозначение $\mathbb{Z}^2 \stackrel{\text{def}}{=} \{(i, j) \mid i, j \in \mathbb{Z}\}$. Пусть $M_{i,j} = (ih', jh'')$, где $h' > 0, h'' > 0, (i, j) \in \mathbb{Z}^2$. Точки $M_{i,j}$ служат вершинами определяемой ниже исходной триангуляции.

Рассмотрим триангуляцию, которая получается объединением таблиц

$$\left\| \begin{array}{ccc} M_{2i,2j} & M_{-2+2i,2j} & M_{-1+2i,1+2j} \\ M_{-2+2i,2+2j} & M_{-2+2i,2j} & M_{-1+2i,1+2j} \\ M_{-2+2i,2+2j} & M_{2i,2+2j} & M_{-1+2i,1+2j} \\ M_{2i,2j} & M_{2i,2+2j} & M_{-1+2i,1+2j} \end{array} \right\|$$

при $(i, j) \in \mathbb{Z}^2$. Далее показано, что эта триангуляция может быть укрупнена с сохранением правильности.

Укрупнение триангуляции будем проводить объединением двух соседних (т.е. имеющих общую сторону) треугольников. Полученные в результате треугольники будем называть *укрупненными* треугольниками.

Рассмотрим такое укрупнение, при котором вершину $M_{0,0}$ будут окружать лишь укрупненные треугольники. Для этого заменим перечисленные ниже соседние треугольники на треугольник, получающийся их объединением. Эквивалентное преобразование таблицы инцидентий состоит в том, что из нее исключаются строки, соответствующие объединяемым треугольникам, и добавляются строки, соответствующие результатам такого объединения – укрупненным треугольникам. Как было отмечено выше, расположение строк в таблице инцидентий не существенно, и потому строки могут быть добавлены между любыми строками упомянутой таблицы. Таким образом, достаточно перечислить выбрасываемые строки таблицы и указать вставляемые в нее строки. Однако, для наглядности преобразования, таблицы инцидентий будем задавать указанием двух строк заменяемых треугольников (в левой от стрелки части формулы) и указанием строки укрупненного треугольника (в правой части формулы). Для удобства чтения проиллюстрируем преобразования рисунком, на котором точки $M_{i,j}$ имеют декартовы координаты (i, j) .

Итак, укрупнение зададим следующим преобразованием таблицы

инциденций:

$$\begin{aligned}
\left\| \begin{array}{ccc} M_{0,0} & M_{-2,0} & M_{-1,1} \\ M_{-2,2} & M_{-2,0} & M_{-1,1} \end{array} \right\| &\longrightarrow \left\| M_{0,0} & M_{-2,0} & M_{-2,2} \right\|, \\
\left\| \begin{array}{ccc} M_{-2,2} & M_{0,2} & M_{-1,1} \\ M_{0,0} & M_{0,2} & M_{-1,1} \end{array} \right\| &\longrightarrow \left\| M_{0,0} & M_{0,2} & M_{-2,2} \right\|, \\
\left\| \begin{array}{ccc} M_{0,0} & M_{0,2} & M_{1,1} \\ M_{2,2} & M_{0,2} & M_{1,1} \end{array} \right\| &\longrightarrow \left\| M_{0,0} & M_{0,2} & M_{2,2} \right\|, \\
\left\| \begin{array}{ccc} M_{2,2} & M_{2,0} & M_{1,1} \\ M_{0,0} & M_{2,0} & M_{1,1} \end{array} \right\| &\longrightarrow \left\| M_{0,0} & M_{2,0} & M_{2,2} \right\|, \\
\left\| \begin{array}{ccc} M_{0,0} & M_{2,0} & M_{1,-1} \\ M_{2,-2} & M_{2,0} & M_{1,-1} \end{array} \right\| &\longrightarrow \left\| M_{0,0} & M_{2,0} & M_{2,-2} \right\|, \\
\left\| \begin{array}{ccc} M_{2,-2} & M_{0,-2} & M_{1,-1} \\ M_{0,0} & M_{0,-2} & M_{1,-1} \end{array} \right\| &\longrightarrow \left\| M_{0,0} & M_{0,-2} & M_{2,-2} \right\|, \\
\left\| \begin{array}{ccc} M_{0,0} & M_{0,-2} & M_{-1,-1} \\ M_{-2,-2} & M_{0,-2} & M_{-1,-1} \end{array} \right\| &\longrightarrow \left\| M_{0,0} & M_{-2,0} & M_{-2,-2} \right\|, \\
\left\| \begin{array}{ccc} M_{-2,-2} & M_{-2,0} & M_{-1,-1} \\ M_{0,0} & M_{-2,0} & M_{-1,-1} \end{array} \right\| &\longrightarrow \left\| M_{0,0} & M_{-2,0} & M_{-2,-2} \right\|.
\end{aligned}$$

Легко видеть, что в результате получается правильная триангуляция.

2. Структура барицентрических звезд исходной триангуляции. Для построения аппроксимации Куранта важна структура барицентрических звезд, соответствующих вершинам рассматриваемой триангуляции.

Для исходной триангуляции имеется два типа барицентрических звезд. К первому типу отнесем барицентрические звезды, содержащие четыре треугольника, а ко второму типу – барицентрические звезды с восемью треугольниками. Барицентрические звезды первого типа имеются у вершин $M_{-1+2i,1+2j}$, а барицентрические звезды второго типа – у вершин $M_{2i,2j}$; здесь $(i, j) \in \mathbb{Z}^2$.

3. Структура барицентрических звезд укрупненной триангуляции. При укрупнении триангуляции появляются дополнительно два типа барицентрических звезд, содержащих по шесть и по восемь треугольников. Барицентрические звезды вершин вида $M_{0,0}$,

$M_{\pm 2,2}$ и $M_{\pm 2,-2}$ состоят из восьми треугольников, а барицентрические звезды вершин вида $M_{0,\pm 2}$ и $M_{\pm 2,0}$ – из шести треугольников.

4. Калибровочные соотношения для функций Куранта. Как известно, функцией Куранта, ассоциированной с выделенной вершиной правильной триангуляции, называется непрерывная функция, равная единице в упомянутой вершине, линейная на каждом треугольнике барицентрической звезды этой вершины и равная нулю вне указанной барицентрической звезды.

Функцию Куранта, соответствующую выделенной вершине $M_{i,j}$ исходной триангуляции, обозначим $\omega_{i,j}$, $(i,j) \in \mathbb{Z}^2$. На исходной триангуляции имеется два типа функций Куранта, соответствующих рассмотренным выше двум типам барицентрических звезд: у функций Куранта с нечетными индексами носитель состоит из четырех треугольников, а у функций Куранта с четными индексами носитель состоит из восьми треугольников.

Введем обозначения

$$\mathbb{I}_1 \stackrel{\text{def}}{=} \{(0,0), (1,1), (-1,1), (1,-1), (-1,-1)\}, \quad \mathbb{I}'_1 \stackrel{\text{def}}{=} \mathbb{I}_1 \setminus (0,0).$$

Для укрупненной триангуляции функцию Куранта, соответствующую выделенной вершине $M_{i,j}$, будем обозначать $\tilde{\omega}_{i,j}$. Поскольку не все вершины исходной триангуляции участвуют в укрупненной триангуляции, то индексы (i,j) пробегают не все множество \mathbb{Z}^2 , а лишь его часть, а именно $(i,j) \in \mathbb{Z}^2 \setminus \mathbb{I}_1 \cap \{(0,0)\}$.

Теорема 1. *Справедливы следующие соотношения*

$$\begin{aligned} \tilde{\omega}_{i,j}(t) &\equiv \omega_{i,j}(t) \quad \forall (i,j) \in \mathbb{Z}^2 \setminus \mathbb{I}_1 \setminus 2\mathbb{I}_1, \\ \tilde{\omega}_{0,0}(t) &\equiv \omega_{0,0}(t) + \frac{1}{2}\omega_{1,1}(t) + \frac{1}{2}\omega_{-1,1}(t) + \frac{1}{2}\omega_{1,-1}(t) + \frac{1}{2}\omega_{-1,-1}(t), \\ \tilde{\omega}_{2,2}(t) &\equiv \omega_{2,2}(t) + \frac{1}{2}\omega_{1,1}(t), \quad \tilde{\omega}_{-2,2}(t) \equiv \omega_{-2,2}(t) + \frac{1}{2}\omega_{-1,1}(t), \\ \tilde{\omega}_{2,-2}(t) &\equiv \omega_{2,-2}(t) + \frac{1}{2}\omega_{1,-1}(t), \quad \tilde{\omega}_{-2,-2}(t) \equiv \omega_{-2,-2}(t) + \frac{1}{2}\omega_{-1,-1}(t). \end{aligned}$$

Доказательство этих соотношений легко получается, если учесть линейность функций Куранта на треугольниках, содержащихся в их носителе.

В дальнейшем вектор (i, j) будем обозначать через α (впрочем, для краткости иногда скобки в обозначении вектора будем опускать). В соответствии с этим обозначением положим

$$M_\alpha \stackrel{\text{def}}{=} M_{i,j}, \quad \omega_\alpha \stackrel{\text{def}}{=} \omega_{i,j}, \quad \tilde{\omega}_\alpha \stackrel{\text{def}}{=} \tilde{\omega}_{i,j},$$

$$\mathbf{0} \stackrel{\text{def}}{=} (0, 0), \quad \mathbf{e} \stackrel{\text{def}}{=} (1, 1), \quad \mathbf{e}^* \stackrel{\text{def}}{=} (-1, 1).$$

В этих обозначениях имеем

$$\mathbb{I}_1 = \{\mathbf{0}, \mathbf{e}, \mathbf{e}^*, -\mathbf{e}, -\mathbf{e}^*\}, \quad \mathbb{I}'_1 = \mathbb{I}_1 \setminus \{\mathbf{0}\}, \quad 2\mathbb{I}_1 = \{\mathbf{0}, 2\mathbf{e}, 2\mathbf{e}^*, -2\mathbf{e}, -2\mathbf{e}^*\},$$

так что теорема 1 принимает вид

Теорема 2. *Справедливы следующие соотношения*

$$\tilde{\omega}_\alpha(t) \equiv \omega_\alpha(t) \quad \text{при } \alpha \in \mathbb{Z}^2 \setminus \mathbb{I}_1 \setminus 2\mathbb{I}_1, \quad \alpha \in \mathbb{Z}^2,$$

$$\tilde{\omega}_\mathbf{0}(t) \equiv \omega_\mathbf{0}(t) + \frac{1}{2}\omega_{\mathbf{e}}(t) + \frac{1}{2}\omega_{\mathbf{e}^*}(t) + \frac{1}{2}\omega_{-\mathbf{e}}(t) + \frac{1}{2}\omega_{-\mathbf{e}^*}(t),$$

$$\tilde{\omega}_{2\mathbf{e}}(t) \equiv \omega_{2\mathbf{e}}(t) + \frac{1}{2}\omega_{\mathbf{e}}(t), \quad \tilde{\omega}_{-2\mathbf{e}}(t) \equiv \omega_{-2\mathbf{e}}(t) + \frac{1}{2}\omega_{-\mathbf{e}}(t),$$

$$\tilde{\omega}_{2\mathbf{e}^*}(t) \equiv \omega_{2\mathbf{e}^*}(t) + \frac{1}{2}\omega_{\mathbf{e}^*}(t), \quad \tilde{\omega}_{-2\mathbf{e}^*}(t) \equiv \omega_{-2\mathbf{e}^*}(t) + \frac{1}{2}\omega_{-\mathbf{e}^*}(t).$$

Краткая запись этой теоремы имеет вид

Теорема 3. *Справедливы следующие соотношения*

$$\tilde{\omega}_\alpha(t) \equiv \sum_{\gamma \in \mathbb{Z}^2} p_{\alpha,\gamma} \omega_\gamma(t) \quad \forall \alpha \in \mathbb{Z}^2 \setminus \mathbb{I}'_1,$$

где

$$p_{\alpha,\gamma} \stackrel{\text{def}}{=} \delta_{\alpha,\gamma} \quad \text{при } \alpha \in \mathbb{Z}^2 \setminus \mathbb{I}_1 \setminus 2\mathbb{I}_1, \quad \gamma \in \mathbb{Z}^2,$$

$$p_{2\alpha,2\alpha} \stackrel{\text{def}}{=} 1 \quad \text{при } \alpha \in \mathbb{I}_1, \quad p_{\mathbf{0},\alpha} \stackrel{\text{def}}{=} 1/2, \quad p_{2\alpha,\alpha} = 1/2 \quad \text{при } \alpha \in \mathbb{I}'_1;$$

здесь $\delta_{\alpha,\alpha'}$ – символ Кронекера.

Полностью упорядочим (произвольным образом) множество $\{\gamma \mid \gamma \in \mathbb{Z}^2\}$ мультииндексов γ (т.е. упорядочим множество \mathbb{Z}^2), зафиксируем этот порядок и будем использовать его в дальнейшем. Будем считать, что полная упорядоченность во множестве $\mathbb{Z}^2 \setminus \mathbb{I}'_1$ индуцирована упорядоченностью во множестве \mathbb{Z}^2 .

В соответствии с этой упорядоченностью введем вектор-столбцы

$$\tilde{\omega} \stackrel{\text{def}}{=} (\tilde{\omega}_\alpha)_{\alpha \in \mathbb{Z}^2 \setminus \mathbb{I}'_1}, \quad \omega \stackrel{\text{def}}{=} (\omega_\gamma)_{\gamma \in \mathbb{Z}^2},$$

а также рассмотрим матрицу

$$\mathfrak{P} \stackrel{\text{def}}{=} (\mathfrak{p}_{\alpha, \gamma}),$$

где $\alpha \in \mathbb{Z}^2 \setminus \mathbb{I}'_1$ – номер строки, а $\gamma \in \mathbb{Z}^2$ – номер столбца.

Из теоремы 1 вытекает соотношение

$$\tilde{\omega} = \mathfrak{P}\omega.$$

Введем линейные пространства

$$\mathbb{S} \stackrel{\text{def}}{=} Cl_p \mathcal{L}(\{\omega_\gamma \mid \forall \gamma \in \mathbb{Z}^2\}), \quad \tilde{\mathbb{S}}_0 \stackrel{\text{def}}{=} Cl_p \mathcal{L}(\{\tilde{\omega}_\alpha \mid \forall \alpha \in \mathbb{Z}^2 \setminus \mathbb{I}'_1\}),$$

где \mathcal{L} означает линейную оболочку множества элементов, указанного в круглых скобках, а Cl_p – замыкание в топологии поточечной сходимости.

Благодаря теоремам 1–3 получаем $\tilde{\mathbb{S}}_0 \subset \mathbb{S}$.

Литература

1. Сиарле Ф. Метод конечных элементов для эллиптических задач. М., 1980. 512 с.
2. Демьянович Ю.К. Локальная аппроксимация на многообразии и минимальные сплайны. СПб., 1994. 356 с.
3. Азаренок Б.Н. О построении структурированных сеток в двумерных невыпуклых областях с помощью отображений // Журнал вычислительной математики и математической физики, 2009. Т. 49. № 5. С. 1–13.
4. Демьянович Ю.К., Зимин А.В. Аппроксимации курантова типа и их вэйвлетные разложения // Проблемы математического анализа, 2008. Т. 37. С. 25–39.

Рябов А.В., Осьминин Л.А.

Санкт-Петербургский государственный университет
ЗАО «Диджитал Дизайн»

Автоматизация процесса разработки нормативного графика движения поездов

*Рекомендовано к публикации старшим преподавателем
Северюковым С.Ю.*

Введение. График движения поездов, наряду с планом формирования поездов, является одним из основных технологических документов в системе железнодорожного транспорта, объединяющим и согласовывающим работу всех его подразделений. По значимости решаемых задач и их сложности технологический процесс разработки нормативного графика движения поездов стоит на одном из первых мест в ряду важнейших задач планирования и управления работой железнодорожного транспорта [1]. Но, несмотря на возрастающую роль графика движения поездов в современных условиях, технология его разработки и сопровождения принципиально не меняется уже многие годы. В настоящее время от начала разработки нормативной базы до окончания разработки графика проходит около года. В результате лежащие в его основе нормативы устаревают, и график приходится оперативно корректировать. В процессе разработки нормативного графика движения поездов используется несколько автоматизированных систем и программно-технических комплексов: для производства тяговых расчетов, расчета станционных и межпоездных интервалов, формирования листов графика движения поездов и др. Зачастую информация передается между системами вручную, что неизбежно порождает ошибки, а также дублирование данных и неэффективное использование ресурсов вычислительной техники и каналов связи.

Анализ. В ходе исследования был проведен анализ технологии разработки нормативного графика движения поездов, производства тяговых расчетов и расчета других элементов графика с целью выявления связей и последовательностей [2]. Результат этой работы был зафиксирован в виде диаграмм, созданных в среде Microsoft Visio. Одна из них представлена на рис. 1.

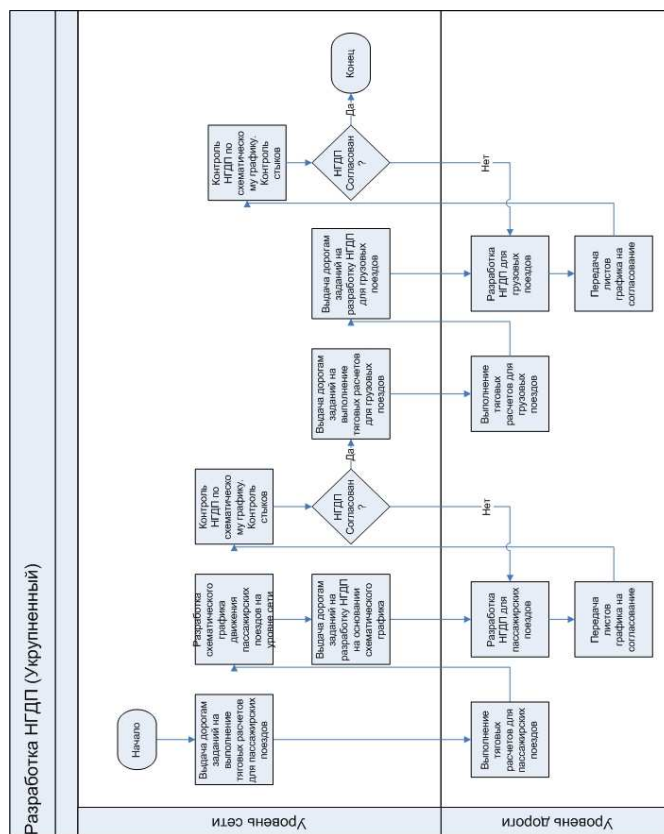


Рис. 1. Диаграмма процесса разработки НГДП

Также был проведен анализ уже существующих систем [3] и выявлены способы и механизмы их взаимодействия. На рис. 2 приведена схема информационных потоков при разработке нормативного графика движения поездов.

Решение. Проведённый анализ выявил ряд проблем, присущих любой попытке автоматизировать и наладить контроль за подобными процессами. В первую очередь, это огромное количество уже действующих в процессе систем, многие из которых не соответствуют новым требованиям и должны быть замещены или модифицированы. При этом замещение должно быть постепенным, не оказывающим существенного влияния на сам процесс. Также подобные

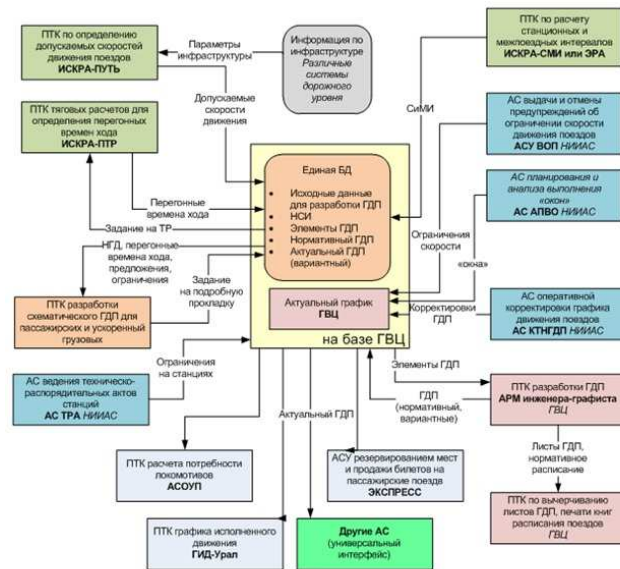


Рис. 2. Схема информационных потоков

системы характеризуются недостаточно точной формализацией требований на этапе проектирования. Это сложно сделать хотя бы из-за огромного количества задействованных в процессе пользователей с принципиально различными задачами и требованиями. Решение должно быть достаточно гибким, чтобы обеспечить быструю модификацию отдельных его частей, не нарушая технологический процесс. Не стоит забывать и о создании единого хранилища данных, совместно используемых различными службами. Локальные модификации общих справочников (в данном случае модификация на дорожном уровне общесетевых справочников) доставляют множество неудобств при объединении и анализе результатов работ отдельных служб и пользователей. Также трудности вызывает развёртывание большого количества клиентских приложений, поскольку для решения локальных проблем, обязательно возникающих вследствие невысокой квалификации персонала в отдалённых районах, приходится посылать инженеров по внедрению в большое количество командировок.

Информация о возможных вариантах решения представлена в открытых источниках довольно скупо. К тому же сложность решений

не позволяет точно сказать, в каком месте была допущена ошибка.

Исходя из вышесказанного, было принято решение строить разработку от процесса контроля. Согласно разработанной архитектуре единое хранилище данных состоит из нескольких логически (а возможно и физически) независимых частей. Взаимодействие пользователей с базой данных (БД) осуществляется через веб-сервисы, для которых будут разработаны соответствующие стандарты. Каждую такую часть можно использовать отдельно, в том числе и в других проектах. Описание процессов и пользовательские представления хранятся отдельно и управляются централизованно. При необходимости предоставить пользователю дополнительно имеющуюся функциональность корректируются правила доступа в ролевой модели. При необходимости разработать новую функциональность правится конкретное пользовательское представление. Уже разработанные пользовательские представления хранятся на сервере в виде готовых dll-сборок. Если изменения требуют модификации БД или создания новых веб-сервисов, то всё это можно сделать, не затрагивая другие части системы. Если пользователю необходимо новое представление, правится схема и пользователю нужно всего лишь перезапустить приложение, что существенно облегчает распространение: достаточно только один раз поставить базовую версию клиентского приложения. Это упрощает реакцию на изменения процесса работы, которые, как показывает практика, происходят достаточно часто, а также позволяет быстро отреагировать на новые требования. Возможности масштабирования ограничены возможностями технической инфраструктуры. Сам процесс не подразумевает обработки большого количества данных в реальном времени, что позволяет использовать пусть и несколько более медленные, но удобные решения. Для замещения какой-либо из устаревших систем, вовлечённых в процесс, достаточно только разработать соответствующие пользовательские представления и связать их с новым решением. Все части системы можно модифицировать отдельно. На рис. 3 приведена высокоуровневая схема архитектуры системы.

Выбранные технологии. Для решения описанной задачи был выбран ряд технологий Microsoft. Ввиду того, что разрабатываемая система будет использоваться на различных функциональных уровнях, число задач пользователя будет существенно отличаться. Было принято решение разработать «тонкий» клиент для пользовате-

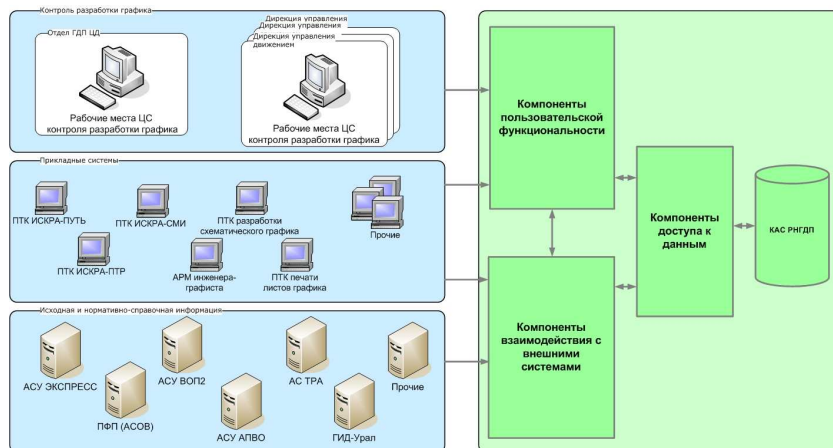


Рис. 3. Высокоуровневое представление архитектуры

лей с небольшим количеством задач и «толстый» для пользователей более высокого уровня, которым требуется выполнять более сложные задачи. Выбор технологии для Web интерфейса стоял между ASP.NET и Silverlight. В итоге был выбран ASP.NET по причине того, что на подавляющем большинстве ПК заказчика установлен Microsoft Internet Explorer 6, при работе с которым Silverlight имеет некоторые проблемы [4], а также имеются и более старые версии браузера, не поддерживающие его. Помимо вышесказанного из-за недостаточного распространения SL он не установлен у большинства пользователей, что может создать дополнительные проблемы с доступом к приложению. Для настольных приложений была выбрана технология Windows Presentation Foundation, так как графической технологией, лежащей в основе WPF, является DirectX, в отличие от Windows Forms, где используется GDI/GDI+. Производительность WPF выше, чем у GDI+ за счёт использования аппаратного ускорения графики через DirectX [5]. В связи с тем, что имеется несколько вариантов уровня представления, было решено вынести уровень бизнес-логики в сервисы. Для их реализации была выбрана технология Windows Communication Foundation, так как WCF делает возможным построение безопасных и надёжных транзакционных систем через упрощённую унифицированную программную модель межплатформенного взаимодействия. Комбинируя функциональность существующих технологий .NET по разработке распределённых приложений, WCF предоставляет единую инфраструктуру

разработки, повышающую производительность и снижающую затраты на создание безопасных, надёжных и транзакционных Web-служб нового поколения. Заложенные в неё принципы интероперабельности позволяют легко добиваться взаимодействия с другими платформами [6].

Вывод. Разработанная архитектура системы решает ряд проблем, общих для всех задач подобного рода, что позволяет успешно использовать подобный подход и в других отраслях.

Литература

1. Кудрявцев В.А., Бадах В.И., Белов К.А.. Организация и управление движением на железнодорожном транспорте. М.: Издат. центр «Академия», 2006. 432 с.
2. Инструкция по разработке графика движения поездов в ОАО «РЖД».
3. Гапанович В.А., Грачев А.А. Системы автоматизации и информационные технологии управления перевозками на железных дорогах / Под редакцией В.И. Ковалева, А.Т. Осьминина, Г.М. Грошева М.: Маршрут, 2006. 544с.
4. <http://stackoverflow.com/questions/672889/any-disadvantages-of-silverlight>
5. Мак-Дональд М. WPF: Windows Presentation Foundation в .NET 3.5 с примерами на C# для профессионалов. 2-е издание: Пер. с англ. М.: ООО «Вильямс», 2008. 928 с.
6. Резник С., Крейн Р., Боуэн К. Основы Windows Communication Foundation для .NET Framework 3.5 / Пер. с англ. А.А. Слинкина М.: ДМК Пресс, 2008. 480 с.

Сафонова А.В.

Санкт-Петербургский государственный университет

Метод машинного обучения с использованием LARS и Lasso в задаче ранжирования поисковой выдачи

Рекомендовано к публикации доцентом Гришкиным В.М.

Цель работы. Цель работы – оценить применимость методов линейной регрессии в задаче ранжирования документов и с их помощью получить формулу ранжирования. Для обучения и проверки используются реальные данные: значения признаков пар «запрос-документ» и оценки релевантности, сделанные ассессорами Яндекса. Данные взяты из открытого источника – задания конкурса Интернет-математика – 2009 [1].

Общая постановка задачи. В данной работе мы имеем дело с характерным примером задачи машинного обучения – обучения по прецедентам или обучения с учителем. Имеется множество объектов и множество возможных откликов или ответов. Существует некоторая неизвестная зависимость между ответами и объектами. Известна только конечная совокупность прецедентов – пар «объект, ответ», называемая обучающей выборкой. На основе этих данных требуется восстановить зависимость, т.е. построить алгоритм, способный для любого нового объекта выдать достаточно точный ответ.

Формализация. Пусть X – множество описаний объектов, Y – множество допустимых ответов. Существует неизвестная целевая зависимость – отображение $y^* : X \rightarrow Y$, значения которой известны только на объектах конечной обучающей выборки

$$X^m = \{(x_1, y_1), \dots, (x_m, y_m)\}.$$

Требуется построить алгоритм $\alpha : X \rightarrow Y$, который приближал бы неизвестную целевую зависимость как на элементах выборки, так и на всём множестве X .

Для оценки качества алгоритма вводится функция потерь $\mathcal{L}(y, y')$, характеризующая величину отклонения ответа $y = a(x)$ от правильного ответа $y' = y^*(x)$ на произвольном объекте $x \in X$. В задачах регрессии это – сумма квадратов невязок $\mathcal{L}(y, y') = (y' - y)^2$.

К сожалению, малое значение функции на обучающей выборке не гарантирует того, что построенный алгоритм будет хорошо восстанавливать целевую зависимость на всём пространстве X . Существует опасность возникновения переобучения – ситуации, когда алгоритм делает попытку описать конкретные данные точнее, чем позволяют уровень шума в данных и погрешность самой модели. Также значительно ухудшает качество прогнозирования значения зависимой переменной то, что начальная обучающая выборка может быть значительно смещена относительно генеральной совокупности возможных значений входных параметров.

Общий метод решения. В данной работе в качестве аппроксимации неизвестной функции используется модель линейной регрессии.

Регрессионная модель есть функция независимой переменной и параметров с добавленной случайной переменной. Параметры модели настраиваются таким образом, что модель наилучшим образом приближает данные. Критерием качества приближения обычно является среднеквадратичная ошибка.

Для построения регрессионной модели здесь используются метод наименьших углов – LARS (Least angle regression) – и его модификация в одном частном случае – алгоритм Lasso (Least absolute shrinkage and selection operator).

LARS. При большом количестве свободных переменных в регрессионной модели возникает проблема неустойчивого оценивания весов модели. LARS предлагает метод выбора такого набора свободных переменных, который имел бы наиболее значимую статистическую связь с зависимой переменной, и метод оценки их весов.

Отличие данного алгоритма от обычной пошаговой регрессии в том, что алгоритм LARS, вместо последовательного добавления свободных переменных, на каждом шаге изменяет их веса. Веса увеличиваются так, чтобы доставить наибольшую корреляцию с вектором регрессионных остатков. Основным достоинством LARS является то, что он выполняется за число шагов, не превышающее число свободных переменных.

Постановка задачи. Задана выборка – матрица X , столбцы которой соответствуют независимым переменным, а строки – элементам выборки, и вектор y , содержащий элементы зависимой пере-

менной. Назначена линейная модель $y = X\beta + \varepsilon$. Обозначим множество столбцов матрицы X как $\{x_1, \dots, x_n\}$. Пусть выполнены условия нормировки

$$|y_1| = 0, \quad |x_j|_1 = 0, \quad |x_j|_2^2 = 1, \quad j = \overline{1, n}.$$

Предполагается, что векторы x_1, \dots, x_n линейно независимы.

Критерием качества модели назначена среднеквадратичная ошибка

$$S(\beta) = |y - \mu|^2,$$

где $\mu = X\beta$.

Требуется найти для каждого шага LARS вектор весов β и набор столбцов, доставляющий наибольшую корреляцию векторов μ и y .

Описание метода. LARS последовательными шагами строит оценку коэффициентов $\hat{\beta}$. На k -ом шаге только k элементов вектора $\hat{\beta}$ отличны от нуля. Алгоритм последовательно вычисляет приближение вектора значений зависимой переменной μ . Для приближений используется вектор корреляций столбцов матрицы X с вектором остатков $y - \mu$

$$c(\mu) = X^T(y - \mu).$$

На k -ом шаге новое значение приближения вектора значений зависимой переменной y вычисляется как

$$\hat{\mu}_k = \hat{\mu}_{k-1} + \gamma_k u_k.$$

Здесь u_k – нормированный вектор, задающий биссектрису между добавляемым вектором – столбцом матрицы X – и вектором регрессионных остатков, который вычисляется следующим образом.

Пусть \mathcal{A} – подмножество индексов $1, \dots, n$ столбцов матрицы X . Это подмножество задает подматрицу

$$X_{\mathcal{A}} = [s_{j_1} x_{j_1}, \dots, s_{j_{|\mathcal{A}|}} x_{j_{|\mathcal{A}|}}], \quad j \in \mathcal{A},$$

где множитель $s \in \{-1, +1\}$ и $|\mathcal{A}|$ – мощность множества \mathcal{A} . Обозначим ковариационную матрицу

$$\mathcal{G} = X_{\mathcal{A}}^T X_{\mathcal{A}} \quad \text{и} \quad A_{\mathcal{A}} = (\mathbf{1}_{\mathcal{A}}^T \mathcal{G}_{\mathcal{A}}^{-1} \mathbf{1}_{\mathcal{A}})^{-\frac{1}{2}},$$

где $\mathbf{1}_{\mathcal{A}}$ – вектор, состоящий из $|\mathcal{A}|$ единиц.

Вычислим единичный вектор

$$u_{\mathcal{A}} = X_{\mathcal{A}} w_{\mathcal{A}}, \quad w_{\mathcal{A}} = A_{\mathcal{A}} \mathcal{G}_{\mathcal{A}}^{-1} \mathbf{1}_{\mathcal{A}}.$$

Вектор u образует со столбцами матрицы $X_{\mathcal{A}}$ одинаковые углы, меньшие $\pi/2$. Описание алгоритма представлено в работе [2].

Lasso. Данный метод оценивания коэффициентов линейной регрессионной модели заключается во введении ограничения на норму вектора коэффициентов модели, что приводит к обращению в 0 некоторых из них. Метод приводит к повышению устойчивости модели в случае большого числа обусловленности матрицы признаков X , позволяет получить интерпретируемые модели – отбираются признаки, оказывающие наибольшее влияние на вектор ответов.

Постановка задачи аналогична таковой для LARS.

Описание метода. Результатом минимизации среднеквадратичной ошибки по β методом наименьших квадратов является вектор $\hat{\beta}$. Пусть $T(\hat{\beta})$ – сумма модулей регрессионных коэффициентов

$$T(\hat{\beta}) = \sum_{j=1}^m |\hat{\beta}_j|.$$

В Lasso $\hat{\beta}$ выбирается из условия минимизации $S(\beta)$ при ограничении

$$T(\hat{\beta}) \leq t,$$

где $t \geq 0$ – параметр регуляризации [3]. За счёт него Lasso обладает большей обобщающей способностью как алгоритм машинного обучения, чем LARS. Подбор параметра регуляризации является отдельной задачей, выходящей за рамки рассмотрения данной работы.

Чем меньше t , тем большее число коэффициентов $\hat{\beta}_i$ принимает нулевое значение. Таким образом, Lasso осуществляет отбор информативных признаков. Задачу можно решать, последовательно вводя ограничения-неравенства.

В данном случае используется модификация стандартного алгоритма [3], в котором на этапе добавления в набор новой активной переменной происходит также удаление из него тех элементов, для которых ограничения-неравенства не выполняются [4].

Оценка качества полученной модели. Таким образом, вышеуказанными методами получаем регрессионную модель, которая

позволяет оценить релевантность пар документ-запрос из тестовой выборки. Спрогнозированные релевантности документов ранжируются внутри каждого запроса по убыванию. По ранжированному списку на основании оценок ассессоров вычисляется метрика качества. Метрикой качества ранжирования здесь является взвешенная совокупная выгода (Discounted cumulative gain – DCG), усредненная по всем запросам. Используется модифицированная формула для вычисления DCG:

$$DCG = \sum_i \frac{rel_i}{\log_2 i + 1}.$$

Практические результаты. В качестве входных обучающих данных был взят файл `imat2009_learning.txt` размером 97280 строк с оценками ассессоров [1]. Он был поделен на обучающую (50003 строк) и тестовую (47287 строк) выборки. Количество признаков для каждой пары (документ-запрос) равняется 245. Реализация алгоритмов LARS и Lasso выполнена в среде Matlab (основана на [4]).

При помощи алгоритмов LARS и Lasso из рассмотрения были исключены 10 признаков как незначащие или сильно коррелирующие с признаками из активного набора.

LARS строит полную регрессию за 235 шагов, равное количеству значимых переменных. Lasso для построения регрессии потребовалось 286 шагов. Значение среднеквадратичного отклонения на тестовой выборке 186,5033.

На тестовых данных минимальная сумма квадратов отклонений предсказанных оценок от оценок ассессоров для Lasso достигается на 178-ом шаге при 156 значащих переменных (174,8731), для LARS – на 167-ом шаге соответственно для 166 значащих переменных (174,8304).

Значение метрики ранжирования DCG на известной тестовой выборке (исключены группы документов, относящихся к запросу с нулевой суммой всех оценок ассессоров) равно 4,226228 на максимальном наборе значащих переменных, 4,224744 – на значащих 156 для Lasso, и 4,217906 – для LARS на 166 переменных.

Базовое значение метрики (baseline) для обучающего алгоритма, обозначенное при предоставлении данных, равнялось 4,041899, лучший результат, полученный в рамках олимпиады – 4,284426 (результаты приведены исключительно для приблизительного сравнения, так как обучение и тестирование в рамках конкурса проводились на выборках большего размера [1]).

Выводы. Модель линейной регрессии является незаменимым и достаточно мощным инструментом для построения искомого приближения неизвестной зависимости между вектором входных переменных и результирующей переменной. Однако, линейная зависимость не может описать всю сложность ранжирующего алгоритма, особенно в условиях ограниченности обучающей выборки. Тем не менее, LARS и Lasso позволяют значительно сократить размерность модели, выделить значащие признаки, построить регрессию на меньшем количестве входных параметров без значительных потерь в точности оценки выходной переменной. По полученным практическим результатам, Lasso является более гибким алгоритмом при решении задачи ранжирования, чем LARS, хотя требует для своего выполнения большего числа итераций.

Развитием регрессионного подхода в задаче аппроксимации целевой функции служат полиномиальная, нелинейная, непараметрическая, символьная регрессия, а также комбинация регрессионных методов со статистическими методами искусственного увеличения размеров обучающего множества и суперпозиции получаемых решающих моделей.

Литература

1. Яндекс. Интернет-математика. Конкурс 2009.
<http://imat2009.yandex.ru/datasets>
2. Efron B., Hastie T., Johnstone J., Tibshirani R. Least Angle Regression // *The Annals of Statistics*, 2004. Vol. 32, № 2. P. 407–499.
3. Tibshirani R. Regression shrinkage and Selection via the Lasso // *Journal of the Royal Statistical Society. Series B (Methodological)*, 1996. Vol. 32. № 1. P. 267–288.
4. Исходные коды Lasso для MatLab. Created by Guilherme Rocha, Peng Zhao and Bin Yu, Department of Statistics, University of California, Berkeley.
<https://mvr.svn.sourceforge.net/svnroot/mvr/mdlselection/lasso>
5. MachineLearning.ru. <http://www.machinelearning.ru/wiki/>

Севостьянов Р.А.

Санкт-Петербургский государственный университет

Программное обеспечение для управления движением робота Futaba RBT-1

Рекомендовано к публикации профессором Веремеем Е.И.

Робот Futaba RBT-1 обладает 20-ю сервоприводами, управляемыми процессорным модулем RPU-11. Управление может осуществляться в двух режимах: режим загруженных демонстрационных движений и режим непосредственного управления от компьютера через последовательный порт. Для второго режима фирмой производителем разработано программное обеспечение, позволяющее синхронизировать робот с компьютером и посылать управляющие команды с помощью интерактивного графического интерфейса [1].

Основная цель изучения и модификации программного обеспечения состоит в обучении робота зрению, т.е. умению различать во внешнем окружении объекты по заданному шаблону и двигаться к ним, обходя препятствия. Для достижения этой цели необходим набор функций, реализованных на языке высокого уровня, с помощью которых можно было бы легко формировать программы управления движением роботом в зависимости от того, что он видит.

Следует отметить, что программное обеспечение, поставляемое производителем в комплекте с роботом, для этого не подходит, так как предусматривает взаимодействие с помощью интерактивного графического интерфейса. Никакие средства для программной реализации алгоритмов управления на языке высокого уровня не поставляются.

В силу отмеченного обстоятельства было решено реализовать собственную библиотеку функций управления роботом для дальнейшего использования в программах непосредственной обработки зрительных сигналов.

Как отмечено выше, обмен данными между компьютером и роботом происходит через последовательный порт. Компьютер отправляет пакеты байтов определенной структуры, процессор робота обрабатывает их и посылает сигналы на сервоприводы, осуществляя движение. Существует три типа пакетов: короткие, длинные и пакеты обратной связи [2]. Короткие пакеты задают изменение положения

отдельного сервопривода, длинные – сразу нескольких (возможно, всех сразу), а пакеты обратной связи служат для отправки данных с робота на компьютер.

Интерес, в основном, представляют длинные пакеты, так как необходимо задавать комплексные движения робота такие, как шаги или повороты, которые требуют изменения состояния одновременно нескольких сервоприводов. Рассмотрим структуру такого пакета, представленную на рис. 1.



Рис. 1. Структура длинного пакета управления сервоприводами

- **Начало** – байт, обозначающий начало пакета данных. Здесь всегда записывается «53h» [3], где h – обозначение шестнадцатеричного числа.
- **Количество байт** – байт, в котором указано, сколько байт данных содержится в пакете [3].
- **Заголовок** – два байта, обозначающих начало сообщения, содержащего информацию для приводов, где всегда записано значение «FAh AFh».
- **ID** – здесь всегда записывается число «00h».
- **Флаги** – здесь всегда записывается число «00h».
- **Адрес** – адрес карты памяти приводов, куда будут записаны данные. Обычно здесь всегда находится значение «1Eh».
- **Длина** – сюда записывается количество байт данных для каждого сервопривода (количество байт, содержащихся в секции «Данные» пакета). Длина = количество байт + 1. Как правило, в этой секции записано значение «05h», причина будет объяснена ниже.
- **Количество** – здесь содержится количество сервоприводов, положение которых необходимо изменить. Обычно посылаются данные на все 20 сервоприводов, даже если их положение не подлежит изменению.
- **VID** – представляет ID сервопривода, на который будет послана информация, содержащаяся в секции «Данные».

- **Данные** – обычно здесь находятся 4 байта информации для сервопривода, ID которого указано в предыдущей секции «VID». Первые два байта содержат информацию о том, в какое положение нужно перевести данный привод (имеется в виду угол поворота). Минимальная единица поворота – 1 градус, что соответствует числу «0Ah» или двум байтам «0Ah 00h», если речь идет непосредственно о пакете.

Стоит отметить, что здесь указывается конечный угол, в который нужно перевести сервопривод, а не количество градусов, на которые осуществляется поворот из данного положения. По умолчанию при включении робота все сервоприводы приводятся в положение «00h 00h».

Следующие два байта задают время, за которое мы хотим перевести данный сервопривод в указанное положение. Минимальная единица времени – примерно 100 мс, что также соответствует числу «0Ah» или двум байтам «0Ah 00h».

- **Сумма** – в эту секцию записывается контрольная сумма пакета. Сумма вычисляется как «исключающее или» (XOR) всех байтов пакета, начиная с ID и до последнего байта последней секции «Данные».

Рассмотрим пример пакета (см. рис. 2), поворачивающего за 400 мс сервопривод с ID 2 на 30 градусов по часовой стрелке, а сервопривод с ID 12 – на 45 градусов против часовой стрелки.

Нач.	к.б.	заголовок	ID	Флаги	Адр	Длн	кол.	VID	Данные	VID	Данные	к.с.																											
	53		12		FA	AF		00		00		1E		05		02		01		2D		01		28		00		0C		3D		FE		28		00		ED	

Рис. 2. Пример длинного пакета данных

Программное обеспечение производителя для сохранения и дальнейшего использования движений использует три типа файлов: файлы позиций, файлы движений и сценарные файлы. Будем использовать именно такие файлы, формируемые с помощью фирменного программного обеспечение, так как это предоставляет большую гибкость и удобство: можно задать и проанализировать, либо изменить любое движение с помощью удобного графического интерфейса, и далее без проблем использовать полученные файлы позиций и движений в качестве исходных для собственной библиотеки. Рассмотрим каждый тип файлов в отдельности.

Файлы позиций задают статические положения робота в какой-то момент времени. В данных файлах содержатся фактические значения углов поворота сервоприводов.

Файлы движений задают единичные движения робота такие, как шаг, поворот и т.д. Движение – это, по сути, последовательность позиций. Таким образом, файл движений содержит имена файлов позиций, из которых состоит данное движение, и моменты времени, в которые робот должен оказаться в соответствующих позициях.

Сценарные файлы содержат информацию о сложных движениях робота, т.е. последовательности единичных движений.

В собственной библиотеке на данный момент предлагается использовать только файлы позиций и движений, так как это предоставляет наибольшую гибкость в управлении роботом, хотя в дальнейшем возможно использование и сценарных файлов, если выяснится, что определенная последовательность движений используется достаточно часто.

Реализация данной библиотеки на языке C++ достаточно проста. Предлагается использовать два базовых класса – класс Position, задающий статические положения, и класс Motion, задающий движения как последовательности позиций. В классе Position определены два метода: void GetPos() и void GetPacket(). Первый считывает из файла позиций информацию о положениях каждого сервопривода и создает массив с координатами положений в десятичной системе, сохраняя его в поле int posit[20]. Метод void GetPacket() переводит координаты из десятичной системы в шестнадцатеричную и формирует массив, представляющий длинный пакет данных для процессора робота, сохраняя его в поле char lpacket[110].

Класс Motion содержит методы для считывания и осуществления движений, определенных в произвольных файлах позиций, а также несколько методов, осуществляющих конкретные, наиболее часто используемые движения. Метод void GetMtn(string *имя_файла*), который считывает файл движений, задает массив данных типа Position (т.е., по сути, пакеты данных для робота) и сохраняет его в поле Position *pos. Метод SendMtn() отсылает эти пакеты, если они уже сформированы. Таким образом, можно один раз считать данные из файла движений и далее пользоваться методом SendMtn() для осуществления этого движения. Кроме того, введены несколько методов, считывающих и реализующих конкретные движения: void StepForward(), void StepBack(), void StepRight(), void StepLeft(), void

TurnRight() и void TurnLeft(). Первые четыре метода осуществляют, соответственно, шаг вперед, шаг назад, шаг влево и шаг вправо, а два оставшихся – соответственно, поворот вправо и поворот влево примерно на 20 градусов.

Для того, чтобы осуществить движение, прежде всего, нужно открыть последовательный порт, к которому подключен робот. Для работы с последовательными портами используется open source библиотека «Serial Gate» [4]. Работа с ней достаточно проста. Создается объект типа SerialGate и с помощью метода Open(int номер_порта, int скорость) открывается порт с номером номер_порта. Скорость, как правило, равна 115200. Далее создается объект типа Motion с заданием в качестве параметра ссылки на объект, определяющего последовательный порт, с которым идет работа. После этого для данного объекта типа Motion можно просто вызывать методы, осуществляющие движения, при этом с помощью методов GetMtn(string имя_файла) и SendMtn() можно осуществить любое произвольное движение, описанное в файле имя_файла.

Таким образом, по сути, используются две библиотеки: библиотека SerialGate.dll для работы с последовательным портом и собственная библиотека Rbt1Control.dll для осуществления управления роботом.

Рассмотрим пример исходного кода программы, осуществляющей четыре шага вперед и поворот влево приблизительно на 40 градусов.

```
//Подключаем заголовочный файл
//библиотеки управления роботом:

#include "Rbt1Control.h"

//Подключаем заголовочный файл
//библиотеки для работы
//с последовательным портом:

#include "SerialGate.h"

//Сообщения для компилятора для корректного
//Подключения библиотек:
#pragma comment(lib, "Rbt1Control.lib")
#pragma comment(lib, "SerialGate.lib")
```

```

int main()
{
//Создаем новый объект типа SerialGate
//и открываем порт с номером 1
//на скорости 115200 байт/сек
SerialGate sg;
sg.Open(1,115200);

//Создаем объект типа Motion и передаем
//ему ссылку на объект, задающий
//последовательный порт:
Motion mtn(&sg);

for (int i=0; i<4; i++)
mtn.StepForward(); // 4 шага вперед;

mtn.TurnLeft();
mtn.TurnLeft(); // Два поворота влево.
}

```

Приведенный код протестирован на роботе, что подтверждает правильность принятого подхода.

Литература

1. RBT-1 Operation Manual. Futaba Corporation official documentation, 2007.
2. RS301CR/302CD User's Manual. Futaba Corporation official documentation.
3. RPU-10 Info. Futaba Corporation official documentation, 2009.
4. Сопряжение компьютеров с внешними устройствами.
<http://www.pcports.ru>

Смирнов М.Н.

Санкт-Петербургский государственный университет

**Алгоритм синтеза управлений, подавляющих
ограниченные внешние воздействия
на морское судно**

Рекомендовано к публикации профессором Веремеем Е.И.

1. Введение. Уровень развития современных компьютерных технологий позволяет устанавливать на современные морские суда системы автоматического управления движением, тем самым облегчая и делая более безопасным для человека выход в открытое море. Такие системы позволяют снизить возможность возникновения аварий из-за человеческого фактора, более точно следовать по заданным маршрутам и огибать различные препятствия с учетом особенностей корабля и действующих на него внешних возмущений.

В связи с этим возникает ряд задач, связанных с построением систем автоматического управления движением морского судна, а именно, минимизация времени совершения маневра, построение оптимальных траекторий движения, подавление различных видов внешних возмущений таких, как ветер и волнение моря. В большинстве задач, рассматриваемых в литературе [1–4], возмущения либо отсутствуют, либо задаются в конкретном виде, в частности, неограниченно убывают с течением времени.

В данной работе рассматривается задача о компенсации (подавлении) ограниченного внешнего возмущения с учетом требований к степени устойчивости. В такой ситуации необходимо построить закон управления, который давал бы наилучший возможный результат по качеству динамики для наилучшего варианта ограниченного возмущения.

2. Постановка задачи. Рассмотрим математическую модель морского судна [1]:

$$\begin{aligned}\dot{\beta} &= a_{11}\beta + a_{12}\omega + b_1\delta + h_1d(t), \\ \dot{\omega} &= a_{21}\beta + a_{22}\omega + b_2\delta + h_2d(t), \\ \dot{\varphi} &= \omega, \\ \dot{\delta} &= u.\end{aligned}\tag{1}$$

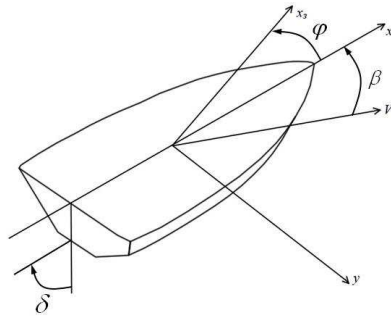


Рис. 1. Динамические переменные определяемое ветром и волнением моря:

На рис. 1 представлены основные динамические переменные. Здесь ω – угловая скорость относительно вертикальной оси, φ – курс (положительным считаем поворот на левый борт), δ – угол отклонения вертикальных рулей, β – угол дрейфа (угол между вектором скорости и продольной осью судна), u – управление, $d(t)$ – внешнее возмущение, ветром и волнением моря:

$$d^T(t)d(t) \leq 1, \quad 0 \leq t < \infty.$$

В качестве компенсирующего управления будем использовать регулятор в форме статической обратной связи по состоянию, представленный уравнением:

$$u = k_1\beta + k_2\omega + k_3\varphi + k_4\delta, \quad (2)$$

где $K = (k_1, k_2, k_3, k_4)$ – коэффициенты, подлежащие выбору в процессе решения задачи синтеза.

На рули и скорость их поворота накладываются ограничения: $|\delta| \leq 30^\circ, |u| \leq 3^\circ/c$.

Для построения данного регулятора воспользуемся подходом, связанным с оценками множества достижимости, т.е. множества всех таких состояний, в которые можно привести систему при помощи допустимого управления.

Введём новые обозначения. Пусть $x = (\beta, \omega, \varphi, \delta)^T$ – вектор состояния, $y(t)$ – выход системы. Тогда систему (1) можно переписать в виде

$$\begin{aligned} \dot{x} &= Ax + Bu + Hd(t), \\ y &= Cx, \end{aligned}$$

где

$$A = \begin{pmatrix} a_{11} & a_{12} & 0 & b_1 \\ a_{21} & a_{22} & 0 & b_2 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix},$$

$$C = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad H = \begin{pmatrix} h_1 \\ h_2 \\ 0 \\ 0 \end{pmatrix}.$$

Замыкая систему (1) управлением (2), получим

$$\begin{aligned} \dot{x} &= A_c x + H_c d(t), \\ y &= Cx, \end{aligned} \quad (3)$$

где $A_c = A + BKC = A + BK$, $H_c = H$.

Рассматриваемый метод решения задачи основан на минимизации оценки l_1 -нормы выхода системы управления. Суть его состоит в нахождении минимального эллипсоида (в смысле следа соответствующей матрицы), в котором содержатся выходы $y(t)$ при любых ограниченных возмущениях $d(t)$.

3. Алгоритм синтеза компенсирующих управлений. Воспользуемся утверждением, доказанным в работе [5].

Пусть матрица A устойчива, пара (A, B) – управляема, а C – матрица полного ранга. Тогда множество достижимых выходов $y(t)$ в системе

$$\begin{aligned} \dot{x} &= Ax + Hd(t), \quad x(0) = 0, \\ y &= Cx, \quad d^T(t)d(t) \leq 1, \quad 0 \leq t < \infty \end{aligned}$$

содержится в эллипсоиде $\varepsilon = \{y : y^T (CPC^T)^{-1}y \leq 1\}$, где P является решением матричного уравнения Ляпунова

$$AP + PA^T + \alpha P + \alpha^{-1}HH^T = 0, \quad \alpha > 0.$$

При этом матрица $P(\alpha)$ должна быть положительно определена. Кроме того, решая однопараметрическую задачу минимизации

$$\min_{\alpha > 0} tr(CP(\alpha)C^T),$$

получим эллипсоид, обладающий минимальным следом среди всех эллипсоидов, содержащих достижимое множество по выходу.

Используя сказанное выше, задачу построения минимального эллипсоида для системы (3) можно сформулировать в виде

$$\begin{aligned} AP + PA^T - \gamma BB^T + \alpha P + \alpha^{-1}HH^T &= 0, \\ \min_{\alpha, \gamma > 0} tr(CP(\alpha, \gamma)C^T). \end{aligned} \quad (4)$$

Требование положительной определенности матрицы $P(\alpha, \gamma)$ сохраняется. Решая задачу (4), получим регулятор в виде $u = -\gamma B^T P^{-1}x$, который обеспечит выполнение неравенства $y^T(CPC^T)^{-1}y \leq 1$ для выхода $y(t)$ системы (3) при любом $\|d\|_\infty \leq 1$.

Для обеспечения требуемой степени устойчивости характеристического полинома системы (3) воспользуемся результатом, доказанным в [6], который указывает способ построения полинома с наперед заданной степенью устойчивости.

Данный подход позволяет найти управление, обеспечивающее не только подавление ограниченного внешнего возмущения, но и необходимую степень устойчивости характеристического полинома замкнутой системы. Окончательный алгоритм состоит в следующем:

1. Следуя методу компенсации внешних воздействий [5], выражаем коэффициенты регулятора k_1, k_2, k_3, k_4 через параметры α, γ .
2. Строим характеристический полином замкнутой системы, зависящий от α и γ .
3. Задаем требуемую степень устойчивости и строим вспомогательный полином с наперед заданными корнями.
4. Приравнявая коэффициенты при соответствующих степенях в характеристическом и вспомогательном полиномах, получаем систему уравнений, решением которой является некоторое множество параметров α и γ .
5. Решаем задачу минимизации (4) на множестве параметров, найденном на предыдущем шаге.

4. Компьютерная модель системы управления движением. Для реализации предложенного алгоритма и моделирования динамических процессов в замкнутой системе удобно использовать пакет MATLAB с подсистемой Simulink. Воспользуемся разработанным алгоритмом для построения системы управления движением морского судна водоизмещением 6000 т. При фиксированной скорости хода коэффициенты в модели (1) принимают следующие значения: $a_{11} = -0,03408$, $a_{12} = 0,56$, $a_{21} = 0,015$, $a_{22} = -0,306$, $b_1 = -0,0099$, $b_2 = -0,00417$, $h_1 = -0,0648$, $h_2 = -0,0046$. Для проведения имитационного моделирования построена Simulink-модель замкнутой системы управления движением судна.

5. Результаты синтеза и моделирования динамики. В результате применения описанного алгоритма к объекту управления с параметрами из пункта 4 получен закон управления с коэффициентами $k_1 = 49,5187$, $k_2 = -108,2410$, $k_3 = 12,3767$, $k_4 = -21,6532$, обеспечивающий желаемую степень устойчивости и динамику замкнутой системы. При этом учитываются технические ограничения на отклонение и скорость поворота рулей. Для проверки качества найденного управления воспользуемся им для автоматического управления морским судном при воздействии ограниченного внешнего возмущения, график которого представлен на рис. 2.

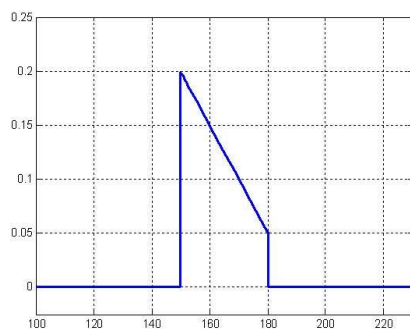


Рис. 2. Внешнее возмущение

На рис. 3 изображены графики изменения курса и отклонения вертикальных рулей.

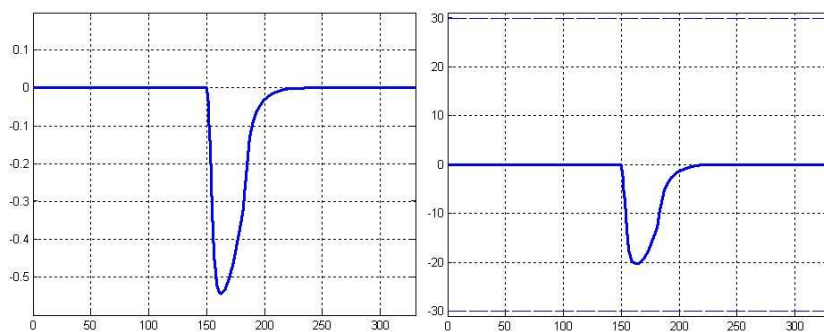


Рис. 3. Изменение курса и отклонение вертикальных рулей

6. Заключение. В работе разработан алгоритм синтеза управлений, компенсирующих ограниченные внешние воздействия, основанный на использовании метода инвариантных эллипсоидов с обеспечением желаемой степени устойчивости. Алгоритм реализован в интегрированной среде MATLAB, построена Simulink-модель системы управления движением. Результаты применения данного алгоритма проиллюстрированы на примере морского судна.

Литература

1. Веремей Е.И., Корчанов В.М., Коровкин М.В., Погожев С.В. Компьютерное моделирование систем управления движением морских подвижных объектов. СПб.: СПбГУ, 2002. 370 с.
2. Веремей Е.И. Синтез законов многоцелевого управления движением морских объектов // Гироскопия и навигация, 2009. Вып. 4. С. 3–14.
3. Лукомский Ю.А., Корчанов В.М. Управление морскими подвижными объектами. СПб.: Элмор, 1996. 320 с.
4. Fossen T.I. Guidance and control of ocean vehicles. John Wiley and Sons. New York, 1999. 494 p.
5. Поляк Б.Т., Щербаков П.С. Робастная устойчивость и управление. М.: Наука, 2002. 303 с.
6. Веремей Е.И. Обеспечение заданной степени устойчивости регуляторами с неполной информацией // Изв. АН СССР. Техническая кибернетика, 1986. Вып. 4. С. 123–130.

Федорова М.А.

Санкт-Петербургский государственный университет

Синтез и компьютерное моделирование астатической системы управления курсом морского судна

Рекомендовано к публикации профессором Веремеем Е.И.

1. Введение. Система автоматического управления курсом судна является важнейшей системой судовой автоматики, от качества работы которой напрямую зависит экономическая эффективность, надежность и безопасность движения морского судна [1, 2]. Подобные системы в линейном приближении должны быть устойчивыми, однако устойчивость является обязательным, но далеко не единственным требованием, предъявляемым к ним.

Важно, чтобы система управления обладала свойством астатизма по регулируемой координате, т.е. способностью приводить ошибку регулирования к нулю при наличии постоянного внешнего воздействия.

Проектирование и построение таких систем является нетривиальной задачей, существенно зависящей от типа судна, его особенностей и назначения.

2. Постановка задачи. Рассмотрим линейную математическую модель бокового движения морского судна [3]:

$$\begin{aligned}\dot{\beta} &= a_{11}\beta + a_{12}\omega + b_1\delta + F, \\ \dot{\omega} &= a_{21}\beta + a_{22}\omega + b_2\delta + M, \\ \dot{\varphi} &= \omega, \\ \dot{\delta} &= u.\end{aligned}\tag{1}$$

На рис. 1 представлены основные динамические переменные. Здесь ω – угловая скорость относительно вертикальной оси, φ – курс (положительным считаем поворот на левый борт), δ – угол отклонения вертикальных рулей, β – угол дрейфа (угол между вектором скорости и продольной осью судна), F – боковая сила, M – момент, u – управление.

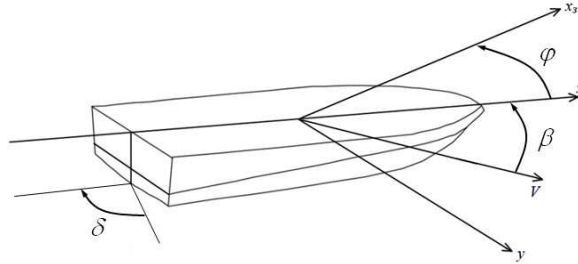


Рис. 1. Динамические переменные

В качестве управления будем использовать регулятор в форме статической обратной связи по состоянию с математической моделью

$$u = k_1\beta + k_2\omega + k_3\varphi + k_4\delta, \quad (2)$$

где $k = (k_1, k_2, k_3, k_4)$ – коэффициенты, подлежащие выбору.

На отклонение рулей и управление налагаются следующие технические ограничения: $|\delta| \leq 30^\circ$, $|u| \leq 3^\circ/c$. Требуется построить систему автоматического управления движением морского судна, обеспечивающую астатизм замкнутой системы по курсу.

3. Алгоритм синтеза астатических законов управления.

Введем понятие астатизма по выходу.

Определение. Линейная стационарная система со входом v и выходом w называется астатической по выходу, если для ступенчатого входного сигнала $v(t) \equiv v_0 \cdot 1(t)$ при любом вещественном v_0 выполняется условие $w_0 = 0$, где w_0 – значение выхода, соответствующее положению равновесия системы с постоянным входным сигналом v_0 .

При обеспечении астатизма замкнутой системы по курсу необходимо учитывать, что коэффициенты регулятора должны также обеспечивать требуемую степень устойчивости. Алгоритм состоит в следующем:

1. Введём функционал

$$J(u) = \int_0^\infty ((\rho_1 + \gamma_1 10^{-4})\omega^2 + (\rho_2 + \gamma_2 10^{-4})\varphi^2 + \rho_3 u^2) dt, \quad (3)$$

где ρ_1, ρ_2, ρ_3 – заданные положительные параметры, γ_1, γ_2 – положительные параметры, которые можно варьировать.

2. Воспользуемся Simulink Response Optimization Toolbox. В качестве настраиваемых параметров выберем $\gamma_1, \gamma_2, \rho_3$. Зафиксируем коэффициенты $\rho_1 = 4,9863, \rho_2 = 0,0557$ [4].
3. Минимизируя функционал (3), найдём вектор коэффициентов регулятора k средствами MATLAB.
4. Исключим из закона управления (2) параметры β и δ , разрешив относительно них первые два уравнения системы (1). Подставим их в закон управления, используя соответствующие величины первых производных. Это даст управление в виде

$$u = K_1\dot{\beta} + K_2\dot{\omega} + K_3\omega + K_4\varphi, \quad (4)$$

где

$$K_1 = \frac{k_1}{(a_{11}/b_1 - a_{21}/b_2)b_1} + k_4 \left(\frac{1}{b_1} - \frac{a_{11}}{b_1^2(a_{11}/b_1 - a_{21}/b_2)} \right),$$

$$K_2 = \frac{-k_1}{(a_{11}/b_1 - a_{21}/b_2)b_2} + k_4 \frac{a_{11}}{b_1 b_2},$$

$$K_3 = \frac{-k_1(a_{12}/b_1 - a_{22}/b_2)}{(a_{11}/b_1 - a_{21}/b_2)} + k_2 + k_4 \left(-\frac{a_{12}}{b_1} + \frac{a_{11}(a_{12}/b_1 - a_{22}/b_2)}{b_1(a_{11}/b_1 - a_{21}/b_2)} \right),$$

$$K_4 = k_3.$$

Таким образом, сначала SRO Toolbox изменяет параметры функционала, затем с помощью среды MATLAB находится вектор коэффициентов регулятора k , который доставляет минимум функционалу (3), по нему вычисляются новые коэффициенты закона управления (4) и изучается динамический процесс. Если график курса $\varphi(t)$ не попадает в заданный коридор, то снова возвращаемся к шагу 1 алгоритма.

4. Компьютерная модель системы управления. Подсистема Simulink пакета MATLAB является наиболее эффективным инструментом для моделирования, имитации и анализа динамических систем.

Сформируем компьютерную модель системы управления движением транспортного судна с водоизмещением 6000 т. Коэффициенты рассматриваемого судна в линейной модели (1) имеют следующие значения: $a_{11} = -0,03408, a_{12} = 0,56, a_{21} = 0,015, a_{22} = -0,306, b_1 = -0,0099, b_2 = -0,00417$.

Для проведения контрольного моделирования зададим возмущения F и M так, чтобы положение равновесия в системе (1) определялось равенствами $\beta_0 = 1, 5^\circ$, $\delta_0 = 12^\circ$. Из системы (1) получаем

$$F = -a_{11}\beta_0 - b_1\delta_0,$$

$$M = -a_{21}\beta_0 - b_2\delta_0.$$

Откуда $F = 0,00296$, $M = 0,000481$.

На рис. 2 представлена Simulink-модель системы управления движением судна, которая реализует описанный выше алгоритм поиска параметров регулятора, обеспечивающего астатизм и требуемую степень устойчивости.

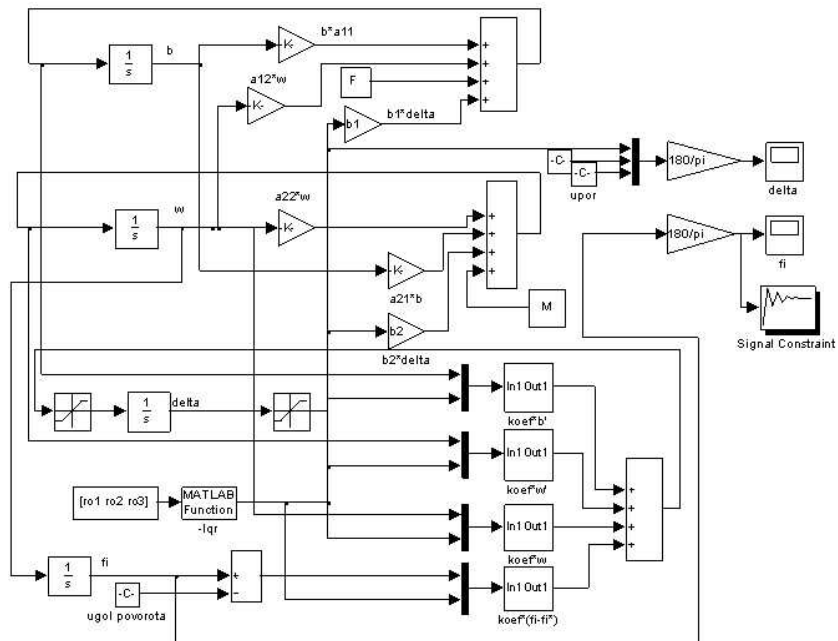


Рис. 2. Simulink-модель

5. Результаты моделирования динамики. В результате компьютерного моделирования с использованием построенной Simulink-модели получены следующие параметры функционала (3)

$$\rho_1 + \gamma_1 10^{-4} = 4,98631, \quad \rho_2 + \gamma_2 10^{-4} = 0,015, \quad \rho_3 = 0,0169$$

и коэффициенты закона управления

$$k_1 = 0,955, \quad k_2 = 6,574, \quad k_3 = 0,941, \quad k_4 = -0,3504,$$

который обеспечивает астатизм замкнутой системы по курсу.

Воспользуемся найденным управлением для автоматического выполнения маневра, состоящего в повороте по курсу на заданный угол $\varphi^* = 30^\circ$, и оценим время переходного процесса. Введём в закон управления (4) командный сигнал φ^* . Тогда

$$u = K_1 \dot{\beta} + K_2 \dot{\omega} + K_3 \omega + K_4 (\varphi - \varphi^*). \quad (5)$$

Время переходного процесса составляет 47 секунд. Замкнем систему (1) управлением (5) и вычислим собственные значения матрицы полученной системы:

$$\lambda_1 = -0,323 + 0,111i, \quad \lambda_2 = -0,323 - 0,111i, \quad \lambda_3 = -0,101, \quad \lambda_4 = -0,05.$$

Таким образом, степень устойчивости равна 0,05.

При расчетах учитываются все технические ограничения морского судна. Графики изменения динамических переменных в процессе маневрирования по курсу представлены на рис. 3.

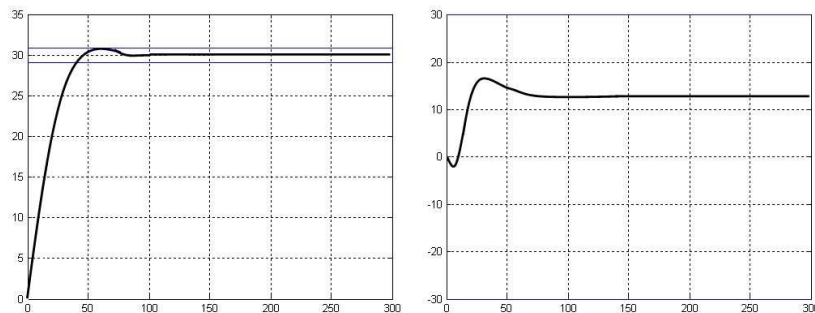


Рис. 3. Изменение курса и отклонение рулей при маневре

6. Заключение. В работе разработан алгоритм построения автоматического управления движением морского судна, обеспечивающего астатизм замкнутой системы по курсу. Указанный алгоритм реализован в пакете MATLAB с использованием подсистемы Simulink. Построена Simulink-модель исследуемой системы управления, найдены параметры регулятора, проведено имитационное моделирование с оценкой качества найденного закона управления.

Литература

1. Лукомский Ю.А., Корчанов В.М. Управление морскими подвижными объектами. СПб.: Элмор, 1996. 320 с.
2. Fossen T.I. Guidance and control of ocean vehicles. John Wiley and Sons. New York, 1999. 494 p.
3. Веремей Е.И., Корчанов В.М., Коровкин М.В., Погожев С.В. Компьютерное моделирование систем управления движением морских подвижных объектов. СПб.: СПбГУ, 2002. 370 с.
4. Веремей Е.И. Синтез законов многоцелевого управления движением морских объектов // Гироскопия и навигация, 2009. Вып. 4. С. 3–14.

Хакимьянова Ю.А.

Санкт-Петербургский государственный университет

Применение подхода BDD к разработке трейдинговых систем

Рекомендовано к публикации доцентом Матросовым А.В.

Введение. Одна из наиболее сложных задач для разработчика – это понимание требований клиента. Чтобы обезопасить себя, заказчики часто формулируют свои требования в формате многостраничных спецификаций, в которых пытаются определить все возможные варианты поведения системы. Разработчик же имеет дело с недokumentированным старым кодом, пытаясь понять, как была задумана работа этого приложения, и одновременно пытается реализовать новые требования, не нарушив функционирования системы. С учетом этих проблем появляются новые подходы к разработке программного обеспечения, которые призваны помогать разработчику в выполнении требований клиентов. Одним из таких подходов является TDD (test-driven-development) – разработка через тестирование [1]. Этот подход стал основой методологии экстремального программирования (XP), где использовалась концепция “test-first”. Суть этого подхода состояла в повторении очень коротких циклов разработки:

1. Написание теста, покрывающего желаемое изменение.
2. Написание кода, который позволит пройти тест.
3. Рефакторинг нового кода к соответствующим стандартам.

Разработка через тестирование очень эффективна, но только в том случае, если в точности выполняются все ее требования. В то же время очевиден главный недостаток TDD – код формируется слишком короткими и частыми итерациями, тесты всё время «падают» после проведенного рефакторинга [2].

Дэн Норт попытался обойти существующие недостатки TDD. В 2003 году он предложил подход к разработке ПО, основанный на поведении системы – BDD (behavior-driven development), который расширяет TDD идеей использования естественного языка при описании тестов. Увидев огромные возможности BDD, Дэн Норт взялся за разработку системы JBehave – BDD-фреймворка для платформы Java, основанного на принципах xUnit [3]. JBehave позволяет

создавать классы для проверки функциональности, так же как и JUnit, однако в случае с JBehave нет необходимости наследовать от какого-либо конкретного базового класса. Как и в случае JUnit, классы JBehave можно запускать в стандартной среде разработки или на предпочитаемой платформе для сборки проекта, такой как Ant, Maven.

Постановка задачи. Была поставлена задача создания модуля, позволяющего проводить разработку финансовых приложений (трейдинговых систем), используя концепцию BDD-тестирования. Необходимость введения программных средств для разработки таких систем через BDD возникает вследствие сложной функциональности финансовых приложений (чаще всего сложные схемы расчёта цен, правила проведения сделок при проведении аукционов и т.д.). Разработанный модуль должен позволять использовать описание поведения системы на естественном языке, сделанное заказчиком. Результатом использования данного модуля является упрощение документирования, вовлечение заказчика в процесс разработки. Преимуществом использования разработанного модуля является тот факт, что над формированием требований к поведению системы и их тестированием трудятся все участники процесса (от пользователей до менеджеров). При разработке указанного модуля использовались библиотеки фреймворка JBehave.

Использование разработанного модуля. В основе методологии BDD лежит сценарий, который служит критерием выполнения требований заказчика по поведению системы и имеет следующий формат описания:

```
Given <условие1>
[And<условие2>]
...
When<событие1>
[And<событие 2>]
...
Then<результат1>
[And<результат2>]
...
```

Иллюстрацию использования разработанного модуля покажем на простом сценарии – сделке купли-продажи. Каждый сценарий созда-

ётся на основе требований, сформулированных заказчиком на естественном языке. Для простой сделки купли-продажи они могут выглядеть следующим образом:

1. При запущенном модуле проведения аукционов. И поднятом контексте мессаджинга системы. И имеющейся на модуле проведения аукционов всей необходимой информации о символе ценной бумаги IBM, паузе между проводимыми аукционами и т.д.
2. Когда пользователь 10004 начинает аукцион, создавая ордер на покупку 200 бумаг IBM по цене 22.30 USD в 15:00:01 1-го февраля 2011 года. И пользователь 10004 посылает ордер для участия в аукционе на продажу 200 бумаг IBM по цене 22.30 USD в 15:00:03 1-го февраля 2011 года.
3. Тогда модуль проведения аукционов получит:
 - для обоих ордеров отчёт (report) о проведении сделки купли-продажи (отчёт со статусом FILLED);
 - и модуль получит сообщение о проведённом аукционе с информацией об инициаторе аукциона, ценной бумаге, количестве и цене проданных и купленных бумаг.

Каждый сценарий хранится в текстовом файле с расширением .feature. Для рассматриваемой сделки купли-продажи он выглядит следующим образом:

```
Scenario: Trading example 1. Filling both received orders.
Given module started AUCTION_ENGINE identified
by id auction_engine
And test ume context is started
And auction_engine contains auction holders:
| symbol | gapMaxInterval | pauseBetweenAuctions |
| IBM    | 1              | 60                    |
When single orders are sent to logical topic
TRADING.AE.REQUEST (auction_engine):
|symbol|price|side|quantity|userId|transactTime      |
|IBM   |22.30|BUY | 200    |10004 |2011-02-01 15:00:01|
|IBM   |22.30|SELL| 200    |10001 |2011-02-01 15:00:03|
When we wait 4 seconds
Then execution reports should be sent to logical topic
TRADING.AE.NOTIFICATION (auction_engine):
```

```
|symbol|side|orderStatus|quantity|leavesQty|cumQty|avgPx |
| IBM  |BUY  | FILLED   | 200   | 0       | 200| 22.30|
| IBM  |BUY  | FILLED   | 200   | 0       | 200| 22.30|
```

And auctions should be sent to logical topic

TRADING.AE.NOTIFICATION (auction_engine):

```
|symbol|auctionType|initiatorId|quantity|clearingPrice|
|IBM   |PUBLIC      |10004      |400     |22.30         |
```

Частью модуля, созданного для тестирования трейдинговых систем, является разработанный класс `AbstractTest`, расширяющий предоставляемый фреймворком `JBehave` класс `JUnitStory`. Задачей этого класса является управление обработкой шагов сценария, подстановка правильных путей к файлам `*.feature` и т.д. Для выполнения сценария создаётся класс с таким же именем как и файл `.feature`, который расширяет `AbstractTest`, и в котором использован метод `createSteps()` – обработчик шагов сценария:

```
public class AuctionEngineOrderProcessing
    extends AbstractTest {
    @Override
    protected List createSteps() {
        return Arrays.asList(new AuctionEngineSteps());
    }
}
```

Запустив сценарий на данном этапе, можно увидеть, что ни один шаг сценария не реализован, а именно, находится в состоянии `PENDING`. Соответственно, далее создаётся связь шагов с конкретным кодом, их реализующим. Для этих целей фреймворк `JBehave` предоставляет аннотации `@Given("...")`, `@When("...")`, `@Then("...")`. Например, реализация отправления сообщения на топик определенного модуля выглядит так:

```
@When('$messages are sent to topic $topic: $table')
def messagesAreSentToTopic(String messages, String topic,
    ExamplesTable table) {
    def objects = toObjects(table, messageClasses[messages])
    objects.each {object->sendMessage(topic, object)}
    Thread.sleep 2000
}
```

Аргументами этого метода являются название класса сообщений, название топика, на который они отправляются, и таблица, представляющая собой инициализирующие поля данного сообщения. После того, как все шаги данного сценария реализованы, проектная команда может убедиться в корректном поведении разрабатываемой системы. Ещё раз выделим последовательность действий, которые надо выполнить при BDD-подходе к разработке ПО:

1. Составить сценарий на естественном языке.
2. Запустить сценарий и посмотреть, для каких шагов ещё нет определений.
3. Написать на используемом языке реализацию соответствующих шагов.
4. Запустить сценарий и проверить, что он корректно выполняется.
5. Периодически «прогонять» данный сценарий в целях проверки того, что процесс дальнейшей разработки не повлиял на правильное функционирование системы.

Возможности разработанного модуля. JBehave позволяет разработчику использовать широкий спектр инструментов для реализации шагов сценария. В нашем случае для реализации шагов использовался язык программирования Groovy, поддержка которого была введена в фреймворк JBehave, начиная с версии 3.1 [4], а также язык Java. Можно выделить основные функциональные шаги, поддерживаемые модулем:

- работа с базой данных (запуск СУБД, создание необходимой БД, выполнение скриптов из текстовых файлов, частичная или полная очистка базы данных);
- получение пути к файлу соответствующего сценария;
- загрузка сценария из текстового файла;
- запуск и остановка различных модулей системы;
- возможность приостановки работы сценария на заданный промежуток времени (пауза);
- создание сообщений, соответствующих спецификации FIX (Financial Information eXchange), и их инициализация в соответствии с их текстовым и табличным представлением;
- возможность отправки и получения различных типов сообщений на различные модули системы;

- возможность обработки полученных сообщений, в том числе проверка того, что список полученных некоторым модулем системы сообщений содержит (не содержит) ожидаемые сообщения;
- и т.д.

Выводы. В представленном модуле был реализован основной набор шагов, характерных для трейдинговых систем. Таким образом, проверка функциональных свойств системы, описание требований к поведению системы сводится к написанию соответствующих сценариев. Наиболее важным является то, что данные сценарии могут составляться и проверяться любым членом проектной команды вне зависимости от его технических навыков (разработчиком, тестировщиком, бизнес-аналитиком, заказчиком и т.д.) Применение BDD-методологии влияет на процесс разработки системы, позволяя:

- связать код с требованиями;
- сформулировать специфичный для предметной области язык;
- записать требования заказчика с помощью обычных фраз;
- улучшить связь с заказчиком.

Литература

1. Beck K. Test-driven development by example. Boston: Addison Wesley, 2003. 240 p.
2. Beust C., Suleiman H. Next generation Java testing: TestNG and Advanced Concepts. Boston: Addison-Wesley Professional, 2007. P. 449–471.
3. North D. Introducing BDD. <http://dannorth.net/introducing-bdd/>
4. Официальный сайт фреймворка Jbehave. <http://jbehave.org/>

Цой О.С.

Санкт-Петербургский государственный университет

Дискретизирующий алгоритм прокладки курса судна с учетом естественных препятствий

Рекомендовано к публикации доцентом Коровкиным М.В.

1. Введение. С времен возникновения мореплавания люди решали задачу прокладки маршрута между двумя точками. Однако эту задачу, которую человек решает с легкостью, трудно решить, используя только алгоритмический подход, т.е. найти алгоритм который строит искомый маршрут для любых исходных точек. В данной работе предложен эвристический алгоритм решения данной задачи.

2. Постановка задачи. Рассмотрим задачу прокладки курса судна с учетом естественных препятствий. Для постановки задачи введем несколько понятий.

Определение 1. *Допустимой* назовем область, допустимую для мореплавания с точки зрения естественных препятствий (островов).

Определение 2. Каждую точку допустимой области (кривую, лежащую в допустимой области) назовем *допустимой*.

Определение 3. Кривую будут называть *оптимальной*, если она имеет минимальную длину на множестве всех допустимых кривых.

Таким образом математически задача ставится так: построить оптимальную допустимую кривую, соединяющую две произвольные, заранее заданные, допустимые точки.

В ходе решения допустимую область будем считать заданной поточно (для каждой точки известно, допустима она или нет). Вследствие этой поправки, получить аналитическое решение этой задачи представляется невозможным.

3. Дискретизирующий алгоритм. В основе его лежит представление о том, что допустимая для плавания зона представляет собой связную область. Пусть также она является замкнутой. Таким образом, любые две допустимые точки можно соединить непрерывной кривой. Укажем один из способов построения этой кривой.

1. Возьмем две допустимые точки и соединим их прямой.
2. Если прямая является допустимой, то кривая построена.
3. Если нет, выбираем точку пересечения прямой с первой преградой и отрезок прямой до этой точки считаем принадлежащим искомой кривой.
4. Исходя из естественных соображений, видно, что преграды имеют замкнутую и непрерывную границу. А так как обе точки являются допустимыми, то существуют две точки пересечения прямой с контуром преграды. Кусок контура, соединяющий две эти точки пересечения, считаем принадлежащим искомой кривой.
5. Будем считать вторую из точек пересечения контура с прямой допустимой и продолжим процесс, начиная с шага 1.

Покажем, как можно адаптировать алгоритм в условиях поточечного задания допустимой области. Считаем, что нам доступны операции:

1. Проверка точки на допустимость.
2. Проверка отрезка (дуги) на допустимость.

Строим путь до первой преграды (аналогично п. 3)

```

BEGIN
Route=add(A)
crs=crsAB
len=AB
WHILE(len>MaxLen)
{
    WHILE(CD - допустим?)
    {
        D=(C, crs, len)
        len=len/2
    }
Route=add(C)
C=D
len=CB
}
END

```

Обходим преграду вдоль контура (аналогично п. 4)

```
BEGIN
Len=CB
WHILE(CB - допустим?)
{
    WHILE(CD - допустим?)
    {
        D=(C, crs, MaxLen)
        crs=crs+MaxAng
    }
Route=add(C)
C=D
crs=crsCB
}
Route=add(B)
END
```

Здесь A, B, C, D – точки; AB, CB, CD – длины отрезков, соединяющих соответствующие точки; crs_{AB} – курс, необходимый для попадания в точку B из точки A ; $D=(C, crs, len)$ означает, что точка D получена продвижением из точки C по курсу crs на расстояние len ; $Route=add(A)$ означает, что добавляем точку A в упорядоченный массив $Route$.

Последовательно соединив точки из массива $Route$, получим искомую кривую.

Для использования алгоритма необходимо задать:

- исходные точки A, B ;
- шаги дискретизации $MaxLen, MaxAng$.

4. Сравнение. Для оценки полученного результата сравним его с маршрутом, построенным с помощью алгоритма Дейкстры [1] на равномерной квадратной сетке. Сравнения проводились для трех различных ситуаций:

1. Без островов – отрезок AB допустим.
2. Единичный остров – отрезок AB пересекается с единственным препятствием.
3. Залив – исходные точки расположены в областях, соединенных узким проходом.

Были проверены результаты выполнения дискретизирующего алгоритма с длиной шага в диапазоне от 5 до 400 км и алгоритма Дейкстры с размерностью сетки от 3×3 до 17×17 . Лучшие результаты по всем параметрам для конкретной ситуации приведены в таблицах:

Таблица 1. Без островов

	Дискретизирующий алгоритм	Алгоритм Дейкстры
Длина шага/размерность графа	любая	3×3
Длина пути	782,658	782,787
Кол-во проверок на допустимость	1	21
Кол-во вычислений	1	37

Таблица 2. Единичный остров

	Дискретизирующий алгоритм	Алгоритм Дейкстры
Длина шага/размерность графа	180	7×7
Длина пути	918,377	936,147
Кол-во проверок на допустимость	7	105
Кол-во вычислений	9	149

Таблица 3. Залив

	Дискретизирующий алгоритм	Алгоритм Дейкстры
Длина шага/размерность графа	150	15×15
Длина пути	415,754	526,612
Кол-во проверок на допустимость	51	465
Кол-во вычислений	53	580

Здесь под количеством вычислений понимаем количество операции двух видов:

1. Найти точку, если известна исходная точка, курс и расстояние.
2. По двум известным точкам вычислить курс и расстояние.

5. Заключение. Предложен алгоритм прокладки курса судна с учетом естественных препятствий. Сравнение его с алгоритмом Дейкстры показало, что он требует меньше вычислительных затрат и, в некоторых случаях, дает меньшую длину пути. При построении маршрута в ситуации «вхождение в узкую зону» предложенный алгоритм обладает рядом преимуществ перед алгоритмом Дейкстры и другими алгоритмами на графах, которые при увеличении размерности графа могут даже не справиться с поставленной задачей.

Литература

1. Кристофидес Н. Теория графов. Алгоритмический подход. М.: Мир, 1978. 432 с.

Якушкин О.О.

Санкт-Петербургский государственный университет

Искусственная нейронная сеть для нахождения сегментов линий

Рекомендовано к публикации доцентом Гришкиным В.М.

Введение. Выделение сегментов линий является важной задачей в распознавании геометрического содержания изображений, применяемой в распознавании прямых, плоскостей и более сложных объектов. Эта проблема была обозначена в 80-х годах [3–5]. В процессе исследований разрабатывался ряд концептуальных методов ее разрешения. Однако, так как поставленная задача не имеет точного решения, каждый из предлагаемых методов давал свой результат, имея разную производительность. В 2010 году был опубликован алгоритм [2] Line Segment Detector (LSD), находящий линии за линейное время. Изучение этого метода стало отправной точкой для идеи создания алгоритма поиска сегментов линий средствами искусственных нейронных сетей (ИНС).

Создание искусственных нейронных сетей развивалось с 60-х годов [1, 6, 7] как попытка смоделировать работу человеческого мозга. Несмотря на неполные, «зашумленные», частично искаженные или отсутствовавшие в обучающей выборке входные данные, ИНС после успешного обучения позволяет получать корректные результаты. Поиск сегментов линий на изображении может являться задачей для работы ИНС.

Постановка задачи. Для создания ИНС, которая сможет осуществить поиск сегментов линий на изображении, изначально упростим задачу. Разработаем схему ИНС для поиска одного сегмента линии на «черно-белых» (в оттенках серого) картинках малого размера. Будем исходить из условия, что на рассматриваемых участках изображения помещается не более одной линии. Искать один сегмент линии означает выявить, есть ли сегмент линии в рассматриваемом регионе изображения и как он повернут относительно своего центра.

Сбор данных для обучения. Для решения поставленной задачи выделим класс картинок для обучения нашей ИНС. Чем уже класс, тем лучше обучится работе с ним ИНС. Как наиболее подходящий пример контрастных, ярко выраженных скоплений линий,

используем изображения урбанистических сред. Приведем картинки к размерности $N \times N$ методом разрезания. Выберем N таким, чтобы свести количество сегментов линий на отрезке до одного. При этом будем рассматривать только «черно-белые» изображения. Для каждого из них проведем поиск сегментов линий существующим алгоритмом LSD. Соберем данные о наличии или отсутствии сегментов линий и повороте найденного сегмента относительно центра на каждой отдельной картинке. Разобьем угол поворота сегмента линии на набор заданных промежутков. В соответствии с описанной методикой сбора данных уточним задачу. Теперь будет целесообразным создание набора классификаторов на базе одной схемы ИНС. Один классификатор будет сигнализировать о наличии или отсутствии сегмента линии на изображении. Для каждого промежутка угла поворота создадим классификатор для определения попадания реального угла поворота сегмента линии в промежуток. Чем больше набор таких классификаторов, тем проще его определять.

Подготовка и нормализация данных. Для оптимизации подготовки и нормализации данных на языке программирования C++ была разработана программа, способная собирать, сегментировать и анализировать информацию об искомых элементах на изображениях произвольного размера, в разных форматах в указываемой папке.

Программа работает по следующему алгоритму. Для каждого изображения в папке:

- производится приведение информации к градациям серого (фильтрация);
- проводится сегментация на квадраты с длиной стороны, заданной пользователем;
- к каждому квадрату:
 - применяется алгоритм поиска сегментов линий LSD (с максимальными параметрами точности);
 - применяется алгоритм нормировки – цвет каждого пикселя в рассматриваемом регионе приводится к значению от «0» до «1», где «1» – значение самого яркого пикселя в регионе и «0» – значение самого темного;
 - по требованию пользователя над нормализованными данными может производиться квантование – приведение значений к набору нулей и единиц.

По требованию пользователя программа может выводить обучающие пары как для всех квадратов, так и только для тех, в которых были найдены сегменты линий.

По умолчанию программа выводит:

- «-1» следом за квадратом, в котором не найдено сегментов;
- для тех квадратов, в которых есть хотя бы один сегмент линии, выводится квадрат, за которым следует значение:
 - «1», если угол поворота сегмента относительно центра попал в регион, заданный пользователем;
 - «-1» в случае, если линия найдена, но угол поворота отличен от запрашиваемого пользователем.

Результирующие обучающие пары выводятся в следующем формате (CSV):

- матрица нормализованных значений;
- значение «1» или «-1».

Выбор топологии сети. Обучающие ИНС данные, полученные от программы, содержат точные «экспертные» оценки. Это позволяет выбрать топологию сети, аналогичную архитектуре многослойного перцептрона.

В данном случае не будем использовать обходные соединения в схеме ИНС, что облегчит практическую реализацию.

Для решения поставленных задач классификации: определения наличия или отсутствия линии и факта попадания угла в заданный промежуток – достаточно одного скрытого слоя и прямого распространения сигнала по ИНС.

На всех нейронах будем использовать логистическую сигмоидальную передаточную функцию. Это предоставит гибкость для подбора коэффициентов сумматоров средствами эмуляции ИНС.

Завершающим элементом ИНС станет сумматор.

Экспериментальный подбор характеристик сети. Прочитаем построчно подаваемое изображение размерности $N \times N$. Для каждой строки создадим сумматор (сумма пикселей, содержащихся в строке с настраиваемыми весами). Результаты работы сумматоров поместим в передаточную сигма-функцию активности нейронов. Создадим сумматор для N значений активности нейронов. Завершением работы ИНС будем считать результат, полученный этим сумматором.

Экспериментальный подбор параметров обучения. На входе в ИНС подаются нормализованные матрицы значений цветов. Каждому элементу матрицы присвоим настраиваемый коэффициент. Матрицы будут построчно считываться сумматорами. Данные, генерируемые сумматорами, будут подаваться в передаточные функции.

Значениями активности нейронов будем считать результаты работы передаточных функций. Для $(N \times N)$ -матрицы будем иметь N нейронов. Значения активности нейронов будут складываться в сумматоре. Присвоим каждому значению активности нейронов свой настраиваемый коэффициент. В сумме получим $N^2 + N$ параметров для обучения.

Обучение сети. В целях апробации описанной ИНС создадим и обучим девять классификаторов. Один – для определения наличия во фрагменте изображения сегмента линии, и восемь – для определения угла поворота сегмента линии. Для обучения ИНС используется пакет математического программирования. Исходя из замечания [2] создателей алгоритма LSD о том, что минимальное представление сегмента линии – это прямоугольник размером восемь на два пикселя, была выбрана минимальная размерность подаваемой на вход ИНС матрицы цветов, равная восьми. Обучение проводится на выборке размером от трехсот обучающих пар.

Проверка адекватности обучения. Выделим критерии адекватности: соотношение верных и неверных попаданий, абсолютная погрешность, средняя погрешность.

Проверка адекватности работы классификаторов проводится на выборке, более чем из 3000 пар, не принимавших участия в обучении. Все классификаторы показывают результат: один промах на пять попаданий, что не превышает расчетную допустимую погрешность в 25 процентов.

Классификация полученной нейронной сети. Полученная ИНС по типу входной информации является аналоговой (использует информацию в форме действительных чисел), по характеру обучения относится к ИНС с обучением с учителем. По характеру настройки синапсов эта система является сетью с фиксированными связями. Характер связей относит ее к классу сетей прямого распространения (Feedforward), т.е. все связи направлены строго от входных нейронов к выходным.

Заключение. Созданный набор классификаторов может быть успешно использован не только на специальной аппаратуре, но и в рамках симулирующих ИНС систем. Набор классификаторов работает за линейное время, что является критерием успешного решения поставленной задачи распознавания сегментов линий.

Литература

1. Мак-Каллок У.С., Питтс В. Логическое исчисление идей, относящихся к нервной активности // В сб.: «Автоматы» под ред. К.Э. Шеннона и Дж. Маккарти. М.: Изд-во ИЛ, 1956. С. 363–384.
2. Grompone R., Jakubowicz J., Randall G. LSD: a fast line segment detector with a false detection control // IEEE Transactions on Pattern Analysis and Machine Intelligence / Ed. by Ramin Zabi, Computer Society, 2010. P. 722–732.
3. Duda R.O., Hart P.E. Use of the Hough transformation to detect lines and curves in pictures // Comm. ACM, 1972. Vol. 15. P. 11–15.
4. Hough P.V. Machine analysis of bubble chamber pictures // Proc. Int. Conf. High Energy Accelerators and Instrumentation, 1959. P. 554–557.
5. Nevatia R., Babu R. Linear feature extraction and description // Computer Graphics and Image Processing. New York: Academic Press, 1980. Vol. 13. P. 257–269.
6. Минский М., Пейперт С. Перцептроны. М.: Мир, 1971. 261 с.
7. Хайкин С. Нейронные сети: полный курс. 2-е изд. М.: Изд. дом «Вильямс», 2006. 1104 с.

Li M.H.^{1,2}, Klyuev V.V.¹, Wu S.H.²
University of Aizu, Aizu-Wakamatsu, Japan¹
Chaoyang University of Technology, Taiwan²

Sentence Alignment from Wikipedia as Comparable Corpora by using STF-IDTF

1. Introduction. Nowadays, bilingual dictionaries and multilingual dictionaries are developed on the Internet. Thus, a monolingual document can be translated into several languages. This situation can help to quickly retrieve the similar content in different languages. Therefore, many researchers are involved in building cross-lingual parallel corpora. Parallel corpora can help for machine translation and cross language information retrieval. Sentence alignment of parallel corpora is one of possible ways to mine knowledge. However, in the real world many documents can be presented in comparable corpora. Comparable corpora refer to the same topic in different languages. Compared to parallel corpora, comparable corpora promote concepts, although the number of comparable corpora is larger than the number of parallel corpora, but detecting sentence pairs with the same meaning is difficult. Wikipedia is a large source of information. Most of the entries are expressed in many different languages. Researchers can employ it as a bilingual dictionary [10], since the same entries in different languages were edited by authors who can edit them. Therefore, Wikipedia can be considered as a set of cross languages dictionaries. Wikipedia content editors are all users of the network; therefore, Wikipedia can be considered as comparable corpora. In this paper, we introduce the technique to select parallel sentences from Wikipedia as comparable corpora, and utilize the parallel sentences to create multilingual parallel corpora.

2. Related work. In fact, Wikipedia provides many entries in many different languages. Martin Hepp [2] describes that Wikipedia entry name almost does not change eventually. Therefore, many researchers use Wikipedia to perform the cross-lingual NLP tasks. Hsu [3] utilized the anchor texts to perform query expansion from Wikipedia as multilingual entries for a cross-lingual information retrieval. Lin in 2010 not only utilized the anchor texts, but also combined a topic-oriented web crawler for a cross-lingual information retrieval [4]. Today, many researchers create parallel corpora, because parallel corpora are useful

to perform the cross-lingual NLP. Ralf [7] developed available parallel corpora containing European Union documents. Its are available in 20 official European Union languages. Corpus consists of almost 8,000 documents per language with an average size of nearly 9 million words per language. Its available for 190 language pair combinations by using two aligners (Vanilla and HunAlign). Some of studies used machine translation to perform a sentence alignment [9], but it needs a lot of time to utilize training data, and the removal of noisy data is difficult. Therefore, when the system performs a sentence alignment from comparable corpora, it will generate a large number of noisy data; it will make identifying parallel sentences more difficult for the sentence pairs. In order to solve this problem, in next section we explain the architecture of proposed system.

3. Performing sentence alignment. Wikipedia provides dump data to users that they have Wikipedia data in different languages. In our research, we perform a sentence alignment between English and Chinese. In fact, our approach is applicable not only to English and Chinese; another language pairs can be used instead. We can perform a sentence alignment between any two languages via change of the bilingual dictionary. We can obtain article pairs from Wikipedia for the same entry in two different languages. Then we select candidate sentences for an alignment. This view allows us to avoid excessive noisy data receiving. Finally, we use the Champollion software [5] to perform the sentence alignment from candidate sentences, and create parallel corpora. In section 3.1, we describe how to obtain article pairs. In section 3.2, we discuss how to select candidate sentences for an alignment and use the Champollion to perform a sentence alignment from candidate sentences.

3.1. Article pairs. We take into account the entries which have cross-lingual links. We can obtain an article pair from the article level when use language with lower number of entries to match other language with higher number of entries. A direction from lower number of entries to higher number of entries can save time. Lower number of entries may be found in other language for higher number of entries.

3.2. Sentence alignment. Although, we can obtain article pairs with the entries for many different languages, but they have a lot of noisy data for a sentence alignment in each article pair, because Wikipedia is comparable corpora. For example, for the entry “Confucius”, English content length is longer than Chinese one. We applied the idea

of STF-IDTF proposed in [5]. The key feature of this technique is to utilize total number of word pairs divided by length of longer sentence. We set a threshold to decide if two sentences in different languages can be candidate sentences for an alignment. It can be computed with

$$\begin{cases} \text{Candidate sentence,} & \frac{\text{Number}(\text{pair}(E,C))}{l} > t, \\ \text{non-Candidate sentence,} & \text{otherwise,} \end{cases}$$

where E and C are two sentences in English and Chinese; pair represents word pairs in two sentences using a bilingual dictionary; l represents longer length among two sentences, and t represents the threshold. This function is applied to each sentence in English and Chinese from the article pair. As a next step, we can perform the sentence alignment for selected candidate sentences. We use the Champollion for this purpose. It considers a match that is possible only when lexical matches are presented; assigns higher weight to less frequent words, which are considered a stronger indication. Then two segments are matched. The Champollion has two main functions: the similarity function and the dynamic programming function. The similarity function uses the idea of TF-IDF (term frequency – inverse document frequency), and the name STF-IDTF stand for segment-wide term frequency – inverse document-wide term frequency. The similarity function works as follows. First, for two sentences in different languages E and C , it defines sets of tokens:

$$\begin{aligned} E &= \{e_1, \dots, e_m\}, \\ C &= \{c_1, \dots, c_n\}, \end{aligned}$$

where e and c are token words. Then it defines k translated word pairs identified between two segments:

$$\text{pair} = \{(pe_1, pc_1), \dots, (pe_k, pc_k)\},$$

where pair represents the word pair selected from a bilingual dictionary. And the STF means the number of occurrences of a term within a segment. The IDTF is a measure of general importance of the term in the document. It is defined as follows:

$$\text{idtf} = \frac{T}{\text{occurences in the document}},$$

where T is total number of terms in the document. Finally, the similarity function computes the similarity value between two sentences in different languages, it is defined as follows:

$$\text{sim}(E, C) = \sum_{i=1}^k \ln(\text{stf}(pe_i, pc_i) \times \text{idtf}(e_i))$$

The dynamic programming function can compute and save the maximum score of iteration. Its core algorithm is very similar to one presented in [1]. It searches for a path with the maximum similarity. Therefore, it can find the optimal path on the base of the maximum score. The Champollion allows 1-0, 0-1, 1-1, 1-2, 2-1, 2-2, 1-3, 3-1, 1-4 and 4-1 sentence alignment types, which mean one source sentence by a zero target sentence, a zero sentence by one target sentence, etc. The dynamic programming algorithm is defined as follows:

$$S(i, j) = \max \left\{ \begin{array}{l} S(i-1, j) + \text{sim}(\text{Seg}_{i,i}, 0) \\ S(i, j-1) + \text{sim}(0, \text{Seg}_{j,j}) \\ S(i-1, j-1) + \text{sim}(\text{Seg}_{i,i}, \text{Seg}_{j,j}) \\ S(i-1, j-2) + \text{sim}(\text{Seg}_{i,i}, \text{Seg}_{j-1,j}) \\ S(i-2, j-1) + \text{sim}(\text{Seg}_{i,i-1}, \text{Seg}_{j,j}) \\ S(i-2, j-2) + \text{sim}(\text{Seg}_{i-1,j}, \text{Seg}_{j-1,j}) \\ S(i-1, j-3) + \text{sim}(\text{Seg}_{i,i}, \text{Seg}_{j-2,j}) \\ S(i-3, j-1) + \text{sim}(\text{Seg}_{i-2,i}, \text{Seg}_{j,j}) \\ S(i-1, j-4) + \text{sim}(\text{Seg}_{i,i}, \text{Seg}_{j-3,j}) \\ S(i-4, j-1) + \text{sim}(\text{Seg}_{i-3,i}, \text{Seg}_{j,j}) \end{array} \right.$$

where S represents a similarity between sentences from the beginning of the document to the i -th source sentence and j -th target sentence, $\text{Seg}[a,b]$ represents all segments numbered between a and b . In section 4, we describe how many article pairs we obtained between English and Chinese, and discuss an example of candidate sentences for the alignment.

4. Experimental results. In the previous section, we described how to perform alignment for the article pairs. We employed cross-lingual links to extracted article pairs through English entries link to Chinese entries from Wikipedia. Table 1 shows the number of Wikipedia entries in two different languages and article pairs. There are more than half of Chinese entries, which have pairs in English. English entries, which have no association with Chinese entries, can be removed. This consideration can help for selecting candidate sentences and the sentence alignment.

Table 1. Total number of Wikipedia entries and article pairs

	English	Chinese	Article pairs
Total number of entries	3,449,018	332,689	224,985

We used an on-line dictionary (Google translation, 2010) to selected candidate sentences for alignment; the threshold was set to 0,5. Table 2 shows an example of candidate sentences for the alignment as a result of this selection. The source language and target language are Chinese and English, respectively. In fact, we still need to adjust the threshold to perform the candidate selection. If we set a high threshold value, although our mechanism will obtain qualified candidate sentence pair for the alignment, but it may also lose other sentences, which can be good as candidate sentences. On the other hand, if we set a lower threshold value, we may receive noisy data. We need to conduct more tests to find the optimal value for the threshold. However, we have an inspiring that candidate sentences for the alignment can be helpful to perform the sentence aligning.

Table 2. A candidate sentences for alignment.

<Title>1378 1378 年 </Title>
<Content>
<Source Language>
11 月 29 日－查理四世（神聖羅馬帝國）於布拉格（生於 1316 年）
</Source Language>
<Target Language>
February 6 Jeanne de Bourbon, queen of Charles V of France (born 1338)
November 29 Charles IV, Holy Roman Emperor (born 1316)
</Target Language>
</Content>

5. Conclusion. In this paper, we introduced the technique of the parallel sentence extraction from Wikipedia as multilingual comparable corpora. Traditional approaches are length-based, they utilize anchor texts and use MT to perform the sentence alignment; therefore, the performance of the alignment is very slow, when applying them. Therefore, with selection candidate sentences for alignment and the Champollion, we can save time and increase the quality of the alignment. In the future, we will perform the parallel sentence extraction for more than two languages, and also create parallel corpora from cross parallel corpora. For example, let us assume that we have two parallel corpora: English-Chinese and English-Japanese. Then we can compare the common part

for Chinese and Japanese to English, and create new parallel corpora, namely, Chinese-Japanese.

References

1. Gale A.W., Church K.W. A program for aligning sentences in bilingual corpora // ACL, 1991.
2. Martin H., Katharina S., Daniel B. Harvesting Wiki Consensus: Using Wikipedia Entries as Vocabulary for Knowledge Management // IEEE Internet Computing, 2007. Vol 11, No 5. P. 54–65.
3. Hsu C.C., Li Y.T., Chen Y.W., Wu S.H. Query Expansion via Link Analysis of Wikipedia for CLIR // Proceedings of NTCIR-7, 2008.
4. Meng-Chun Lin, Ming-Xiang Li, Chih-Chuan Hsu, Shih-Hung. Query Expansion from Wikipedia and Topic Web Crawler on CLIR // Proceedings of NTCIR-8 Workshop Meeting, 2010. No 6.
5. Xiaoyi Ma. Champollion: A Robust Parallel Text Sentence Aligner // In Proceedings of LREC, 2006.
6. Smith J., Quirk C., Toutanova K. Extracting parallel sentence from comparable corpora using document level alignment // In North American Chapter of the Association for Computational Linguistics (NAACL), 2010.
7. Steinberger R., Pouliquen B., Widiger A., Ignat C., Erjavec T., Tufis D., Varga D. The JRC-Acquis: A Multilingual Aligned Parallel Corpus with 20+ Languages // In proceeding of the 5th International Conference on Language Resource and Evaluation, 2006.
8. Tyers M.F., Pieanaar J.A. Extracting Bilingual Word Pairs from Wikipedia // In LREC 2008, SALTMIL Workshop, 2008.
9. Keiji Yasuda, Eiichiro Sumita. Method for Building Sentence-Aligned Corpus from Wikipedia // AAAI, 2008.
10. Kun Yu, Junichi Tsujii. Bilingual Dictionary Extraction from Wikipedia // Proceeding of the 12-th Machine Translation, 2009.

Tikhomirov V.A.
Saint-Petersburg State University

Spectral image compression using different clustering methods

Recommended by professor Veremey E.I.

1. Introduction. This research is devoted to improvement of spectral image compression methods. Spectral image is the set of pixels where each pixel belongs to the spectral space. Notice that the quality of images for this spectral representation is very high, but the disadvantage of spectral approach is that this space has large dimension. Its value can vary from 61 to 81 depending on a device taking a spectral picture. Thus the spectral image file can be very large and it is uncomfortable to show these pictures in slide-show or to send it through the Internet because needs to wait for a long time. The main idea of research is to divide points of spectral image to some clusters and then compress each cluster with the help of Principal Component Analysis (PCA) method. One of the interesting questions is the choice of the space for clustering. This article presents clustering only for RGB model, so it is necessary to transform spectral image to RGB color space before the clustering. On the other hand image compression with the help of PCA is performed in spectral space. This article describes two approaches of clustering application: well-known clustering method “k-means” and new method of clustering using some special planes.

2. Problem definition. Let the original image be a spectral image for the following type: $I_o = \{X_i\}_{i=1}^{mn}$, $X_i \in R^l$, where m is a length of spectral image, n is a width of spectral image, l is a spectral space dimension. Size of the original image is designated by S_o with mn . Compression of spectral image can be performed by two steps. First, elements of spectral image are divided with some clustering algorithm, but before the spectral image is transformed into the RGB color model. Clustering algorithm can be treated as an operator K dividing original image $I_o = \{X_i\}_{i=1}^{mn}$ into h disjoint sets called clusters. That is each element of original image is associated with the number of certain cluster by the following operator $K : X_i \in R^l \rightarrow y_j \in Y_j$, where $j = \overline{1, h}$.

On the second step each cluster is compressed with the help of PCA approach [2], which is a method for transformation a certain number of

possibly correlated variables into a smaller number of uncorrelated variables called principal components. PCA was developed by Karl Pearson in 1901. The main idea of this algorithm is to calculate covariance matrix $C = \frac{\sum_{i=1}^{mn} (X_i - \mu_x)(X_i - \mu_x)'}{mn}$, where $X_i \in R^l$ and $\mu_x = E\{X_i\}$ is a mathematical expectation. Let find eigenvectors and eigenvalues of matrix C and arrange them in the order of descending eigenvalues (largest first). Thus the algorithm finds directions where the data set has the most significant scattering. After that it is necessary to choose finite number of eigenvectors b . The result of this algorithm application is the following compressed form of spectral image:

$$y_i = V(X_i - \mu_x), \quad (1)$$

where V is matrix of eigenvectors, $\dim(V) = lb$. It is obvious that parameter b directly influences on the compression ratio. So, compressed image is the object of the following type: $I_c = \{W, y, \mu\}$, where W is the matrix of eigenvectors for all clusters: $W = \{V_i\} i = 1, h$, $\dim(W) = lbh$, y is compressed form of image, $\dim(y) = nm(b + 1)$. The term "1" is used in this formula for additional layer for clustering matrix containing the numbers of clusters which are placed to the same positions as the correspondent pixels of the original image. Parameter μ is a mathematical expectation of the central point for each cluster, $\dim(\mu) = lh$. The compression algorithm is an operator F_c , which divides spectral image to clusters with the help of clustering algorithm K and then transforms these parts of original image to compressed image using PCA method: $I_c = F_c(I_o)$. So, the size of compression file be $S_c = lbh + nm(b + 1) + lh$.

The decompression algorithm is an operator F_r , which transforms compressed image to restored image: $I_r = F_r(I_c)$. Using PCA the restored image I_r is calculated by $X_i^* = V'y_i + \mu_x$. Restored image is the spectral image, which is obtained by transformation from compressed image using decompression algorithm F_r : $I_r = \{X_i^*\}_{i=1}^{mn}$, $X_i^* \in R^l$. Size of restored image is equal to size of original image: $S(I_o) = S(I_r)$. The original and restored images can be watched after transformation to the RGB color model [1]. Remark that compressed image is only the mathematical form of original image which is used only for storage and can not be watched directly: firstly it is necessary to apply decompression algorithm F_r .

To estimate the quality analysis of image compression it is convenient to introduce functionals of quality:

- 1) $F = \frac{\sum_{i=1}^{mn} (X_i - X_i^*)^2}{mn}$ is a spectral functional of quality. This functional is mean square error (MSE) which measures the average of the squares of the "errors".
- 2) $E = \frac{\sum_{i=1}^{mn} \sqrt{(L_i - L_i^*)^2 + (a_i - a_i^*)^2 + (b_i - b_i^*)^2}}{mn}$ is a Lab [1] functional of quality. This functional describes the average difference special for human eyes, because color space Lab was created for mimic the nonlinear response of the human eye.
- 3) $E = \frac{\sum_{i=1}^{mn} \sqrt{(T_i - T_i^*)^2}}{mn}$ is a RGB functional of quality. This is the same functional as the first one but in other color space RGB, where $T_i = (X_i)_{RGB}$.

The main idea of the optimization problem is to minimize one of quality functionals F , E , D

$$\min_{h,b} F, \min_{h,b} E, \min_{h,b} D. \quad (2)$$

Also this optimization problem depends on clustering algorithm K which was discussed before. In this article clustering is performed in RGB color model but it is possible to choose another color space. Using these parameters algorithm can to solve optimization problem but image file will be too large. That is why special condition is introduced to be fulfilled: the ratio between size of compressed image S_c and original image S_o should be less than a given constant value q : $S_c/S_o < q < 1$.

The problem of the choosing parameters h, b appears during the solution process of optimization problem. All these parameters are integer numbers, but the result depends on the clustering algorithm and color space which are fixed before. So, this algorithm needs to solve the following problem of clustering: using clustering algorithm K divide original image $I_o = \{X_i\}_{i=1}^{mn}$ to clusters by such a way that the correspondent cluster elements be close with metrics ρ .

3. The ways of solution. This section presents a description of the ways to solve the problem presented above in the section 2. The main idea of this research is to apply clustering to compression. So, if the set is divided to different clusters and then each cluster is compressed by PCA the results will be better than if the whole set is compressed only by PCA. First part of this section presents an application description of the well-known method of clustering k-means [3] and the second part presents the application description of new proposed method.

3.1. K-means clustering. Let the input data be original image $I_o = \{X_i\}_{i=1}^{mn}, X_i \in R^l$. Before the input data clustering let transform initial image to the RGB model, because the clustering is performed in RGB color space [2]. The main idea of this method is to divide points to clusters by definition which point is closer to which center.

The algorithm of this method consists of the following steps:

1. Choose the number of clusters h and the centers of clusters randomly $C_j, j = \overline{1, h}$.
2. Calculate the distances in Euclidean metric between each point and each center of cluster: $D_{ij} = \sqrt{(X_i - C_j)^2}, i = \overline{1, mn}, j = \overline{1, h}$.
3. Calculate the minimum distance from each point for each cluster $\overline{D}_i = \min_j D_{ij}$. This parameter shows which point is closer to which center of cluster.
4. If the center of cluster is the closest center to this point then this point belongs to this cluster.
5. Change the centers of clusters. New center calculates as arithmetical mean of all points of each cluster. $\overline{C}_j = (\sum_{h=1}^{p_j} X_h^j)/p_j$, where $p_j, j = \overline{1, h}$ is the amount of points in each cluster.
6. The algorithm stops if the centers of prior step is equal to the centers of current step; otherwise new centers equate to the old centers and the algorithm continues from the step 2.

Output data is presented by the matrix I_{cl} , $\dim(I_{cl}) = mn$ where the position $I_{cl}(e, t)$ is placed by the number of cluster according to the location on original image. This method is simple, but there are some problems. Firstly, this algorithm requires a lot of time for clustering because it uses all the points of input data during each iteration. This algorithm strongly depends on the centers which are chosen in the beginning of the algorithm. Secondly, this algorithm requires to know the number of clusters.

3.2. Planes clustering. Let the input data be original image $I_o = \{X_i\}_{i=1}^{mn}, X_i \in R^l$. Before the input data clustering initial image must be transformed to the RGB model, because the clustering is performed in RGB color space [2]. This method divides the input set by the special planes.

The method consists of the following steps:

1. Calculate two eigenvectors by PCA, which show the direction of maximum dispersion of the input set points. These vectors have orthogonal characteristic.

2. Using these eigenvectors and one point which was chosen randomly draw the primary plane $\pi_0 : Ax + By + Cz + D = 0$.
3. Calculate the distance between each point and primary plane using following formula: $d(\pi_0, X_i) = \frac{|Ax_i + By_i + Cz_i + D|}{\sqrt{A^2 + B^2 + C^2}}$. During this process count the number of points which are from both side of the primary plane. After that calculate the parameter $G = \frac{c_+}{c_-}$, where c_+ is the number of points from positive side and c_- is the number of points from negative side. If c_- is bigger than c_+ G is calculated by the another formula: $G = \frac{c_-}{c_+}$. So, the condition $G \geq 1$ is always fulfilled.
4. Calculate the distance between primary plane and the point which has the maximum distance from primary plane: $d_{max} = \max_i |d(X_i, \pi_0)|$.
5. Choose the value G_{fix} . Against to this parameter the method draws cluster planes to one side or to both sides from the primary plane.
6. Choose the number of clusters h . The number of clusters corresponds with number of planes.
7. If $G \geq G_{fix}$ then the algorithm calculates step $s = \frac{d_{max}}{h}$ and builds $(h-1)$ planes with step s from one side of the primary plane where the count of points is bigger.
 - (a) If $G < G_{fix}$ and h is odd number the step is calculated by $s = \frac{2d_{max}}{h}$ and builds $\frac{(h-1)}{2}$ planes with step s from both sides of the primary plane;
 - (b) If $G < G_{fix}$ and h is even number the step is calculated by $s = \frac{2d_{max}}{h}$ and builds $\frac{h}{2}$ planes with step s from both sides of the primary plane and the primary plane isn't a cluster plane, it is used only for building another planes which are cluster planes.
8. If $d(X_i, F_j) = \min_j d(X_i, F_j)$ then point X_i belongs to cluster j .

Output data is the matrix I_{cl} , $\dim(I_{cl}) = mn$ where on $I_{cl}(e, t)$ location there is the number of cluster according to the location on original image.

4. Experiments and results. In this section the tables of the results are presented for the developed programming support. All the experiments were obtained when $b = 2$, $Ratio = \frac{S_o}{S_c}$. The file size was 6,7 Mb. The limit of q is $\frac{1}{3}$.

Table 1. PCA+k-means results

h	F	E	D	Ratio	File size(Kb)
4	1.1648	8.2826	0.0601	30.86	213
5	1.1045	7.6814	0.0520	30.63	215
6	1.0668	7.3484	0.0451	30.45	216
7	0.8499	7.2482	0.0387	30.24	217

Table 2. PCA+planes results

h	F	E	D	Ratio	File size(Kb)
4	1.0284	10.047	0.0501	30.86	213
5	1.0556	9.7249	0.0482	30.63	215
6	0.9588	9.1215	0.0449	30.45	216
7	0.8259	8.5569	0.0403	30.24	217

Table 3. PCA results

b	F	E	D	Ratio	File size(Kb)
2	1.3644	10.256	0.0602	31.68	210

5. Conclusion. Proposed clustering methods improve the quality of spectral image, but the file size grows in a small range. The comparison of the table 1 and table 2 the functionals F and D have better results in planes clustering than in k -means clustering. But the results of functional E are not too bad.

Maybe if the compression is not in RGB color space the results will be better. This research can be improved by this idea.

References

1. Andriashin A. Master Thesis. University of Eastern Finland. Spectral color analysis, 2004. 74 p.
2. Pomerantsev A.L. Principal Component Analysis (PCA), 2008.
3. Likas A., Vlassis N., Verbeek J.J. The global k -means clustering algorithm, 2002.

Ueno R., Klyuev V.V.
University of Aizu, Aizu-Wakamatsu, Japan

Semantic Search Engine Query Expansion Using WordNet

1. Introduction. Nowadays, the World Wide Web is a useful source of information, and the Web is huge and still growing explosively. Therefore, search engines, such as Google, Yahoo!, Baidu and so on, are frequently used on the Web to search for information needed by users. But searching accurately, utilizing search engines is still difficult for many users, especially for beginners of the Internet.

One of the reasons for it is the difficulty in expressing queries, applying a key word language. So many Internet users often cannot find right key words and then they express ambiguous queries. With ambiguous queries, search engines might output results that are not wanted by users.

For getting over that difficulty, the method called a query expansion has been introduced in the research area of information retrieval. Query expansion is the process that expands the query consisting of keywords typed by the users. Research on query expansion is also related to the research area of natural language processing, because queries could be not only lists of the words but also sentences.

The query expansion system, such as “Google Suggest” [1], generates expanded queries based on queries typed by users, and shows those expanded queries to the users for selections. Many systems of this type use statistical way to perform query expansion. But in that way, sometimes the system does not work accurately. So in that way, ambiguous queries are often not expanded in accordance with users’ wishes.

In this research, we propose a semantic approach to query expansion using WordNet [2] and its API, WordNet::SenseRelate::AllWords [8]. WordNet is a lexical database in English. According to our proposal, the system analyses and disambiguates words in the query and expands the query based on that semantic analysis. This approach is one of the possible solutions for the problem of design an intelligent search system utilizing query expansion.

2. Related Work. Many studies are carried out in the research area of query expansion. Therefore, there are various approaches to the query expansion. In this section, we review two major types of approaches to build a system that performs query expansion.

2.1. Basic Idea. Query expansion is the process that analyses the user's query and adds terms related to the user's query words. To process that analysis, query expansion systems scan information such as relations and relevance of the words, the data about the user, the search history and so on. And the systems get the query words that make the user's questions more precise.

2.2. Statistical Approaches. These approaches use statistical data such as searched queries, access frequency of search results, etc. For example, Google Suggests and some studies [3, 4] adopt these approaches. So these approaches are based on the logs of user's actions when he or she searches for information on the Internet. In these approaches, query expansion systems discover significant relationships between words based mainly on a term co-occurrence analysis, then expand the queries by using those relationships. Those relationships often have not linguistic interpretations. We should prepare a lot of logs if we use these approaches, because these approaches don't perform well in case of few data in the logs.

2.3. Lexical Approaches. These approaches add words to the query using language properties such as synonyms and other linguistic word relationships. For instance, systems introduced in [5-7] adopt these approaches. Applying them, query expansion systems discover semantic relationships between words typically based on thesauri, then expand the queries for preventing from omissions of the words that are similar to the user's query words by using those relationships. So the words of the queries expanded in that way might have linguistic interpretations. We need to prepare the thesaurus database such as WordNet if we use it.

The approach proposed in this research is one of such approaches.

3. Methodology. As we mentioned before, searching for information on the Internet is very difficult for novice users. The key reason is in the ambiguity of queries typed by users. Such queries will typically have multiple meanings. For example, the word "dog" has several meanings: the animal's name, hotdog, and so on. Search engines retrieve docu-

ments for all meanings. So, ambiguous queries return a large number of non-relevant documents, and it is very difficult for the Internet users to find relevant information from a large list of URLs.

3.1. Tools used. The key idea of the proposed solution is to disambiguate and expand user's queries automatically. In our research, we use synonyms for expanding queries. Our assumption is as follows: Query expansion with synonyms that have same meaning can help to generate precise queries. We apply this understanding utilizing the following tools: "Google AJAX Search API" [8], "WordNet" , "WordNet::SenseRelate::AllWords" [9], and "JWNL (Java WordNet Library)" [10].

Google AJAX Search API is one of the API tools used to get search results from the Google web search. We use it to get titles and URLs from Google search results to the user query.

WordNet is a large lexical database in English. WordNet has two main functionalities; one supports a dictionary access to words definitions. When a user inputs a term, the corresponding dictionary record is returned. Another functionality provides a user with means to access the net of words, that allows us to get synonyms. This is a key feature of WordNet. When a user inputs a term, relevant words (such as synonyms, hypernyms, hyponyms and so on) are returned in response.

Both of WordNet::SenseRelate::AllWords and JWNL are APIs of WordNet. WordNet::SenseRelate::AllWords is a Perl module for the word sense disambiguation. And JWNL is a Java wrapper for WordNet, it allows us to use the WordNet system in Java applications.

We use Java and Perl to implement our prototype.

3.2. System Architecture. The main scenario for users to utilize the prototype is: first, a user inputs a query consisting of two or more terms (because the system uses relationships between word's senses to disambiguate word's senses, therefore, if there is only one term, the system can't disambiguate and expand queries) These terms go to the finding synonyms method. It works as follows: at the beginning, WordNet::SenseRelate::AllWords disambiguates the sense of each term. Then a synonym that has the same sense is found using WordNet for each term. If the synonym can't be found for a term, that term is not expanded. And that, each term in the user's query is expanded, the number of terms in the query is twice as many as the initial terms in the query. Then

the system moves the expanded query to the searching function. Finally, we send the query to the Google web search, and results of retrieval are returned to a user. These steps are done automatically.

4. Evaluation. Evaluation of the performance of the system is a labor-intensive and expensive process. Organizers of TREC [14] and NTCIR [15] conferences involve experts for the evaluation.

Our evaluation method is a human inspection of retrieved results by the prototype in response to the test set of queries. The number of assessors is one. The number of queries is 100. We understand that such an evaluation is subjective.

4.1. Queries for Tests. To prepare our tests, we collected 100 queries using Google insights for Search [11] and Google AdWords Keyword Tool [12]. We collected the most popular 10 queries in the United States of America for the period from January 2004 to January 2011 from Google insights for Search. Each of queries consists of one term. And then we collected the most relative 10 queries that has two or more terms based on each of the most popular queries using Google AdWords Keyword Tool. All queries were in English.

4.2. Accuracy Evaluation. We measured the quality of our prototype with the average of variation in page contents. We checked the top 8 pages of the search results for user query and the search results for query using our prototype, and measured whether the contents of the page include relevant information, such as explanation of terms. The results are shown in Table 1.

Table 1. Results of the evaluation

	Average precision
User queries	36% (285/800)
Our Prototype	29% (233/800)

Then we tested the results using t-test [13] to see the difference. The mean of Group of user queries minus the mean of Group of our prototype equals 0.06500. 95% confidence interval of this difference is (0,01884; 0,11116). The two-tailed P value equals $0,0063 < 0,01$. By conventional criteria, this difference is considered to be statistically significant.

The results produced by our prototype include about 7% fewer relevant pages than the results for original user queries. Especially for cases of queries with proper nouns, such as names, our prototype can't perform well. 51 queries (51%) include proper nouns. And also the number

of queries without expansion by our prototype are 34 (34%) in this evaluation. On the other hand, for some cases of queries without proper nouns, such as “injury lawyers for you”, our prototype demonstrated better precision compared to the search results for original user queries. The results without proper nouns are shown in Table 2.

Table 2. Results without proper nouns


	Average precision
User queries	27% (93/344)
Our Prototype	28% (95/344)

5. Conclusion. Our prototype demonstrated a semantic approach to expand queries utilizing WordNet and its API. The preliminary tests show its good performance with queries without proper names. And the system often needs too much time to expand the query. It is bad for the users, especially for beginners.

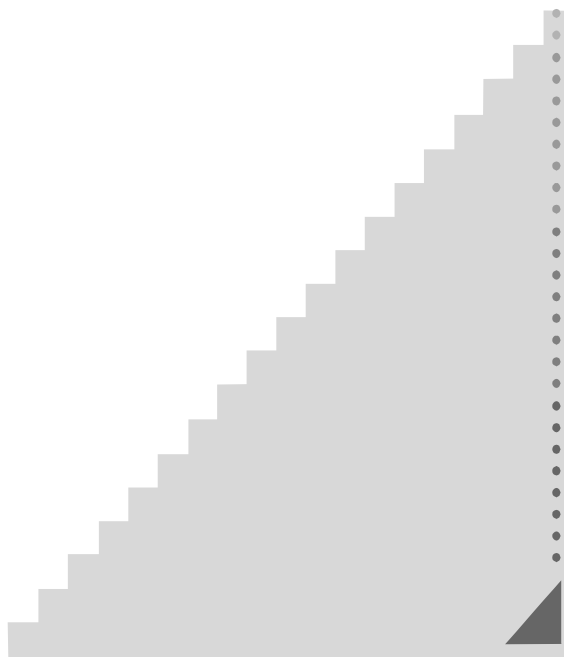
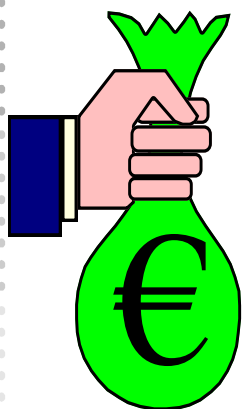
References

1. Google. Features: Autocomplete. <http://www.google.com/support/websearch/bin/answer.py?hl=en&answer=106230>
2. Princeton University. About WordNet. <http://wordnet.princeton.edu>
3. David E. Losada. Statistical query expansion for sentence retrieval and its effects on weak and strong queries. // Information Retrieval. USA, MA: Kluwer Academic Publishers Hingham, 2010. Vol. 13, No 5. P. 485–506.
4. Youngjoong Ko. Pseudo-relevance feedback and statistical query expansion for web snippet generation. // Information Processing Letters. Netherlands:Elsevier North-Holland, Inc. Amsterdam, 2008. Vol. 109. P. 18–22.
5. Tokunaga Takenobu. Query expansion using LMF-compliant lexical resources // Proceedings of the 7th Workshop on Asian Language Resources (ALR-7), 2009. P. 145–152.

6. Francisco J.P. Joining automatic query expansion based on thesaurus and word sense disambiguation using WordNet // International Journal of Computer Applications in Technology, 2008. Vol. 33, No 4. P. 271–279.
7. Lili Kotlerman. Directional distributional similarity for lexical expansion // Proceedings of the ACL-IJCNLP 2009 Conference Short Papers, 2009. P. 69–72.
8. Google. Google AJAX Search API Tutorial.
<http://code.google.com/edu/ajax/tutorials/ajax-search-api.html>
9. Ted Pedersen. WordNet::SenseRelate.
<http://www.d.umn.edu/~tpederse/senserelate.html>
10. Brett Walenz. JWNL (Java WordNet Library).
<http://sourceforge.net/projects/jwordnet/>
11. Google. Google Insights for Search.
<http://www.google.com/insights/search/>
12. Google. Keyword Tool – Google AdWords – Online Advertising by Google. <http://adwords.google.com/select/KeywordToolExternal>
13. Student. The probable error of a mean // Biometrika, 1908. Vol. 6, No. 1. P. 1–25.
14. Texas Real Estate Commission. Available:
<http://www.trec.state.tx.us/>
15. NTCIR Project. <http://research.nii.ac.jp/ntcir/index-en.html>



*5. Управление
социально-
экономическими
системами*



Андриенко В.А.

Санкт-Петербургский государственный университет

Об одной оптимизационной задаче в модели межотраслевого баланса

Рекомендовано к публикации профессором Смирновым Н.В.

Введение. В настоящее время в органах государственного управления и, особенно, в Правительстве РФ наблюдается повышенный интерес к модели межотраслевого баланса. Об этом свидетельствует, например, распоряжение Правительства РФ от 14 февраля 2009 г. № 201-р, предписывающее Росстату осуществлять разработку базовых таблиц «затраты-выпуск» на регулярной основе один раз в пять лет. Кроме того, как сообщает Российская бизнес-газета (№ 712), в ряде субъектов РФ самостоятельно возвращаются к составлению межотраслевых балансов (МОБ): «О целесообразности составления и реализации МОБов шла речь на состоявшемся в июне 2009 г. круглом столе по сотрудничеству между российскими регионами. По прозвучавшим на этом форуме экспертным оценкам использование МОБов в экономической политике и хозяйственной деятельности может минимум на 20% повысить рентабельность отраслей и, самое меньшее, на 15% увеличить доходы по внешнеэкономическим связям».

Необходимо также заметить, что опыт стран с развитой рыночной экономикой, использующих таблицы «затраты-выпуск» на протяжении более 50-ти лет (например, США) [1] показывает, что составление МОБ на регулярной основе – объективная необходимость, позволяющая эффективно управлять рыночной макроэкономической системой региона и государства в целом. Таким образом, модель межотраслевого баланса, отвергнутая в нашей стране на волне рыночных преобразований как элемент плановой экономики, переживает в настоящее время второе рождение. Прежде всего это связано с очевидными диспропорциями в развитии различных отраслей народного хозяйства и пониманием того, что на основе данной модели можно эти диспропорции выявить и устранить. Для этой цели представляется возможным ставить и решать различные оптимизационные задачи в модели МОБ.

В данной статье в качестве примера рассматривается постановка и решение одной из оптимизационных задач в гипотетической мо-

дели МОБ некоторого региона. Для большей наглядности, задача рассматривается в связи с актуальными на сегодняшний день проблемами повышения энергоэффективности и модернизации российской экономики, рассматриваются последствия внедрения технологических инноваций на предприятиях. Кроме того, показано, как на основе модели МОБ рассчитать оптимальные объёмы валовых выпусков отраслей и оптимальный уровень цен на продукцию, оценить экономическую эффективность внедрения технологических инноваций, выявить причины и проследить направленность инфляционных процессов, определить необходимый объём государственного субсидирования отраслей.

Описание макроэкономической системы на основе модели МОБ. Распределение продукции каждой отрасли в межотраслевом балансе экономики страны описывается следующим соотношением [2]:

$$x_i = \sum x_{ij} + f_i, \quad i, j = 1, \dots, n, \quad (1)$$

где x_i – валовая продукция в натуральном исчислении, произведённая i -й отраслью; x_{ij} – поставки i -й отрасли в j -ю отрасль; f_i – конечная продукция i -й отрасли. В состав конечной продукции f_i входит потребление, создание запасов, инвестиции, экспорт, импорт, государственные расходы. Ради её получения в обществе и организуется многоотраслевое производство на основе специализации отраслей.

Обычно межотраслевая модель строится на следующих предположениях:

- в каждой отрасли имеется единственная технология производства;
- нормы производственных затрат не зависят от объёма выпускаемой продукции;
- не допускаются замещения в производстве одних видов продукции другими.

В силу данных предположений величина x_{ij} для всех i и j должна иметь следующий вид:

$$x_{ij} = a_{ij}x_j, \quad i, j = 1, \dots, n.$$

Величина a_{ij} называется коэффициентом прямых затрат. Этот коэффициент показывает, какое количество продукта i -й отрасли необходимо затратить на производство единицы продукты j -й отрасли.

Запишем уравнение (1) в матричной форме:

$$(E - A)X = F, \quad (2)$$

где E – единичная $(n \times n)$ -матрица, A – продуктивная технологическая $(n \times n)$ -матрица, X – n -мерный вектор валового выпуска продукции, F – n -мерный вектор конечной продукции. Уравнение (2) называется *экономико-математической моделью МОБ* или *моделью Леонтьева «затраты-выпуск»*, записанное в натуральной форме учёта продукции.

Уравнение (2) может быть записано также в стоимостной форме учёта единицы продукции:

$$(E - A^T)P = R, \quad (3)$$

где E, A – матрицы из (1), P – n -мерный вектор цен на единицу продукции, R – n -мерный вектор долей добавленной стоимости на единицу продукции. Уравнение (3) называется *уравнением равновесных цен*.

Как видно из (3), конечная цена P на единицу продукции складывается из цен $A^T P$ на потребляемые отраслью ресурсы и величины добавленной стоимости R , причём последняя зависит от величины входящих в неё компонент, а именно:

$$R = r_{sal} + r_{tall} + r_{grant} + r_{credit} + r_{invest} + r_{netprof}, \quad (4)$$

где r_{sal} – доля заработной платы; r_{tall} – доля налогов; r_{grant} – доля дотаций (субсидий, субвенций); r_{credit} – доля кредитов; r_{invest} – доля амортизационных отчислений; $r_{netprof}$ – доля чистой прибыли (за вычетом амортизационных отчислений).

Вектор R может быть записан также в виде суммы затратной и доходной части: $R = r_{cost} + r_{prof}$. Очевидно, что вектор r_{cost} включает в себя все компоненты, кроме r_{grant} и $r_{netprof}$. Введём в рассмотрение n -мерный вектор рентабельности отраслей $profit$. Его компоненты могут быть вычислены по следующей формуле:

$$profit_i = r_{prof\ i} / (A_i^T P + r_{cost\ i}). \quad (5)$$

Здесь A_i^T – i -я строка матрицы A^T .

Таким образом, со стороны государства на макроэкономическом уровне возможно управляющее воздействие средствами денежно-кредитной и налогово-дотационной политики на рентабельность и инвестиционную активность отраслей, причём вектор R можно рассматривать в качестве управляющего параметра (переобозначим его через R_{dir}).

Замечание. Несмотря на очевидные достоинства модели МОБ в области исследования макроэкономики, следует признать что она, как и всякая модель, имеет свои недостатки. Основным недостатком данной модели является значительный временной лаг между сбором и обработкой информации и составлением модели МОБ, который увеличивает вектор ошибки управления системой. Данная проблема будет решаться по мере развития информационных технологий.

Постановка и решение оптимизационной задачи в модели МОБ. В качестве одного из вариантов постановки прямой задачи линейного программирования в модели МОБ можно рассмотреть задачу поиска оптимального вектора валовых мощностей народного хозяйства, обеспечивающего выпуск конечной продукции в количестве не менее ожидаемого спроса C [3]:

$$\begin{aligned} \min Z(X), \quad Z(X) &= R_{dir1}X_1 + \dots + R_{dirN}X_N, \\ (E - A)X &= F, \\ F &\geq C, \quad X \geq 0, \end{aligned} \tag{6}$$

Двойственной к (6) является задача поиска оптимального вектора цен на продукцию отраслей, обеспечивающего выгодность производственной деятельности:

$$\begin{aligned} \max Y(P), \quad Y(P) &= F_1P_1 + \dots + F_NP_N, \\ (E - A^T)P &= R_{dir}, \\ P &\geq 0, \end{aligned} \tag{7}$$

Замечание. В рамках данной статьи, в целях упрощения модели, вводится лишь по одному ограничению в каждой задаче. В реальной жизни, ввиду сильной корреляции между экономическими параметрами, следует учитывать множество других ограничений

(ограничения на трудовые ресурсы, инвестиционную и внешнеторговую деятельность, инфляционные ожидания и т.д.), оказывающих серьёзное влияние на формирование оптимальных значений вектора валовых мощностей и рыночных равновесных цен [4].

Рассмотрим гипотетическую модель МОБ некоторого региона, не учитывая его внешнеторговую деятельность. Пусть для экономики данного региона основными являются следующие отрасли: энергетика (ЭН), добыча сырья (ДС), сельское хозяйство (СХ), строительство (СТ), металлургия (МТ), машиностроение (МШ), пищевая промышленность (ПП). Технологическая матрица, ожидаемый спрос на продукцию конечного потребления (КП, в млн. руб.), а также вектор R_{dir} и его компонента r_{prof} (в тыс. руб.) представлены в следующей таблице.

Таблица. Технологическая матрица, ожидаемый спрос, R_{dir}, r_{profit}

	ЭН	ДС	СХ	СТ	МТ	МШ	ПП	КП
ЭН	0,1	0,2	0,1	0,1	0,15	0,15	0,1	12500
ДС	0,2	0,1	0,05	0,1	0,2	0,05	0,05	5600
СХ	0,01	0,01	0,2	0,01	0,01	0,01	0,25	3300
СТ	0,1	0,1	0,1	0,1	0,15	0,15	0,1	11800
МТ	0,1	0,1	0,05	0,15	0,15	0,2	0,05	6400
МШ	0,15	0,15	0,05	0,15	0,1	0,1	0,05	7800
ПП	0,04	0,04	0,02	0,03	0,04	0,04	0,03	4200
R_{dir}	5	2,3	1,2	7	3,2	4,4	1,8	
r_{prof}	2,7	1,9	0,6	3,7	1,9	2,5	1,0	

Таким образом, мы имеем все необходимые данные для решения поставленной задачи. Оптимальный вектор валовых мощностей X_{opt} (в млн. руб.) и оптимальный вектор цен P_{opt} (в тыс. руб.) найдём с помощью пакета средств оптимизации Excel Solver: $X_{opt} = (34823; 26363; 9221; 31487; 28099; 28321; 10343)^T$ и соответственно $P_{opt} = (14, 4; 11, 94; 8, 02; 15, 82; 14, 1; 14, 22; 9, 11)^T$. Объём продукции в натуральном исчислении легко узнать по формулам $X_{nati} = X_{cost i} / P_i$, $i = 1, \dots, n$. Используя (5), вычислим установившуюся рентабельность отраслей (%):

$$profit = (23, 08; 18, 92; 8, 08; 30, 53; 15, 6; 21, 33; 12, 33)^T.$$

Замечание. Поскольку в условиях рыночной экономики государство не должно напрямую вмешиваться в процесс ценообразования, вектор оптимальных цен может быть использован, например,

в качестве важной информации для ФАС, позволяющей выявлять ценовые сговоры на монопольных рынках.

Из теории двойственности и анализа чувствительности известно, что существуют интервалы изменения коэффициентов целевой функции (в данной постановке – R_{dir}), при которых текущее оптимальное решение прямой задачи сохраняется [5]. Однако необходимо помнить о том, что это повлечёт за собой изменение оптимальных переменных двойственной задачи, т.е. вектора P_{opt} . Применив достаточное правило оптимальности, можно увидеть, что допустимые интервалы изменения вектора R_{dir} для данных отраслей лежат в интервале от 0 до $+\infty$. Это означает, что при изменении вектора R_{dir} в данных пределах вектор X_{opt} не изменится.

Оценка экономической эффективности внедрения технологических инноваций. Пусть в результате внедрения инновационных технологий на части предприятий машиностроительной отрасли её общая энергоэффективность выросла на 67%. Этому соответствует изменение элемента a_{16} технологической матрицы A . Теперь он стал равен 0,1. Тогда новые значения векторов P_{opt} , $profit$ будут $P_{opt} = (14, 07; 11, 62; 7, 83; 15, 51; 13, 78; 13, 19; 8, 91)^T$ и $profit = (23, 75; 19, 55; 8, 3; 31, 33; 15, 99; 23, 39; 12, 64)^T$. Общая экономическая эффективность от внедрения энергосберегающих технологий в машиностроительной отрасли при условии оптимальных валовых выпусков отраслей составит 20 939 млн.руб. Таким образом, данный пример подтверждает, что интенсивному пути развития отраслей соответствует уменьшение значений элементов матрицы A .

Изучение инфляционного процесса с помощью модели МОБ. Предположим, что в результате ценового сговора, энергетическая отрасль решила увеличить свою чистую прибыль на 18,5% за счёт повышения цен на свою продукцию. В этом случае, повысится на 0,5 тыс.руб. доля чистой прибыли данной отрасли, а её рентабельность увеличится на 3,86%. Однако, это приведёт к повышению оптимального уровня цен на продукцию всех отраслей на $\Delta P = (0, 65; 0, 26; 0, 17; 0, 2; 0, 26; 0, 23; 0, 18)^T$, а также к снижению рентабельности всех отраслей, кроме первой, на $\Delta profit = (3, 86; -0, 51; -0, 2; -0, 53; -0, 34; -0, 5; -0, 28)^T$. Очевидно, что данные отрасли будут вынуждены вновь поднимать цены на свою продукцию, стремясь как минимум восстановить свою прежнюю рентабельность. Однако, это вызовет лишь очередной виток ин-

фляции. Очевидно, что к инфляции издержек приводит увеличение любой компоненты вектора R_{dir} , а также увеличение значения элементов технологической матрицы A . Эти выводы могут показаться очевидными, однако следует подчеркнуть, что модель МОБ позволяет делать точные количественные оценки и отслеживать направленность инфляционных процессов.

Управление за счёт государственного субсидирования отраслей. Предположим теперь, что сельскохозяйственная отрасль из-за конкуренции с внешними поставщиками не может поднять цены на свою продукцию и восстановить прежнюю рентабельность. Это может привести к её полному разорению и внешнеэкономической зависимости региона. В целях обеспечения продовольственной безопасности правительство региона решило восстановить рентабельность данной отрасли до прежнего уровня. После вычислений получим необходимый объём государственного субсидирования отрасли, равный 17 085 236 руб. Важно, что при этом цена на продукцию данной отрасли остаётся без изменения, а это, в свою очередь, не порождает инфляционных процессов в других отраслях.

Заключение. Таким образом, на конкретных примерах было показано, что вектор R_{dir} действительно можно рассматривать в качестве средства управления макроэкономической системой, а модель межотраслевого баланса необходимо использовать в целях государственного регулирования рыночной экономики.

Литература

1. Леонтьев В.В. Избранные произведения. В 3-х томах. М.: Экономика, 2006. 1364 с.
2. Прасолов А.В. Математические методы экономической динамики. СПб.: Лань, 2008. 352 с.
3. Ефимов В.А. Методология экономического обеспечения демографической политики устойчивого развития. СПб.: СЗАГС, 2007. 184 с.
4. Тарасевич Л.С., Гребенников П.И., Леусский А.И. Макроэкономика: Учебник. М.: Высшее образование, 2007. 654 с.
5. Хемди А. Таха Исследование операций. М.: Вильямс, 2007. 912 с.

Афанасьева Е.В.

Санкт-Петербургский государственный университет

Моделирование распределения исследователей по областям науки с помощью вероятностных цепочек

Рекомендовано к публикации доцентом Ампиловой Н.Б.

1. Введение. Всякий раз, когда исследователю приходится иметь дело со статистическими данными так или иначе появляется понятие вероятностного вектора, который может отражать, например, распределение некоторого ресурса между слоями общества, групп населения между сферами деятельности и т.д. Последовательности таких вероятностных векторов, рассматриваемых за каждый год из некоторого промежутка времени, образуют вероятностные цепочки, которые могут служить инструментом моделирования поведения социальных систем, а также получения статистических оценок относительной динамики социально-экономических процессов.

Дадим определение вероятностной цепочки, следуя [3], где приведены основные свойства этого понятия.

Определение 1. Дискретная вероятностная $(1, n)$ -цепочка – это последовательность векторов вида

$$\bar{p}_t = \begin{pmatrix} p_{1t} \\ \vdots \\ p_{nt} \end{pmatrix}, \quad t = 0, 1, \dots,$$
$$0 \leq p_{kt} \leq 1, \quad \sum_{k=1}^n p_{kt} = 1.$$

Любую нелинейную вероятностную цепочку можно представить набором строго положительных порождающих функций – преобразований над вероятностными векторами. Так, i -ый элемент вероятностного вектора в момент времени t будет иметь вид

$$p_{k,t+1} = \frac{P_k(p_0, \dots, p_t)}{\sum_{i=1}^n P_i(p_0, \dots, p_t)},$$

где $P_k(p_0, \dots, p_t)$ – строго положительная порождающая функция, и $k = 1, \dots, n$, $t = 0, 1, \dots$

В зависимости от порождающих функций, которыми они задаются, выделяют различные виды вероятностных цепочек. В данной работе используются логарифмически-линейные вероятностные цепочки, для которых порождающие функции выглядят следующим образом

$$P_k(p) = A_k p_1^{a_{k1}} \cdot \dots \cdot p_n^{a_{kn}}, \quad -\infty \leq a_{kj} \leq +\infty, \quad A_k > 0, \quad k = 1, \dots, n,$$

где A_1, \dots, A_n – суммарные преимущества групп $1, \dots, n$, а элементы матрицы $[a_{ij}]_{i,j=1}^n$ отражают суммарные характеристики «гибкости» относительного прироста долей данных групп. Так, например, положительное значение коэффициента a_{ij} говорит о том, что увеличение доли j -ой группы влечет за собой увеличения доли i -ой группы, и наоборот. Отрицательное значение свидетельствует о том, что, напротив, при увеличении доли j -ой группы доля i -ой группы снижается, и наоборот. В литературе, где описывается применение этой модели, такие отношения между группами i и j называются соответственно дополняющими и конкурирующими.

Таким образом, логарифмически-линейные вероятностные цепочки могут быть представлены в следующем виде:

$$p_{k,t+1} = \frac{A_k p_{1,t}^{a_{k1}} \cdot \dots \cdot p_{n,t}^{a_{kn}}}{\sum_{s=1}^n A_s p_{1,t}^{a_{s1}} \cdot \dots \cdot p_{n,t}^{a_{sn}}}, \quad (1)$$

где $k = 1, \dots, n$, $t = 1, \dots, T$. Согласно социально-экономической интерпретации данной модели, предложенной в [3], под n может подразумеваться количество групп населения или территорий, между которыми распределяется некоторый ресурс, или, например, количество территорий или сфер деятельности, между которыми распределяется популяция, и т.д. Впервые эта модель была описана в работе [2], а впоследствии неоднократно применялась в исследованиях распределения валового регионального продукта по регионам таких стран, как США, Бразилия, Греция, Индонезия, Испания и Португалия [1].

В данной работе описанная модель применяется к изучению того, каким образом в Российской Федерации происходит распределение исследователей между областями научного знания.

2. Постановка задачи. Для построения модели, основанной на логарифмически-линейных вероятностных цепочках, использовались данные Федеральной службы государственной статистики РФ за 1994–2009 гг. по числу исследователей в естественных, технических, медицинских, сельскохозяйственных, общественных и гуманитарных науках (всего шесть областей). Под исследователями здесь подразумеваются работники, профессионально занимающиеся исследованиями и разработками и непосредственно осуществляющие создание новых знаний, продуктов, методов и систем, а также управление указанными видами деятельности.

3. Построение модели. Перед построением модели перепишем систему (1) в более удобном виде

$$\begin{cases} p_{1,t+1} = \frac{1}{\sum_{j=2}^n A_j p_1^{a_{j1}} \cdot \dots \cdot p_n^{a_{jn}}}, \\ \frac{p_{k,t+1}}{p_{1,t+1}} = A_k p_1^{a_{k1}} \cdot \dots \cdot p_n^{a_{kn}}, \end{cases}$$

где k , t и p_t принимают те же значения, что и в (1). Не умаляя общности, можно рассматривать изменение долей исследователей в различных областях науки относительно соответствующих изменений доли для первой области, а в данном случае – для естественных наук. Динамику изменения долей исследователей во всех областях науки, кроме первой, удобно переписать следующим образом

$$\ln p_{k,t+1} - \ln p_{1,t+1} = \ln A_k + \sum_{j=1}^n a_{kj} \ln p_{j,t}, \quad (2)$$

где $k = 2, \dots, n$, $t = 1, \dots, T$. Система (2) представляет собой систему внешне не связанных уравнений (Seemingly unrelated regressions, SUR), т.е. такую систему, которая характеризуется наличием корреляций между ошибками в ее уравнениях. Для оценивания подобных систем используется трехшаговый метод наименьших квадратов, однако выделяют два условия, при которых можно использовать обычный метод наименьших квадратов, не проигрывая при этом в эффективности оценки: во-первых, когда ошибки не коррелируют между уравнениями системы, и, во-вторых, когда во всех уравнениях системы присутствуют одни и те же объясняющие переменные.

В нашем случае выполняется второе условие, так как в каждом из уравнений системы (2) в правой части участвуют переменные $\ln p_{1,t}, \dots, \ln p_{6,t}$. Следовательно, для оценки неизвестных параметров можно использовать обычный метод наименьших квадратов.

4. Оценивание модели. Оценка параметров исследуемой модели методом наименьших квадратов была проведена с помощью приложения, разработанного в рамках данного исследования специально для работы с некоторыми видами нелинейных вероятностных цепочек. Помимо непосредственной оценки неизвестных коэффициентов A_1, \dots, A_n и матрицы $[a_{ij}]_{i,j=1}^n$, осуществлялась также проверка основных предпосылок регрессионного анализа, поскольку их выполнение требуется для получения несмещенных, состоятельных и эффективных оценок. Средствами разработанного приложения и частично средствами статистического пакета SPSS была проведена проверка таких условий, как отсутствие автокорреляции случайных остатков, равенство нулю их математического ожидания, отсутствие корреляций с объясняющими переменными и постоянство дисперсии (гомоскедастичность), а также нормальность распределения случайных остатков.

Для выявления автокорреляций использовались тест Дарбина – Уотсона и метод рядов. Для проверки гомоскедастичности и отсутствия корреляций между случайными остатками и объясняющими переменными – тест ранговой корреляции Спирмена. Проверка нормальности распределения случайных остатков проводилась с помощью теста Шапиро – Уилка, а проверка равенства их математического ожидания – с помощью одновыборочного t -критерия Стьюдента.

Затем по критерию Фишера была осуществлена проверка значимости уравнений системы, за исключением второго и четвертого уравнений, поскольку для них тест Шапиро – Уилка выявил отклонения случайных остатков от нормального распределения на уровнях значимости 0,01 и 0,05 соответственно. Кроме того, для всех уравнений системы, кроме указанных, была проведена проверка значимости полученных коэффициентов по t -критерию Стьюдента.

5. Интерпретация результатов. На основе построенной модели была проведена интерполяция распределения исследователей по областям науки на промежуток с 1995 по 2009 год. Для иллюстрации ее точности на рис. 1, 2 приведены два графика, отражающие реальные и смоделированные доли исследователей соответственно.

На основе результатов проведенных проверок были выделены наиболее значимые коэффициенты взаимного влияния:

$$a_{6,1} = 7,375 (\alpha = 0,1), \quad a_{6,2} = 23,811 (\alpha = 0,1),$$

$$a_{6,6} = 1,361 (\alpha = 0,01).$$

Напомним, что в качестве исследуемых шести областей были взяты естественные, технические, медицинские, сельско-хозяйственные, общественные и гуманитарные науки, причем изменения долей исследователей в последних пяти сферах рассматривались относительно изменения доли исследователей в естественных науках. Приве-

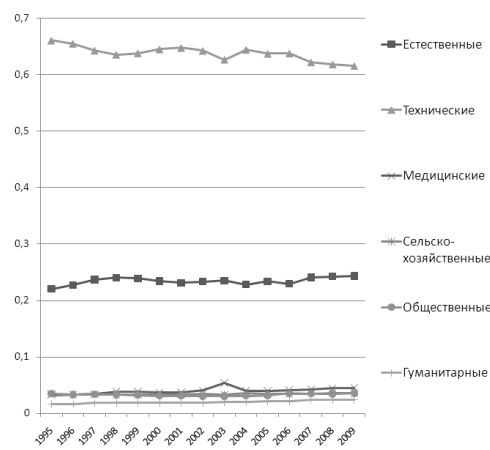


Рис. 1. Исходные данные за 1995–2009 гг.

денные коэффициенты взаимного влияния говорят о том, что доля исследователей в гуманитарных областях получает положительное входное влияние от самой себя, а также от долей исследователей в естественных и технических науках. Это означает, что увеличение доли исследователей естественнонаучных и особенно технических направлений приводит к увеличению доли исследователей в гуманитарной сфере, и наоборот. Кроме того, увеличение доли исследователей в гуманитарных науках приводит к еще большему увеличению доли исследователей в этой сфере.

Однако при этом, по данным за 2009 год, доля исследователей в гуманитарных науках сама по себе весьма незначительна (всего 2,4%), в то время как доли исследователей в естественно-научной и технической областях составляют соответственно 24,3% и 61,6%.

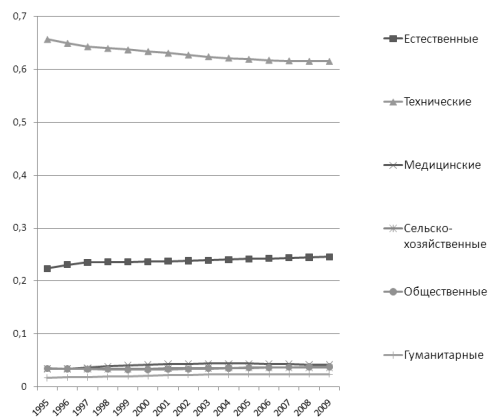


Рис. 2. Смоделированные данные за 1995–2009 гг.

Литература

1. Dall'erba S. Competition, complementarity and increasing disparities among the regions of Spain and Portugal // *Revue d'Economie Regionale et Urbaine*, 2004. № 2. P. 311–330.
2. Dendrinos D.S., Sonis. M. Nonlinear discrete relative population dynamics of the U.S. regions // *Journal of Applied Mathematics and Computation*, 1988. № 22. P. 25–44.
3. Sonis M. Discrete non-linear probabilistic chains // *Functional differential equations*, 2003. Vol. 10. № 3–4. P. 593–639.

Барченко М.К.

Санкт-Петербургский государственный университет

Некрамеровская модель разорения страховой компании¹

Рекомендовано к публикации доцентом Иголкиным В.Н.

При исследовании деятельности страховой компании рассматривается эволюция капитала страховой компании $u(t) = u + \xi_t - \eta_t$, где ξ_t – поток страховых премий, η_t – поток страховых исков. Эти случайные процессы, очевидно, зависимы. В принятой модели капитал страховой компании $u(t) = u + ct - \sum_k x_k = u + \sum_k (c\tau_k - x_k)$, т.е. поток премий является детерминированным с постоянной интенсивностью. Этот случайный процесс $u(t)$ называют процессом риска. Он рассматривался многими авторами, в основном, шведской школы. Хороший обзор этих работ дан в книге [1]. Связь между ξ_t и η_t в этой модели выражена через средние значения τ_k и x_k .

Рассмотрим некрамеровскую модель, в которой поступление страховой премии связано с поступлением иска. Эта модель предложена в работе [2]. Разберем ее более детально и постараемся довести до вычислительного алгоритма.

Пусть имеется последовательность интервалов определенной длины τ . Имеется простейший поток с параметром λ клиентов, вносящих случайные премии с распределением $r(x)$. Каждый клиент инициирует приход иска. Используя терминологию теории массового обслуживания, будем говорить, что пришедший клиент становится на обслуживание с экспоненциальным временем обслуживания $\mu e^{-\mu t}$. Рассматриваемая система имеет бесконечное число приборов. Для вычисления капитала нужно знать число необслуженных клиентов, равное разности числа пришедших и обслуженных. Распределение числа пришедших на интервале длины τ является Пуассоновским и определяет величину X_n суммарную премию. Будем считать, что все расчеты происходят в конце интервала. Тогда на интервале τ поступает случайное число ξ клиентов $\xi = 0, 1, \dots$ и $P\{\xi = n\} = \frac{(\lambda\tau)^n}{n!} e^{-\lambda\tau}$. Полученная за время τ премия X также является случайной и ее

¹Работа выполнена при финансовой поддержке РФФИ, проект № 09-01-00360

распределение $Q(x)$ равно

$$Q(x) = \sum_{n=0}^{\infty} \frac{(\lambda\tau)^n}{n!} e^{-\lambda\tau} r^{*n}(x),$$

где $r^{*n}(x)$ означает n -кратную свертку функции $r(x)$. Будем рассматривать систему в моменты $\tau, 2\tau, \dots$ и характеризовать ее в эти моменты текущим капиталом и числом необслуженных клиентов.

Обозначим $P_{k,m}(\tau)$ – вероятность прихода m заявок и ухода $m-k$ на интервале $(0, \tau)$.

$$P_{k,m}(\tau) = e^{-\lambda\tau} \lambda^m \int_{D_m} \sum_C \prod_{i=1}^k e^{-\mu(\tau-x_{s_i})} \times \\ \times \prod_{s=1, s \neq s_i}^m (1 - e^{-\mu(\tau-x_s)}) dx_1 \dots dx_m.$$

Здесь C произвольное сочетание номеров s_1, \dots, s_k по k из чисел $(1, \dots, m)$. После преобразований

$$P_{k,m}(\tau) = e^{-\lambda\tau} \frac{\lambda^m}{k!(m-k)!} \left(\frac{1 - e^{-\mu\tau}}{\mu} \right)^k \left(\tau - \frac{1 - e^{-\mu\tau}}{\mu} \right)^{m-k}.$$

Пусть в начале рассматриваемого (n по порядку) интервала имеется l_n клиентов, и из числа этих клиентов к концу интервала будет обслужено l_1 . Обозначим ξ – число обслуженных клиентов к концу интервала τ , ξ принимает значения $0, 1, \dots, l_n$, q_{l_1, l_n} – вероятность получить l_1 обслуженных клиентов. Для нахождения q_{l_1, l_n} можно построить следующую систему дифференциальных уравнений

$$\begin{cases} \dot{q}_{0, l_n} = -l_n \mu q_{0, l_n}, \\ \dot{q}_{1, l_n} = -(l_n - 1) \mu q_{1, l_n} + l_n \mu q_{0, l_n}, \\ \dots \\ \dot{q}_{j, l_n} = -(l_n - j) \mu q_{j, l_n} + (l_n - j + 1) \mu q_{j-1, l_n}, \\ \dots \\ \dot{q}_{l_1, l_n} = \mu q_{l_1-1, l_n} \end{cases}$$

с начальными условиями $q_{0, l_n}(0) = 1$, $q_{j, l_n}(0) = 0$ для остальных j . Решая ее, получаем

$$q_{l_1, l_n} = C_{l_n}^{l_1} \sum_{i=0}^{l_1} (-1)^i C_{l_1}^i e^{-(l_n - l_1 + i)\mu\tau}.$$

В течение интервала пришло m клиентов, из которых k осталось необслужены. Тогда в конце интервала будет $m - k + l_1$ обслуженных клиентов.

Вычислим теперь положительный X и отрицательный Y вклады в капитал за интервал длины τ . Имеющийся капитал на начало интервала обозначим $U_n = U(n\tau)$. Для капитала имеем рекуррентное соотношение $U_{n+1} = U_n + X_n - Y_n$. Условные распределения X_n и Y_n равны соответственно $r^{*m}(x)$ и $f^{*(m-k+l_1)}(x)$. Тогда $X_n - Y_n$ имеет распределение

$$\phi_{m,k,l_1} = \int_{-\infty}^{\infty} r^{*m}(u) f^{*(m-k+l_1)}(u+x) du.$$

Распределение приращения капитала на этом интервале при фиксированном начальном или конечном состоянии равно

$$\phi_{l_n, l_{n+1}}(x) = \sum_{l_{n+1}=l_n-l_1+k} \phi_{m,k,l_1}(x) P_{k,m}(\tau) q_{l_1, l_n}(\tau).$$

Суммирование ведется по m от 0 до ∞ , по k от 0 до m , по l_1 от 0 до l_n при условии $l_{n+1} = l_n - l_1 + k$.

Вероятности неразорения в двух соседних моментах времени связаны следующим рекуррентным соотношением

$$P_{n+1, l_n}(U) = \sum_{l_{n+1}} \int_0^{\infty} \phi_{l_n, l_{n+1}}(x-U) P_{n, l_{n+1}}(x) dx.$$

Полагая $l_n = j$, $l_{n+1} = k$, можно переписать эти соотношения в виде

$$P_{n+1, j}(U) = \sum_k \int_0^{\infty} \phi_{j, k}(x-U) P_{n, k}(x) dx.$$

Начиная с $P_{0, j}(U) = 1$, можно найти монотонно убывающую последовательность $P_{n, j}(U)$. Она имеет предел. Если этот предел равен нулю, то соответствующая вероятность неразорения на бесконечном интервале равна нулю.

Если перейти к пределу $n \rightarrow \infty$, то для вероятностей неразорения на бесконечном интервале получим систему линейных интегральных уравнений

$$P_j(U) = \sum_k \int_0^{\infty} \phi_{j, k}(x-U) P_k(x) dx, \quad j = 0, 1, \dots$$

Интересно, что эта модель сводится к Марковскому варианту основной модели Лунберга – Крамера, но с бесконечным числом состояний. Введем в рассмотрение марковские события (состояния марковской цепи), j, k – число необлуженных заявок в начале и конце интервала. Тогда условные вероятности переходов

$$\pi_{j,k} = \int_{-\infty}^{\infty} \phi_{j,k}(x) dx,$$

и система линейных интегральных уравнений запишется в виде

$$P_j(U) = \sum_k \pi_{j,k} \int_0^{\infty} \frac{\phi_{j,k}(x-U)}{\pi_{j,k}} P_k(x) dx, \quad j = 0, 1, \dots$$

В вычислительном отношении рассматриваемая модель приводит к очень сложным вычислениям. В то же время структура программы нахождения вероятности неразорения в этой модели методом Монте-Карло (моделирование процесса) несложна.

Рассмотрим следующий случай: страховая премия является постоянной $(1 + \theta)m = \text{const}$, а выплата по иску имеет дискретное распределение (см. таблицу).

Таблица. Распределение выплаты по иску

p_i	3/4	1/8	1/16	1/16
x_i	0	m	$8m$	$6m$

Функция плотности интервала между поступлением страховых премий имеет вид $\lambda e^{-\lambda t}$. Каждый договор обслуживается в течение случайного интервала с функцией плотности $\mu e^{-\mu t}$. Также имеется последовательность интервалов $t_0 = 0, t_1 = \tau, \dots, t_N = T = N\tau$, где N – число рассматриваемых интервалов. Для примера можно взять $\tau = 1/4$ года, а $T = 10$ лет, тогда $N = 40$.

Для получения выборочных значений случайной величины с экспоненциальным распределением используем генератор случайных чисел α , равномерно распределенных на интервале $(0, 1)$, а именно, $1 - e^{-\lambda x} = \alpha, x = -\frac{1}{\lambda} \ln(1 - \alpha)$.

Метод моделирования дискретной случайной величины основан на следующем соотношении [3]

$$P\left(\sum_{k=0}^{l-1} p_k \leq \alpha < \sum_{k=0}^l p_k\right) = p_l, \quad p_l = P(\xi = x_l).$$

В данном случае функция для ее получения такова:

```

float isk(float a, float m){
    if (a>=0 && a<3/4) return 0;
    if (a>=3/4 && a<7/8) return m;
    if (a>=7/8 && a<15/16) return 8*m;
    if (a>=15/16 && a<=1) return 6*m;
}

```

Строится несколько циклов. Моделируем моменты прихода клиентов $s_j, j = 1, 2, \dots$, пока $s_j = \sum_{i=1}^j \xi_i < T$, где ξ_i – значение случайной величины с функцией плотности $\lambda e^{-\lambda t}$, и моменты прихода исков $q_j = s_j + \eta_j$, где η_i – значение случайной величины с функцией плотности $\mu e^{-\mu t}$. Далее для каждого интервала считаем количество пришедших клиентов b_1 и исков b_2 и вычисляем общий капитал компании

```
u[k+1] = u[k] + b1*(1+theta)*m - b2*isk(alpha(),m);
```

Если при некотором k значение $u[k] < 0$, то считаем компанию разорившейся. В случае $u[N] \geq 0$ эксперимент удачен. Проводя h раз такие испытания и вычисляя количество удачных экспериментов h_1 , получаем эмпирическую вероятность неразорения страховой компании $P = \frac{h_1}{h}$.

Если взять значения параметров $\lambda = 10^3$, $\mu = 3/2$, $m = 5 \cdot 10^3$, $\theta = 0, 2$, $u_0 = 0$ и проводить по тысяче испытаний, то при нескольких запусках программы P примет значения 0,911, 0,933, 0,923. Можно рассмотреть и другие варианты премий и исков, алгоритм программы при этом существенно не изменится.

Литература

1. Grandell J. Aspects of risk theory (Springer series in statistics: probability and its applications). New York: Springer-Verlag, 1991. 175 p.
2. Иголкин В.Н. , Ковригин А.Б. Финансовые потоки и их флуктуации. СПб.: Изд-во СПбГУ, 2005. 134 с.
3. Ермаков С.М., Михайлов Г.А. Курс статистического моделирования. М.: Наука, 1976. 320 с.

Басков О.В.

Санкт-Петербургский государственный университет

Последовательный алгоритм построения двойственного конуса и его применение в принятии решений

Рекомендовано к публикации профессором Ногиньим В.Д.

1. Постановка задачи. Пусть заданы образующие выпуклого конечнопорожденного конуса

$$\mathbb{C} = \text{cone} \{c^1, \dots, c^p\} = \left\{ x \in \mathbb{R}^m \mid \exists \alpha_1, \dots, \alpha_p \geq 0 : x = \sum_{i=1}^p \alpha_i c^i \right\}.$$

Для упрощения записи набор $c^i, i \in \mathbb{K}$ будем обозначать $c^{\mathbb{K}}$, где \mathbb{K} – некоторое множество индексов. Задача построения двойственного конуса состоит в нахождении образующих конуса

$$\mathbb{D} = \{x \in \mathbb{R}^m \mid \forall y \in \mathbb{C} \ xy \geq 0\}$$

по известным $c^{\mathbb{I}}$, где $\mathbb{I} = \{1, \dots, p\}$.

2. Каноническое представление конуса. Так как целью данной работы является разработка алгоритма решения поставленной задачи, пригодного для программной реализации, необходимо задать способ хранения конуса в памяти машины.

Определение. Будем называть образующую a острого выпуклого конечнопорожденного конуса *существенной*, если $\text{cone} \{a\}$ является гранью этого конуса.

Напомним, что выпуклое множество \mathbb{G} называется гранью выпуклого множества \mathbb{F} , если $\forall x, y \in \mathbb{F} : \frac{x+y}{2} \in \mathbb{G} \Rightarrow x \in \mathbb{G}$. С учетом этого несущественные образующие оказываются представимыми в виде неотрицательной линейной комбинации других образующих, и поэтому их исключение не изменяет конуса.

Можно показать, что любой выпуклый конечнопорожденный конус \mathbb{F} представим в виде

$$\mathbb{F} = \text{cone} \{a^{\mathbb{J}}\} + \mathcal{L} \{b^{\mathbb{H}}\}, \quad (1)$$

где \mathbb{J} и \mathbb{H} – некоторые конечные, быть может, пустые, множества индексов. При этом конус $\text{cone}\{a^{\mathbb{J}}\}$ острый, а все его образующие существенны, попарно неколлинеарны и ортогональны $\mathcal{L}\{b^{\mathbb{H}}\}$. Представление (1) будем называть каноническим.

По ходу алгоритма необходимо хранить не только каноническое представление текущего конуса, но и информацию обо всех его гранях. Справедливо утверждение, согласно которому любая грань конуса (1) представима в виде $\text{cone}\{a^{\mathbb{K}}\} + \mathcal{L}\{b^{\mathbb{H}}\}$ для некоторого $\mathbb{K} \subseteq \mathbb{J}$. Отметим, что сам конус всегда является своей собственной гранью.

3. Алгоритм. Конус \mathbb{D} , двойственный к заданному конечнополющенному конусу $\mathbb{C} = \text{cone}\{c^{\mathbb{I}}\}$, представим в виде пересечения полупространств

$$\mathbb{D} = \bigcap_{i \in \mathbb{I}} \{x \in \mathbb{R}^m \mid c^i x \geq 0\}.$$

Будем последовательно строить конусы

$$\mathbb{D}_k = \bigcap_{i=1}^k \{x \in \mathbb{R}^m \mid c^i x \geq 0\},$$

начиная с $\mathbb{D}_0 = \mathbb{R}^m = \mathcal{L}\{e^1, \dots, e^m\}$, где e^i – орты \mathbb{R}^m . Сразу оговоримся, что если конус \mathbb{C} задан в виде суммы $\text{cone}\{c^{\mathbb{I}}\} + \mathcal{L}\{b^{\mathbb{H}}\}$, то следует взять $\mathbb{D}_0 = \mathcal{L}^{\perp}\{b^{\mathbb{H}}\}$. У конуса \mathbb{D}_0 единственная грань – он сам.

Опишем s -ый шаг алгоритма, на котором по каноническому представлению конуса

$$\mathbb{D}_{s-1} = \text{cone}\{d_{s-1}^{\mathbb{J}_{s-1}}\} + \mathcal{L}\{f_{s-1}^{\mathbb{H}_{s-1}}\}$$

строится конус $\mathbb{D}_s = \mathbb{D}_{s-1} \cap \{x \in \mathbb{R}^m \mid c^s x \geq 0\}$.

Сперва необходимо определить, ортогонален ли вектор c^s пространству $\mathcal{L}\{f_{s-1}^{\mathbb{H}_{s-1}}\}$. В зависимости от этого рассмотрим два случая.

Если $c^s \notin \mathcal{L}^{\perp}\{f_{s-1}^{\mathbb{H}_{s-1}}\}$, то пусть c_0^s – проекция c^s на $\mathcal{L}\{f_{s-1}^{\mathbb{H}_{s-1}}\}$. Далее полагаем $\mathbb{J}_s = \mathbb{J}_{s-1}$ и для $i \in \mathbb{J}_s$ строим

$$d_s^i = d_{s-1}^i - \frac{c^s d_{s-1}^i}{c_0^s c_0^s} c_0^s.$$

Каноническое представление конуса \mathbb{D}_s имеет вид $\mathbb{D}_s = \text{cone} \{d_s^{\mathbb{J}}\} + \mathcal{L} \{f_s^{\mathbb{H}_s}\}$, где $\mathcal{L} \{f_s^{\mathbb{H}_s}\} = \mathcal{L} \{f_{s-1}^{\mathbb{H}_{s-1}}\} \cap \mathcal{L}^\perp \{c_s^s\}$. При этом грани \mathbb{D}_s получаются из граней \mathbb{D}_{s-1} естественным образом: если $\text{cone} \{d_{s-1}^{\mathbb{K}}\} + \mathcal{L} \{f_{s-1}^{\mathbb{H}_{s-1}}\}$ для некоторого $\mathbb{K} \subseteq \mathbb{J}_{s-1}$ является гранью \mathbb{D}_{s-1} , то $\text{cone} \{d_s^{\mathbb{K}}\} + \mathcal{L} \{f_s^{\mathbb{H}_s}\}$ – грань \mathbb{D}_s .

Теперь обратимся к ортогональному случаю: $c^s \perp \mathcal{L} \{f_{s-1}^{\mathbb{H}_{s-1}}\}$. Множество \mathbb{J}_{s-1} разбивается на три:

$$\mathbb{A} = \{i \in \mathbb{J}_{s-1} \mid d_{s-1}^i c^s > 0\},$$

$$\mathbb{B} = \{i \in \mathbb{J}_{s-1} \mid d_{s-1}^i c^s < 0\}, \quad \mathbb{S} = \{i \in \mathbb{J}_{s-1} \mid d_{s-1}^i c^s = 0\}.$$

Вводится также множество

$$\mathbb{P} = \left\{ (i, j) \in \mathbb{A} \times \mathbb{B} \mid \text{cone} \{d_{s-1}^i, d_{s-1}^j\} + \mathcal{L} \{f_{s-1}^{\mathbb{H}_{s-1}}\} - \text{грань } \mathbb{D}_{s-1} \right\}.$$

Полагаем $\mathbb{J}_s = \mathbb{A} \cup \mathbb{S} \cup \mathbb{P}$. Для $(i, j) \in \mathbb{P}$ строим векторы

$$d_s^{(i,j)} = (c^s d_{s-1}^i) d_{s-1}^j - (c^s d_{s-1}^j) d_{s-1}^i.$$

Для $i \in \mathbb{A} \cup \mathbb{S}$ полагаем $d_s^i = d_{s-1}^i$. Кроме того, $\mathbb{H}_s = \mathbb{H}_{s-1}$ и $f_s^i = f_{s-1}^i$ для $i \in \mathbb{H}_s$.

Рассмотрим грань конуса \mathbb{D}_{s-1} , представленную в виде

$$\text{cone} \{d_{s-1}^{\mathbb{K}}\} + \mathcal{L} \{f_{s-1}^{\mathbb{H}_{s-1}}\}, \quad \mathbb{K} \subseteq \mathbb{J}_{s-1}.$$

Если $\mathbb{K} \cap \mathbb{B} = \emptyset$, то множество $\text{cone} \{d_s^{\mathbb{K}}\} + \mathcal{L} \{f_s^{\mathbb{H}_s}\}$ является гранью \mathbb{D}_s . В случае $\mathbb{K} \cap \mathbb{B} \neq \emptyset$ и $\mathbb{K} \cap \mathbb{A} \neq \emptyset$ строим две грани \mathbb{D}_s :

$$\text{cone} \left\{ d_s^{(\mathbb{K} \cap (\mathbb{A} \cup \mathbb{S})) \cup (\mathbb{P} \cap ((\mathbb{K} \cap \mathbb{A}) \times (\mathbb{K} \cap \mathbb{B})))} \right\} + \mathcal{L} \{f_s^{\mathbb{H}_s}\},$$

$$\text{cone} \left\{ d_s^{(\mathbb{K} \cap \mathbb{S}) \cup (\mathbb{P} \cap ((\mathbb{K} \cap \mathbb{A}) \times (\mathbb{K} \cap \mathbb{B})))} \right\} + \mathcal{L} \{f_s^{\mathbb{H}_s}\}.$$

В остальных случаях граней \mathbb{D}_s не получаем.

Каноническое представление конуса \mathbb{D}_s в случае $\mathbb{B} = \emptyset$ совпадает с представлением \mathbb{D}_{s-1} ; в случае $\mathbb{B} \neq \emptyset, \mathbb{A} \neq \emptyset$ имеет вид $\mathbb{D}_s = \text{cone} \{d_s^{\mathbb{P} \cup \mathbb{A} \cup \mathbb{S}}\} + \mathcal{L} \{f_s^{\mathbb{H}_s}\}$; если $\mathbb{B} \neq \emptyset, \mathbb{A} = \emptyset, \mathbb{S} \neq \emptyset$, то соответственно $\mathbb{D}_s = \text{cone} \{d_s^{\mathbb{S}}\} + \mathcal{L} \{f_s^{\mathbb{H}_s}\}$; а если $\mathbb{A} = \mathbb{S} = \emptyset$, конус $\mathbb{D}_s = \mathcal{L} \{f_s^{\mathbb{H}_s}\}$.

Таким образом, на каждом шаге s алгоритма по конусу \mathbb{D}_{s-1} строится каноническое представление \mathbb{D}_s . Алгоритм завершается на шаге p построением искомого конуса $\mathbb{D}_p = \mathbb{D}$.

4. Применение к задаче принятия решения. Рассмотрим задачу, в которой лицу, принимающему решение (ЛПР), требуется из множества возможных вариантов \mathbb{X} выбрать некоторое подмножество выбираемых вариантов $\mathbb{C}(\mathbb{X})$. Каждый вариант оценивается набором числовых критериев, т.е. задано отображение $f : \mathbb{X} \rightarrow \mathbb{R}^m$. Пусть, кроме того, предпочтения ЛПР выражаются с помощью бинарного отношения $\succ_{\mathbb{X}} \subseteq \mathbb{X} \times \mathbb{X} : x' \succ_{\mathbb{X}} x'' \Leftrightarrow \mathbb{C}(\{x', x''\}) = \{x'\}$.

Обозначим множество возможных векторов $\mathbb{Y} = f(\mathbb{X})$. На нем отношением $\succ_{\mathbb{X}}$ индуцируется $\succ_{\mathbb{Y}}$. Будем считать, что выполнены аксиомы «разумного выбора» [1], в соответствии с которыми существует иррефлексивное транзитивное конусное продолжение \succ отношения $\succ_{\mathbb{Y}}$ на все пространство \mathbb{R}^m . Напомним, что конусность отношения означает существование выпуклого конуса $\mathbb{M} \subseteq \mathbb{R}^m$ такого, что $x \succ y \Leftrightarrow x - y \in \mathbb{M}$. Из транзитивности и иррефлексивности отношения следует, что конус \mathbb{M} острый и не содержит нуля.

В соответствии с принципом Эджворта – Парето всякое множество выбираемых вариантов является подмножеством множества Парето $\mathbb{P}_f(\mathbb{X}) = \{x \in \mathbb{X} \mid \nexists x^* \in \mathbb{X} : f(x^*) \geq f(x)\}$, где \geq – отношение Парето: $y' \geq y''$ означает, что по всем компонентам $y'_i \geq y''_i$, и хотя бы по одной компоненте имеет место строгое неравенство.

На практике нередко оказывается, что множество Парето, оценивающее сверху множество выбираемых вариантов, все еще довольно широкое, и произвести из него выбор затруднительно. Поэтому возникает задача сужения множества Парето.

Было предложено много подходов к ее решению, обзор которых можно найти в [2]. В данной работе для построения более точной оценки множества $\mathbb{C}(\mathbb{X})$ будет использоваться информация о готовности ЛПР идти на компромисс. Если ЛПР готово ради увеличения оценок по одной группе критериев терпеть убытки по другой группе, то будем говорить, что первая группа критериев важнее второй с параметрами, отражающими конкретные числовые значения выигрышей и убытков. Математически эту информацию можно задать соотношением $u \succ 0$, причем вектор u , называемый «сообщением», должен иметь хотя бы одну строго положительную и хотя бы одну строго отрицательную компоненты.

Задача учета одного сообщения полностью решена в [1]. Испол-

зую идею ее доказательства, предложим способ учета произвольного конечного набора сообщений u^1, \dots, u^p . Отметим, что набор сообщений может оказаться противоречивым в том смысле, что одновременное выполнение соотношений $u^k \succ 0$ невозможно.

Справедливость соотношений $u^k \succ 0$ означает $u^k \in \mathbb{M}$. Если ввести конус $\mathbb{Q} = \text{cone} \{u^1, \dots, u^p, e^1, \dots, e^m\} \setminus \{0\}$, то

$$\mathbb{R}_+^m \subseteq \mathbb{Q} \subseteq \mathbb{M}, \quad (2)$$

где \mathbb{R}_+^m – неотрицательный ортант, конус отношения Парето \geq . Конус \mathbb{Q} можно задать в виде $\mathbb{Q} = \{y \in \mathbb{R}^m \mid g(y) \geq 0\}$, где вектор-функция g имеет компоненты $g_i(y) = d_i y$, а d_i – образующие двойственного к \mathbb{Q} конуса. Введем новый векторный критерий $h(x) = g(f(x))$. С помощью (2) можно показать включение $\mathbb{C}(\mathbb{X}) \subseteq \mathbb{P}_h(\mathbb{X}) \subseteq \mathbb{P}_f(\mathbb{X})$. Таким образом, образующие двойственного к \mathbb{Q} конуса позволяют построить новый векторный критерий, множество Парето относительно которого дает более точную оценку множества выбираемых вариантов.

Построение образующих двойственного к \mathbb{Q} конуса можно вести с помощью алгоритма предыдущего раздела. При этом начинать построение можно с самодвойственного конуса \mathbb{R}_+^m , тогда остается лишь пересечь его с полупространствами, задаваемыми нормальными u^1, \dots, u^p , причем на каждом шаге имеет место ортогональный случай. Наконец, можно показать, что критерием противоречивости набора является выполнение на некотором шаге алгоритма соотношения $\mathbb{A} = \emptyset$.

Таким образом, алгоритм, решающий чисто математическую задачу построения двойственного конуса, может быть применен в задаче многокритериального выбора для решения проблемы учета произвольного конечного набора сообщений об относительной важности критериев.

Литература

1. Ногин В.Д. Принятие решений в многокритериальной среде: количественный подход (2-е изд., испр. и доп.). М.: Физматлит, 2005. 176 с.
2. Ногин В.Д. Проблема сужения множества Парето: подходы к решению // Искусственный интеллект и принятие решений, 2008. № 1, С. 98–112.

Богданова О.К.

Санкт-Петербургский государственный университет

Математическое моделирование распространения нового продукта в социальных сетях

Рекомендовано к публикации профессором Буре В.М.

1. Постановка задачи. Рассмотрим социальную сеть, состоящую из n агентов, образующих конечное множество $N = \{1, \dots, n\}$. Исследуем распространение некоторого нового продукта на рынке, который можно представить такой сетью.

В рассматриваемой сети возможны следующие состояния: $(0, n)$, $(1, n-1), \dots, (n, 0)$. Пара $(i, n-i)$, $i = \overline{0, n}$ – это состояние сети, при котором i – это число активных агентов, а $(n-i)$ – число восприимчивых агентов [1]. Активные игроки уже обладают продуктом, а восприимчивые – нет. Сеть может переходить из одного состояния в другое ниже описанным способом.

Предположим, что существует некоторый фактор воздействия на агентов сети. Например, это может быть реклама продукта. Обозначим уровень рекламы через $\lambda = \lambda(c)$, где $0 < c < c_0$ – это количество денег, вложенное в рекламу, $\lambda \geq 0$ – неубывающая функция.

Обозначим вероятность перехода из восприимчивого состояния в активное через $p = p(\lambda, A)$, где A – это число активных агентов, $p \in [0, 1]$ – неубывающая по A функция [2]. Предположим, что обратный переход (из активного состояния в восприимчивое) невозможен.

Процесс распространения продукта можно описать с помощью цепи Маркова с конечным пространством состояний $\{(0, n), (1, n-1), \dots, (n, 0)\}$ и следующей матрицей вероятностей перехода [3, 4]

$$P = \begin{pmatrix} C_n^0(1-p)^n & C_n^1p(1-p)^{n-1} & \dots & C_n^n p^n \\ 0 & C_{n-1}^0(1-p)^n & \dots & C_{n-1}^{n-1}p^{n-1} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{pmatrix}.$$

Пусть $\lambda(c)$ и $p(\lambda, A)$ известны. Предположим, что на рынке существует одна фирма и ее цель – максимизировать прибыль в долгосрочном периоде. Прибыль фирмы можно описать функцией

$b(\lambda, c, A, r) = \sum_i b_i \delta^{i-1}$, где $b_i = E[A(i) - A(i-1)]r(i) - c(i)$ – прибыль фирмы на i -ом этапе, $\delta \in (0, 1)$ – функция дисконтирования (обесценивания), $r \in [r_0, R]$ – цена продукта, r_0 – его себестоимость. Будем считать, что продолжительность игры равна l .

Стратегией фирмы является вектор $\{(r_j, c_j)\}$, $j = 1, \dots, l$, где (r_j, c_j) – это стратегия фирмы в состоянии j .

Задача заключается в нахождении оптимальной стратегии фирмы, т.е. стратегии, максимизирующей ее общую прибыль.

2. Пример. Рассмотрим некоторую сеть, состоящую из трех агентов, каждый из которых изначально является восприимчивым. Таким образом, на нулевом этапе сеть находится в состоянии $(0, 3)$. Предположим, что λ и p известны и постоянны. Рассмотрим фирму, распространяющую новый продукт. Допустим, цель этой фирмы – максимизировать свою прибыль за два этапа.

Рассматриваемому случаю соответствует следующая матрица вероятностей перехода

$$\Pi = \begin{pmatrix} (1-p)^3 & 3p(1-p)^2 & 3p^2(1-p) & p^3 \\ 0 & (1-p)^2 & 2p(1-p) & p^2 \\ 0 & 0 & (1-p) & p \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

В начальном состоянии число активных агентов равно 0, т.е. $A(0) = (1, 0, 0, 0)$. Значит математическое ожидание числа активных игроков на этом этапе $E[A(0)] = 0$.

Из начального состояния с нулем активных игроков можно перейти в одно из следующих четырех состояний: $(0, 3)$, $(1, 2)$, $(2, 1)$, $(3, 0)$ с соответствующими вероятностями. Вектор распределения активных агентов на данном этапе вычисляется по формуле $A(1) = A(0)\Pi$, следовательно,

$$A(1) = ((1-p)^3, 3p(1-p)^2, 3p^2(1-p), p^3).$$

Таким образом, на первом этапе математическое ожидание числа активных игроков может быть вычислено следующим образом:

$$E[A(1)] = 3p(1-p)^2 + 6p^2(1-p) + 3p^3 = 3p.$$

Теперь можно найти прибыль фирмы на данном этапе:

$$b_1 \delta^0 = E[A(1) - A(0)]r(1) - c(1) = 3pr(1) - c(1).$$

Вектор распределения активных игроков на втором этапе вычисляется аналогичным образом:

$$\begin{aligned}
A(2) &= \Pi A(1) = A(0) \Pi^2 = \\
&= A(0) \begin{pmatrix} (1-p)^3 & 3p(1-p)^2 & 3p^2(1-p) & p^3 \\ 0 & (1-p)^2 & 2p(1-p) & p^2 \\ 0 & 0 & (1-p) & p \\ 0 & 0 & 0 & 1 \end{pmatrix}^2 = A(0) \times \\
&\times \begin{pmatrix} (1-p)^6 & 3p(1-p)^4(2-p) & 3p^2(1-p)^2(2-p)^2 & p^3(2-p)^3 \\ 0 & (1-p)^4 & 2p(1-p)^2(2-p) & p^2(2-p)^2 \\ 0 & 0 & (1-p)^2 & p(2-p) \\ 0 & 0 & 0 & 1 \end{pmatrix} = \\
&= ((1-p)^6, 3p(1-p)^4(2-p), 3p^2(1-p)^2(2-p)^2, p^3(2-p)^3).
\end{aligned}$$

Найдем математическое ожидание числа активных игроков на втором этапе

$$E[A(2)] = 3p(1-p)^4(2-p) + 6p^2(1-p)^2(2-p)^2 + 3p^3(2-p)^3 = 6p - 3p^2.$$

Прибыль фирмы на данном этапе можно вычислить следующим образом:

$$b_2\delta = (E[A(2)] - A(1))r(2) - c(2))\delta = ((3p - 3p^2)r(2) - c(2))\delta.$$

Общая прибыль фирмы за два этапа будет являться суммой прибылей на каждом из них:

$$\begin{aligned}
b(\lambda, c, A, r) &= b_1 + b_2\delta = \\
&= (3pr(1) - c(1)) + ((3p - 3p^2)r(2) - c(2))\delta,
\end{aligned}$$

где $r(i) \in [r_0, R]$ и $0 < c(i) < c_0$, $i = 1, 2$.

Получили линейную функцию многих переменных. Следовательно, максимум данной функции будет достигаться на концах отрезков, ограничивающих параметры $r(i)$ и $c(i)$, $i = 1, 2$. Нетрудно убедиться, что в данном случае максимум функции $b(\lambda, c, A, r)$ достигается при $r(i) = R$ и $c(i) = 0$, $i = 1, 2$, т.е.

$$\max b(\lambda, c, A, r) = 3pR(1 + (1 - p)\delta).$$

Литература

1. Богданова О.К., Парилина Е.М. Эмпирическое исследование социальных сетей // Процессы управления и устойчивость: Труды 41-й международной научной конференции аспирантов и студентов / Под ред. Н.В. Смирнова, Г.Ш. Тамасяна. СПб.: Издат. Дом С.-Петерб. гос. ун-та, 2010. С. 559–564.
2. Губанов Д.А., Новиков Д.А., Чхартишвили А.Г. Модели влияния в социальных сетях // Управление большими системами, 2009. Вып. 27. С. 205–281.
3. Майн Х., Осаки С. Марковские процессы принятия решений. М.: Наука, 1977. 176 с.
4. Petrosyan L.A., Baranova E.M. Cooperative stochastic games in stationary strategies // Game Theory and Applications, 2006. № 11. P. 1–7.

Бровкина Л.А.

Санкт-Петербургский государственный университет

Модель прогнозирования объемов перевозок грузов

*Рекомендовано к публикации старшим преподавателем
Лежниной Е.А.*

Введение. Данная работа посвящена построению математической модели прогнозирования объемов перевозок грузов. Исходным материалом для задачи является некоторый набор данных о процессе, полученный через равные промежутки времени – временной ряд. В общем случае временной ряд рассматривается как смесь четырех составляющих [1]:

- 1) тренда или систематического движения,
- 2) колебания относительно тренда,
- 3) сезонной компоненты,
- 4) остатка или несистематического случайного компонента.

Постановка задачи. На основе реальных данных транспортно-экспедиторской фирмы о перевозках груза в течение дня для 53 предприятий в период с 06 января 2009 г. по 28 июля 2010 г. требуется построить математическую модель соответствующего временного ряда и прогноз с целью получения как можно более достоверной информации о количестве груза, доставляемого на каждое предприятие.

Основные результаты. Для предварительного анализа была разработана программа в пакете MATLAB.

Известно [1], что для выявления структуры процесса нужно сделать оценку средних значений и среднеквадратических разбросов данных, анализ корреляционной и спектральной структуры процесса. О нестационарности процесса можно судить по графику корреляционной функции, на котором после периода затухания появляется несколько сравнительно больших значений коэффициентов корреляции. По наличию пиков на графике периодограммы можно сделать вывод о присутствии периодической составляющей [2]. Для анализа трендов строилась линейная аппроксимация процесса по методу наименьших квадратов. Амплитуда периодических колебаний временного ряда примерно постоянна.

Исходя из этих результатов, можно предложить аддитивную модель процесса перевозок:

$$x_t = u_t + v_t + \varepsilon_t,$$

где x_t – уровни временного ряда; u_t – трендовая составляющая; v_t – периодическая составляющая; ε_t – остаточная компонента.

Для выделения периодической составляющей процесс подвергался дискретному преобразованию Фурье (ДПФ), в котором выделялись частоты на уровне трех средних значений средней абсолютной величины (см. рис. 1). Выделенная часть подвергалась обратному преобразованию Фурье.

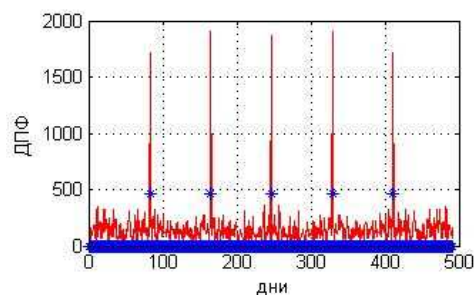


Рис. 1. График дискретного преобразования Фурье процесса $x_t - u_t$

При анализе остатков были получены следующие результаты: средняя величина остатков близка к нулю и равна 10^{-14} , а корреляционная функция вырождена. Не прослеживается направленность в расположении остатков, что говорит о гомоскедастичности. По графику гистограммы и корреляционной функции можно сделать вывод, что распределение остаточного процесса близко к нормальному распределению. Это означает, что для остаточного процесса можно принять модель белого шума, т.е. последовательности независимых случайных величин с нулевым средним. Следовательно, построенную модель можно использовать для прогноза.

Для оценки потенциальной точности прогнозирования оценивались значения среднеквадратических разбросов: s_x – исходного процесса, s_u – процесса $x_t - u_t$, s_ε – остаточного процесса. Результаты для первого предприятия представлены в таблице 1.

Таблица 1. Среднеквадратический разброс

s_x	s_u	s_ε
11,0310	10,8689	6,9956

Величина s_x характеризует ошибку прогноза процесса с помощью его среднего значения (постоянного процесса), s_u – с помощью трендового (линейного) процесса, s_ε – с помощью комбинации трендового процесса и периодической составляющей.

Из таблицы 1 видно, что линейный прогноз слабо уменьшает ошибку прогноза, прогноз с учетом периодической составляющей уменьшает ошибку прогноза примерно на 35%. Для других предприятий результаты имеют сходный характер.

Результаты статистического анализа позволяют предложить алгоритм прогнозирования следующей структуры:

- выделение трендовой и периодической составляющей по участку процесса, предшествующему прогнозируемой дате (участок анализа);
- прогнозирование на основе экстраполяции выделенных составляющих на прогнозируемую дату.

Важным вопросом при этом является длина участка анализа. Для ответа на этот вопрос была разработана программа моделирования для различных длительностей участков анализа N (в неделях). Результаты моделирования приведены на рис. 2.

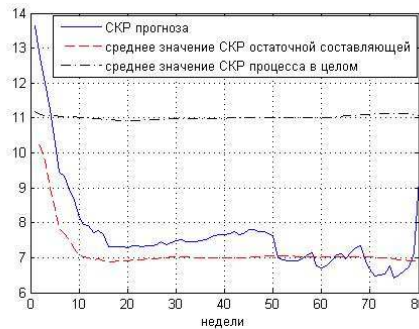


Рис. 2. Зависимость значений среднеквадратического разброса (СКР) от длительности участка

Из представленных результатов видно, что среднее значение СКР остаточной составляющей близко к среднеквадратической ошибке

прогноза, начиная с величин N порядка 20. В связи с этим, длина участка выбрана $N = 20$ недель.

Далее была разработана программа моделирования алгоритма одношагового прогноза. Участок графика результатов прогноза представлен на рис. 3.

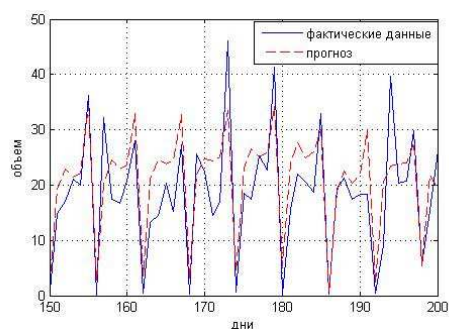


Рис. 3. График прогноза

В таблице 2 для первого предприятия представлены среднее значение СКР процесса δ_x по участку длительности $N = 20$, среднее значение СКР остаточной составляющей δ_ε и среднеквадратическая ошибка прогноза δ_p .

Таблица 2. Значения СКР по участку длительности $N = 20$

δ_x	δ_ε	δ_p
10,9179	6,8763	7,325

Из таблицы 2 видно, что среднеквадратическая ошибка прогноза близка к СКР остаточной составляющей, и прогнозирование на основе трендовой и периодической составляющей позволяет уменьшить ошибку прогноза примерно на 30%, что соответствует результатам проведенного статистического анализа. Для остальных предприятий результаты имеют сходный характер.

Литература

1. Кендалл М., Стьюарт А. Многомерный статистический анализ и временные ряды. М.: Наука, 1976. 736 с.
2. Бушманова М.В., Иванова Т.А., Мельникова Г.Г., Реент Н.А., Трофимова В.Ш. Анализ временных рядов и прогнозирование: Учебное пособие. Магнитогорск: Изд-во МГТУ, 2006. 142 с.

Буре В.М., Сергеева А.А.
Санкт-Петербургский государственный университет

Оптимальные стратегии выбора логистической фирмы

Введение. Логистика в настоящее время широко используется во всем мире, и без решения логистических задач трудно победить в конкурентной борьбе. По мере развития товарно-денежных отношений рынки увеличиваются, потребности в товаре растут, а следовательно, возникает необходимость в своевременной доставке нужного товара от производителя до конечного потребителя. С вовлечением в процесс товарооборота все новых участников структура рынка значительно усложняется, и задача оперативной обработки заказа клиента становится первоочередной. Вопросами своевременного размещения заказов, складирования, хранения, транспортировки, сопровождения занимается логистика. Существует большое количество публикаций, в которых применяются различные математические методы для изучения проблем логистики (см., например, [1–3]). В данной работе применяется математическая теория игр, теория вероятностей и элементы теории массового обслуживания для исследования логистической модели.

Будем рассматривать логистическую систему следующего вида. Пусть на рынке логистических перевозок имеются две фирмы, которые осуществляют услуги по транспортировке товара от производителя до конечного потребителя. Каждая фирма предлагает свою политику формирования конечной стоимости выполнения заказа клиента. Фирма 1 выполняет заказы клиентов в порядке очереди, фирма 2 выполняет все заказы сразу. При этом стоимость выполнения заказа фирмой 1 является фиксированной, а фирма 2 включает в стоимость цену за единицу времени выполнения заказа. Также необходимо учесть потери клиентов, связанные с ожиданием выполнения заказа той или иной фирмой. Клиент выбирает фирму для транспортировки товара, стараясь минимизировать свои издержки.

Постановка задачи. Каждая фирма представляет собой обслуживающее устройство. Таким образом, в системе имеется два обслуживающих устройства, каждое из которых определяет свой порядок обслуживания [4]. Фирма 1 обслуживает клиентов в порядке очереди, фирма 2 обслуживает всех клиентов сразу. Длительности обслу-

живания клиентов фирмами 1 и 2 являются независимыми случайными величинами с экспоненциальными распределениями, длительность обслуживания клиентов фирмой 1 определяется плотностью $f_1(t) = \frac{1}{\mu_1} e^{-\mu_1 t}$, $t > 0$, а длительность обслуживания клиентов фирмой 2 – плотностью $f_2(t) = \frac{1}{\mu_2} e^{-\mu_2 t}$, $t > 0$.

Обозначим через τ_1 – время пребывания в системе клиента при выборе фирмы 1. Это время складывается из двух слагаемых: $\tau_1 = \tau_{11} + \tau_{12}$, где τ_{11} – время ожидания начала выполнения заказа фирмой 1, τ_{12} – время обслуживания фирмой 1. Через τ_2 обозначим время пребывания в системе клиента при выборе фирмы 2. Так как время ожидания начала обслуживания в фирме 2 равно нулю, то $\tau_2 = \tau_{22}$, где τ_{22} – время обслуживания фирмой 2. Параметры τ_1 и τ_2 являются случайными величинами.

Определим затраты клиентов на обслуживание в каждой из фирм.

Пусть c_1 – стоимость выполнения заказа фирмой 1. Она является фиксированной и не зависит от длительности выполнения заказа клиента. Пусть далее $c_2 = c_{21} + c_{22}\tau_{22}$ – стоимость выполнения заказа фирмой 2, зависящая от длительности обслуживания клиента фирмой 2. Здесь c_{21} – фиксированная стоимость, взимаемая за выполнение заказа, c_{22} – стоимость единицы времени обслуживания клиента фирмой 2.

Помимо издержек на выполнение заказа клиенты несут потери, связанные с ожиданием выполнения заказа. Пусть r_1 – удельные потери, которые несет клиент при ожидании выполнения заказа у фирмы 1, r_2 – удельные потери, которые несет клиент при ожидании выполнения заказа у фирмы 2. Можно определить полные потери, связанные с ожиданием выполнения заказа фирмой 1, которые будут складываться из двух слагаемых: $r_1\tau_1 = r_1(\tau_{11} + \tau_{12})$, а полные потери, связанные с ожиданием выполнения заказа фирмой 2, будут выглядеть следующим образом: $r_2\tau_2 = r_2\tau_{22}$.

Теперь можно определить полные потери клиентов на обслуживание фирмами 1 и 2 следующим образом:

$$\begin{aligned}\widetilde{Q}_1 &= r_1\tau_1 + c_1, \\ \widetilde{Q}_2 &= r_2\tau_2 + c_2 = (r_2 + c_{22})\tau_{22} + c_{21}.\end{aligned}$$

Тогда средние потери клиентов на обслуживание при условии обращения клиента в соответствующую фирму определяются математическими ожиданиями

$$Q_1 = E\widetilde{Q}_1 = r_1(E\tau_{11} + E\tau_{12}) + c_1,$$

$$Q_2 = E\widetilde{Q}_2 = (r_2 + c_{22})E\tau_{22} + c_{21}.$$

Предположим, что в данный момент времени на обслуживание поступает группа, состоящая из n клиентов, при этом известно, что на обслуживании в фирме 1 находится k клиентов (из них $k - 1$ стоит в очереди на обслуживание). Каждый клиент решает, какую фирму выбрать для осуществления перевозки товара. Пусть p_i – вероятность того, что клиент i выберет фирму 1, соответственно $1 - p_i$ – вероятность того, что клиент i выберет фирму 2.

Данная модель приводит к игре n лиц [5], в которой клиенты представляют собой игроков, выбирающих логистическую фирму для транспортировки необходимого товара. Рассмотрим неантагонистическую игру в нормальной форме:

$$\Gamma = \langle N, \{p_i\}_{i \in N}, \{H_i\}_{i \in N} \rangle,$$

где $N = \{1, \dots, n\}$ – множество игроков, $\{p_i\}_{i \in N}$ – множество стратегий игроков, $p_i \in [0, 1]$, $\{H_i\}_{i \in N}$ – множество функций выигрыша игроков,

$$H_i = -(p_i Q_{1i} + (1 - p_i) Q_{2i}) = -(p_i(Q_{1i} - Q_{2i}) + Q_{2i}).$$

Ожидаемые потери игрока i на обслуживание при условии обращения в фирму 1: $Q_{1i} = r_1 t_{1i} + c_1$, где t_{1i} – среднее время до полного выполнения заказа при условии, что игрок i обратился в фирму 1. Ожидаемые потери игрока i на обслуживание при условии обращения в фирму 2: $Q_{2i} = (r_2 + c_{22})t_{2i} + c_{21}$, где t_{2i} – среднее время до полного выполнения заказа при условии, что игрок i обратился в фирму 2, $i = 1, \dots, n$.

Далее будем использовать функцию потерь игроков: $h_i = -H_i$, $i = 1, \dots, n$.

Будем предполагать, что рассматривается игра с полной информацией.

Основные результаты. Сформулирована и доказана следующая теорема.

Теорема. В рассмотренной игре $\Gamma = \langle N, \{p_i\}_{i \in N}, \{H_i\}_{i \in N} \rangle$ существует единственная ситуация равновесия (p_1^*, \dots, p_n^*) , которая определяется следующим образом:

- если $r_1\mu_1(\frac{n}{2} + k + \frac{1}{2}) - (r_2 + c_{22})\mu_2 + c_1 - c_{21} < 0$, то $p_i^* = 1$, $i = 1, \dots, n$;
- если $r_1\mu_1(k+1) - (r_2 + c_{22})\mu_2 - c_{21} + c_1 > 0$, то $p_i^* = 0$, $i = 1, \dots, n$;
- если $r_1\mu_1(k+1) \leq (r_2 + c_{22})\mu_2 + c_{21} - c_1 \leq r_1\mu_1(\frac{n}{2} + k + \frac{1}{2})$,
то $p_i^* = \frac{2((r_2 + c_{22})\mu_2 - r_1(k+1)\mu_1 - c_1 + c_{21})}{r_1\mu_1(n-1)}$, $i = 1, \dots, n$.

где $k = 0$, если у фирмы 1 нет клиентов в очереди и на обслуживании, $k = 1$, если у фирмы 1 на обслуживании находится один клиент и нет клиентов в очереди, $k > 1$, если у фирмы 1 на обслуживании находится один клиент и $k - 1$ клиент в очереди.

Заключение. В данной работе рассматривается логистический рынок при наличии двух фирм, осуществляющих перевозку товара от производителя клиенту. Клиенты являются игроками, выбирающими фирму для обслуживания с целью минимизировать свои затраты. Найдены оптимальные стратегии поведения игроков.

Литература

1. Ghiani G., Laporte G., Musmanno R. Introduction to logistics systems planning and control. London: John Wiley and Sons, 2004. 360 p.
2. Daganzo C. Logistics system analysis. Berlin: Shpringer, 1996. 272 p.
3. Langevin A., Riopel D. Logistics systems: design and optimization. New York: Springer, 2005. 392 p.
4. Буре В.М. Теоретико-игровая модель одной системы массового обслуживания // Вестник СПбГУ. Сер. 1, 2002. Вып. 2. С. 3–5.
5. Петросян Л.А., Зенкевич Н.А., Семина Е.А. Теория игр: Учебное пособие для университетов. М.: Высшая школа, 1998. 304 с.

Горбати М.В.

Санкт-Петербургский государственный университет

Динамическая модель лицензирования на рынке одежды

Рекомендовано к публикации доцентом Шевкопляс Е.В.

1. Постановка задачи. Рассмотрим теоретико-игровую модель лицензирования на рынке одежды [1]. В дифференциальной игре участвуют два игрока, а именно, владелец известного бренда первой линии (игрок L) и получатель лицензии (игрок l), т.е. фирма, которой владелец бренда L предоставляет права на производство и продажу бренда второй линии с денежной компенсацией. Предполагается, что лицензионный контракт предоставляется владельцем бренда на фиксированное время $[0, T]$, $T > 0$.

Пусть $S(t)$ – это общий объем продаж предоставленного по лицензии бренда в момент $t \in [0, T]$, тогда $\dot{S}(t)$ – скорость продаж в момент $t \in [0, T]$. Предполагаем, что стоимость рекламы для игрока $i \in \{L, l\}$ в момент $t \in [0, T]$ имеет вид $K_i(a_i(t)) = \frac{k_i}{2} a_i^2(t)$, $k_i > 0$, где k_i – ценовой параметр, $a_i(t)$ – рекламные расходы игрока $i \in \{L, l\}$ в момент $t \in [0, T]$. Пусть $r \in (0, 1)$ – доля выручки игрока L от продаж бренда второй линии. Тогда общий доход игрока L за предоставленную лицензию на торговый сезон $[0, T]$ вычисляется по формуле

$$R_L = \int_0^T r \dot{S}(t) dt = rS(T).$$

Доход игрока l за тот же период имеет следующий вид

$$R_l = \int_0^T (1 - r) \dot{S}(t) dt = (1 - r)S(T).$$

Согласно модели [1] полагаем, что на репутацию бренда второй линии не влияет репутация бренда первой линии. Следовательно, будем учитывать только стоимость деловой репутации бренда второй линии. Кроме того, предполагается, что динамика деловой репутации

зависит от рекламной политики, проводимой игроками L и l . Пусть $G(t)$ – стоимость деловой репутации, причем её изменение описывается дифференциальным уравнением

$$\dot{G}(t) = \gamma_L a_L(t) + \gamma_l a_l(t) - \beta G(t), \quad G(0) = \bar{G}, \quad (1)$$

где $\beta > 0$ – скорость ухудшения деловой репутации, $\bar{G} > 0$ – начальное значение стоимости деловой репутации, $\gamma_i > 0$, $i \in \{L, l\}$ – эффективность рекламной политики игрока i . Изменение объема продаж задается уравнением

$$\dot{S}(t) = \alpha G(t), \quad S(0) = 0, \quad (2)$$

где $\alpha > 0$ – коэффициент пересчета стоимости деловой репутации в общий объем продаж. В отличие от базовой модели [1] будем считать, что прибыль игрока дисконтируется во времени при помощи экспоненциальной функции $e^{-\rho(t)t}$. Целью игроков L и l является максимизация собственной прибыли π_i , $i \in \{L, l\}$:

$$\begin{aligned} \max_{a_L(t) \geq 0} \pi_L &= \rho_L G(T) e^{-\rho T} + r S(T) e^{-\rho T} - \int_0^T \frac{k_L a_L^2(t)}{2} e^{-\rho t} dt, \\ \max_{a_l(t) \geq 0} \pi_l &= \rho_l G(T) e^{-\rho T} + (1-r) S(T) e^{-\rho T} - \int_0^T \frac{k_l a_l^2(t)}{2} e^{-\rho t} dt, \end{aligned}$$

где $\rho_i > 0$, $i \in \{L, l\}$, $r \in (0, 1)$.

2. Равновесие по Штакельбергу для постоянного дисконтирования. Пусть $\rho(t) = \rho$, тогда справедливо следующее

Утверждение 1. *Равновесные по Штакельбергу [2] рекламные расходы при постоянном дисконтировании имеют вид*

$$\begin{aligned} a_l^c &= \frac{(1-r)\alpha\gamma_l}{\beta k_l} e^{\rho(t-T)} + \frac{\gamma_l}{k_l} \left(\rho_l - \frac{(1-r)\alpha}{\beta} \right) e^{(\rho+\beta)(t-T)}, \\ a_L^c &= \frac{r\alpha\gamma_L}{\beta k_L} e^{\rho(t-T)} + \frac{\gamma_L}{k_L} \left(\rho_L - \frac{\alpha r}{\beta} \right) e^{(\rho+\beta)(t-T)}. \end{aligned}$$

Доказательство. В качестве управления игрока $i \in \{L, l\}$ выступает $a_i(t)$. При помощи принципа максимума сначала найдем оптимальное управление для игрока l . Гамильтониан для игрока l имеет вид

$$H^l = -\frac{k_l a_l^2}{2} e^{-\rho t} + \psi(\gamma_L a_L + \gamma_l a_l - \beta G) + \zeta \alpha G,$$

где ψ, ζ – сопряженные переменные к системе (1), (2), которые должны удовлетворять системе уравнений

$$\dot{\psi} = -H_G^l = \psi\beta - \zeta\alpha, \quad \psi(T) = \rho l e^{-\rho T}, \quad (3)$$

$$\dot{\zeta} = -H_S^l = 0, \quad \zeta(T) = (1-r)e^{-\rho T}, \quad (4)$$

$$H_{a_l}^l = -k_l a_l e^{-\rho t} + \psi\gamma_l = 0. \quad (5)$$

Решая (4), получим

$$\zeta = \text{const} = \zeta(T) = (1-r)e^{-\rho T}.$$

Подставим ζ в (3)

$$\dot{\psi} = \psi\beta - \alpha(1-r)e^{-\rho T}.$$

Решение (3) есть

$$\psi = \frac{(1-r)\alpha}{\beta} e^{-\rho T} + \left(\rho l - \frac{(1-r)\alpha}{\beta} \right) e^{-\rho T + \beta(t-T)}.$$

Из (5) находим

$$a_l = \frac{\psi\gamma_l}{k_l} e^{\rho t}$$

и подставляем найденное ψ :

$$a_l^c = \frac{(1-r)\alpha\gamma_l}{\beta k_l} e^{\rho(t-T)} + \frac{\gamma_l}{k_l} \left(\rho l - \frac{(1-r)\alpha}{\beta} \right) e^{(\rho+\beta)(t-T)}.$$

Гамильтониан для игрока L имеет вид

$$H^L = -\frac{k_L a_L^2}{2} e^{-\rho t} + \nu(\gamma_L a_L + \gamma_l a_l^c - \beta G) + \lambda \alpha G,$$

где ν, λ – сопряженные переменные к системе (1), (2). Применяя принцип максимума Понтрягина аналогично тому, как было сделано выше для игрока l , получим систему

$$\begin{aligned}\dot{\nu} &= -H_G^L = \nu\beta - \lambda\alpha, & \nu(T) &= \rho_L e^{-\rho T}, \\ \dot{\lambda} &= -H_S^L = 0, & \lambda(T) &= r e^{-\rho T}, \\ H_{a_L}^L &= -k_L a_L e^{-\rho t} + \nu\gamma_L = 0.\end{aligned}$$

Из системы находим оптимальное управление игрока L :

$$a_L^c = \frac{r\alpha\gamma_L}{\beta k_L} e^{\rho(t-T)} + \frac{\gamma_L}{k_L} \left(\rho_L - \frac{\alpha r}{\beta} \right) e^{(\rho+\beta)(t-T)}.$$

3. Равновесие по Штакельбергу для непостоянной функции дисконтирования. Рассмотрим задачу с непостоянным дисконтированием $\rho(t)$. В этом случае прибыли игроков L и l имеют вид

$$\begin{aligned}\max_{a_L(t) \geq 0} \pi_L &= \rho_L G(T) e^{-\rho(T)T} + r S(T) e^{-\rho(T)T} - \int_0^T \frac{k_L a_L^2(t)}{2} e^{-\rho(t)t} dt, \\ \max_{a_l(t) \geq 0} \pi_l &= \rho_l G(T) e^{-\rho(T)T} + (1-r) S(T) e^{-\rho(T)T} - \int_0^T \frac{k_l a_l^2(t)}{2} e^{-\rho(t)t} dt,\end{aligned}$$

где $\rho_i > 0, i \in \{L, l\}, r \in (0, 1)$.

Утверждение 2. Равновесные по Штакельбергу [2] рекламные расходы при непостоянном дисконтировании $\rho(t)$ имеют вид

$$\begin{aligned}a_l^{nc} &= \frac{(1-r)\alpha\gamma_l}{\beta k_l} e^{\rho(t)t - \rho(T)T} + \frac{\gamma_l}{k_l} \left(\rho_l - \frac{(1-r)\alpha}{\beta} \right) e^{\rho(t)t + \beta(t-T) - \rho(T)T}, \\ a_L^{nc} &= \frac{r\alpha\gamma_L}{\beta k_L} e^{\rho(t)t - \rho(T)T} + \frac{\gamma_L}{k_L} \left(\rho_L - \frac{\alpha r}{\beta} \right) e^{\rho(t)t + \beta(t-T) - \rho(T)T}.\end{aligned}$$

Доказательство. Проведем вычисления аналогично доказательству утверждения 1. В качестве управления игрока $i \in \{L, l\}$ выступает $a_i(t)$. Тогда Гамильтониан для игрока l

$$H^l = -\frac{k_l a_l^2}{2} e^{-\rho(t)t} + \psi(\gamma_L a_L + \gamma_l a_l - \beta G) + \zeta \alpha G,$$

где ψ, ζ – сопряженные переменные к системе (1), (2), удовлетворяющие системе

$$\begin{aligned}\dot{\psi} &= -H_G^l = \psi\beta - \zeta\alpha, & \psi(T) &= \rho_l e^{-\rho(T)T}, \\ \dot{\zeta} &= -H_S^l = 0, & \zeta(T) &= (1-r)e^{-\rho(T)T}, \\ H_{a_l}^l &= -k_l a_l e^{-\rho(t)t} + \psi\gamma_l = 0.\end{aligned}$$

Окончательно получаем

$$a_l^{nc} = \frac{(1-r)\alpha\gamma_l}{\beta k_l} e^{\rho(t)t - \rho(T)T} + \frac{\gamma_l}{k_l} \left(\rho_l - \frac{(1-r)\alpha}{\beta} \right) e^{\rho(t)t + \beta(t-T) - \rho(T)T}.$$

Далее запишем Гамильтониан для игрока L :

$$H^L = -\frac{k_L a_L^2}{2} e^{-\rho(t)t} + \nu(\gamma_L a_L + \gamma_l a_l^{nc} - \beta G) + \lambda \alpha G,$$

где ν, λ – сопряженные переменные к системе (1), (2), удовлетворяющие системе

$$\begin{aligned}\dot{\nu} &= -H_G^L = \nu\beta - \lambda\alpha, & \nu(T) &= \rho_L e^{-\rho(T)T}, \\ \dot{\lambda} &= -H_S^L = 0, & \lambda(T) &= r e^{-\rho(T)T}, \\ H_{a_L}^L &= -k_L a_L e^{-\rho(t)t} + \nu\gamma_L = 0.\end{aligned}$$

Окончательно получаем равновесные по Штакельбергу расходы фирмы L :

$$a_L^{nc} = \frac{r\alpha\gamma_L}{\beta k_L} e^{\rho(t)t - \rho(T)T} + \frac{\gamma_L}{k_L} \left(\rho_L - \frac{\alpha r}{\beta} \right) e^{\rho(t)t + \beta(t-T) - \rho(T)T}.$$

4. Проверка системы на устойчивость по Ляпунову. Запишем систему (1), (2) в матричной форме

$$\begin{pmatrix} \dot{G}(t) \\ \dot{S}(t) \end{pmatrix} = \begin{pmatrix} -\beta & 0 \\ \alpha & 0 \end{pmatrix} \begin{pmatrix} G(t) \\ S(t) \end{pmatrix} + \begin{pmatrix} \gamma_L a_L(t) + \gamma_l a_l(t) \\ 0 \end{pmatrix}.$$

Соответствующая ей однородная система имеет вид

$$\begin{pmatrix} \dot{G}(t) \\ \dot{S}(t) \end{pmatrix} = \begin{pmatrix} -\beta & 0 \\ \alpha & 0 \end{pmatrix} \begin{pmatrix} G(t) \\ S(t) \end{pmatrix}. \quad (6)$$

Пусть $A = \begin{pmatrix} -\beta & 0 \\ \alpha & 0 \end{pmatrix}$. Найдем собственные числа матрицы A :

$$\det(A - \lambda E) = \begin{vmatrix} -\beta - \lambda & 0 \\ \alpha & -\lambda \end{vmatrix} = (-\lambda)(-\beta - \lambda) = 0.$$

Тогда $\lambda_1 = 0$, $\lambda_2 = -\beta < 0$ – собственные числа с неположительной вещественной частью.

Заметим, что система (6) удовлетворяет соотношению

$$n - \text{rank}(A - \lambda_m E) = k_m,$$

где n – размерность матрицы A , k_m – кратность собственных чисел λ_m с нулевой вещественной частью, $m \leq n$. В нашем случае $n = 2$, $k_m = 1$ и

$$\text{rank}(A - \lambda_m E) = \text{rank}(A - \lambda_1 E) = \text{rank} A = 1.$$

Тогда, согласно [3, 4] система (6) устойчива по Ляпунову, и, следовательно, устойчива по Ляпунову и система (1), (2).

Литература

1. Buratto A., Zaccour G. Coordination of advertising strategies in a fashion licensing contract // Journal of Optimization Theory and Applications, 2009. Vol. 142, p. 31–53.
2. Osborne M.J., Rubenstein A. A course in game theory. Cambridge, Massachusetts: MIT Press, 1994. 352 p.
3. Александров А.Ю., Александрова Е.Б., Екимов А.В., Смирнов Н.В. Сборник задач и упражнений по теории устойчивости. СПб.: «СОЛО», 2003. 162 с.
4. Демидович Б.П. Лекции по математической теории устойчивости М.: Наука, 1967. 472 с.

Григорова Т. В.

Санкт-Петербургский государственный университет

Оптимизация коалиционной структуры в иерархических играх

Рекомендовано к публикации профессором Петросяном Л.А.

1. Введение. В работе рассматриваются иерархические коалиционные двухуровневые игры. Ищется равновесие по Нэшу, а также рассматривается задача нахождения оптимального коалиционного разбиения.

2. Основные понятия. Рассмотрим иерархическую игру со следующей структурой.

Пусть A_0 – центр распределения. B_k , $k = 1, \dots, n$ – центр производства, производящий l видов товаров. Пусть $u_k = (u_1^k, \dots, u_m^k)$ – вектор, определяющий количество ресурсов, передаваемых центром A_0 центру производства B_k . Таким образом, центр A_0 может ограничивать возможности, подчиненных ему центров производств. В свою очередь, каждый производственный центр B_k производит из поступивших ресурсов u_k товары $x_k = (x_1^k, \dots, x_l^k)$, $x_k \geq 0$, в этом выражении x_l^k обозначает количество товара вида l , произведенного центром B_k . При этом имеют место следующие ограничения

$$x_k A_k \leq u_k + \alpha_k, \quad k = 1, \dots, n,$$

$$\sum_{k=1}^n u_k = b, \quad u_k \geq 0.$$

Здесь $A_k = (\alpha_{ij}^k)$, где $\alpha_{ij}^k \geq 0$ – количество товаров i -го вида, которое центр B_k произвел, используя единицу j -го вида ресурса, направленного центром A_0 в B_k . Пусть A^k – это технологическая матрица B_k -го производственного центра. Обозначим $\alpha_k \geq 0$ как вектор наличных ресурсов B_k -го производственного центра.

Под стратегией центра распределения A_0 будем понимать вектор $u = (u_1, \dots, u_n)$.

Под стратегией центра производства B_k будем понимать вектор $x_k = (x_1^k, \dots, x_l^k)$, т.е. вектор произведенных из ресурсов u_k товаров.

Определим функции выигрышей игроков.

Для центра распределения A_0 функция выигрыша будет иметь вид

$$\sum_{k=1}^n c_k x_k, \quad c_k \geq 0.$$

Здесь c_k можно интерпретировать как вектор цен, реализованных центром распределения A_0 на все виды производственных товаров $x_k(u_k)$.

Для центров производства B_k функция выигрыша определим следующим образом

$$(d_k, x_k), \quad d_k \geq 0,$$

где d_k можно интерпретировать как вектор цен, реализованных центром производства B_k на все виды производственных товаров x_k .

Пусть $\bar{x}_k(u_k)$ – решение задачи линейного программирования

$$\begin{cases} \max_{x_k} (d_k, x_k), \\ x_k A_k \leq u_k + \alpha_k, \\ x_k \geq 0. \end{cases}$$

Пусть \bar{u}_k – решение задачи нелинейного программирования

$$\begin{cases} \max_{u_k} \sum_{k=1}^n c_k \bar{x}_k(u_k), \\ \sum_{k=1}^n u_k = b, \\ u_k \geq 0. \end{cases}$$

Можно показать, что $(\bar{u}, \bar{x}_1(\bar{u}_1), \dots, \bar{x}_n(\bar{u}_n))$ является равновесием по Нэшу [1].

3. Коалиционная модель поведения. Рассмотрим коалиционную модель поведения центров производства B_k , $k = 1, \dots, n$. Предположим, что центры производства B_1, \dots, B_n получили возможность объединяться в коалиции [2]. Рассмотрим разбиение M_1, \dots, M_s

$$\bigcup_{r=1}^s M_r = \bigcup_{k=1}^n B_k,$$

$$M_i \cap M_j = \emptyset, \quad i \neq j.$$

Допустим, что центр производства B_k входит в коалицию M_r . Далее будем рассматривать ситуацию, в которой распределение ресурсов производится между коалициями M_1, \dots, M_s . Введем следующие обозначения

$$m_r = \{k : B_k \in M_r\}, \quad u_{M_r} = \{u_k : B_k \in M_r\}.$$

Определим функцию выигрыша для коалиции M_r

$$\sum_{k \in m_r} (d_k, x_k)$$

и характеристическую функцию коалиции M_r

$$V(M_r) = \max_{x_k} \sum_{k \in m_r} (d_k, x_k).$$

Положим, что $A_k = A_k(M_r)$ – технологическая матрица производственного центра B_k , т.е., технологическая матрица зависит от того, в какой именно коалиции находится центр B_k . Чтобы определить $\bar{x}_k(u_k)$, $k \in m_r$ – вектор произведенных товаров, доставляющий максимум функции выигрыша коалиции M_r – рассмотрим задачу максимизации при следующих ограничениях

$$\begin{cases} \max_{x_k} \sum_{k \in m_r} (d_k, x_k), \\ x_k A_k(M_r) \leq u_k + \alpha_k, \\ x_k \geq 0, \\ k \in m_r, u_k \in u_{M_r}. \end{cases}$$

Далее найдем значение вектора распределения ресурсов $\bar{u}_k \in u_{M_r}$, дающее максимальное значение функции выигрыша центра распределения A_0

$$\begin{cases} \max_{u_k \in u_{M_r}} \sum_{k \in m_r} c_k \bar{x}_k(u_k), \\ \sum_{u_k \in u_{M_r}} u_k = b. \end{cases}$$

Таким образом, можно определить значение максимального выигрыша коалиции M_r

$$\sum_{k \in m_r} (d_k, \bar{x}_k(\bar{u}_k)).$$

Чтобы распределить максимальный выигрыш коалиции M_r между игроками $B_k \in M_r$, будем использовать вектор Шепли. Но для этого необходимо найти максимальные значения функции выигрыша для всех коалиций, состоящих из игроков коалиции M_r .

Введем обозначения

$$\gamma = \{k : B_k \in S\}, \quad u_\gamma = \{u_k : B_k \in S\}, \quad x_\gamma = \{x_k : B_k \in S\}.$$

Определим функцию выигрыша для коалиции S

$$\sum_{k \in \gamma} (d_k, x_k)$$

и определим характеристическую функцию для коалиций $S \subset M_r$

$$V(S) = \max_{x_k} \sum_{k \in \gamma} (d_k, x_k), \quad S \subset M_r, \quad x_k \in x_\gamma.$$

Найдем из следующей задачи максимизации вектор производственных товаров \bar{x}_k , $k \in \gamma$, доставляющий максимум функции выигрыша коалиции S , учитывая, что технологическая матрица каждого производственного центра зависит только от коалиции M_r

$$\begin{cases} \max_{x_k} \sum_{k \in \gamma} (d_k, x_k), \\ x_k A_k(M_r) \leq u_k + \alpha_k, \\ x_k \geq 0, \quad k \in \gamma, \\ x_k \in x_\gamma, \quad u_k \in u_\gamma. \end{cases}$$

Найдем вектор распределения ресурсов $\bar{u}_k \in u_\gamma$, доставляющий максимум значения функции выигрыша центра распределения A_0 , рассмотрим следующую задачу максимизации:

$$\begin{cases} \max_{u_k} \sum_{k \in S} c_k \bar{x}_k(u_k), \\ \sum_{k \in \gamma} u_k = \bar{u}_{M_r}, \\ u_k \in u_\gamma. \end{cases}$$

Итак, теперь можно определить значение характеристической функции для коалиции $S \subset M_r$, оно будет выглядеть следующим образом

$$V(S) = \sum_{k \in S} (d_k, \bar{x}_k(\bar{u}_k)).$$

Тем самым получим возможность для вычисления вектора Шепли [3] для всех $B_k \in M_r$

$$Sh_k = \sum_{S \in M_r} \frac{(|S| - 1)! (|M_r| - |S|)!}{|M_r|!} (V(S) - V(S/k)).$$

3.1. Оптимальное коалиционное разбиение. В этом разделе попытаемся определить, какое коалиционное разбиение даст максимальный суммарный выигрыш для всех участников игры. Пусть Δ – множество, состоящее из всех возможных коалиционных разбиений δ_p [4]. Множество векторов распределения ресурсов при коалиционном разбиении δ_p будем обозначать

$$U_{\delta_p} = \{u_k, k = 1, \dots, n\}.$$

Множество векторов производственных товаров для коалиции $M_r \in \delta_p$ будем обозначать

$$X_{\delta_p}^{M_r} = \{x_k, k = 1, \dots, |M_r|\}.$$

Пусть

$$H(A_0, \delta_p, u_k) = \sum_{k=1}^n c_k x_k(u_k), \quad u_k \in U_{\delta_p}, \quad x_k \in X_{\delta_p}^{M_r}, \quad c_k \geq 0$$

это функция выигрыша распределительного центра A_0 в случае коалиционного разбиения δ_p , а

$$H(M_r, \delta_p, x_k) = \sum_{k \in M_r} (d_k, x_k), \quad x_k \in X_{\delta_p}^{M_r}$$

это функция выигрыша коалиции M_r в случае коалиционного разбиения δ_p .

Пусть

$$\sum_{M_r \in \delta_p} H(M_r, \delta_p, x_k) + H(A_0, \delta_p, u_k),$$

$$x_k \in X_{\delta_p}^{M_r}, \quad u_k \in U_{\delta_p}$$

является функцией суммарного выигрыша для всех коалиций и распределительного центра A_0 в случае разбиения δ_p . Оптимальное коалиционное разбиение можно определить, решив следующую задачу

$$\bar{\delta} = \operatorname{argmax}_{\delta_p \in \Delta} \left(\sum_{M_r \in \delta_p} H(M_r, \delta_p, x_k) + H(A_0, \delta_p, u_k) \right), \quad \bar{\delta} \in \Delta.$$

4. Заключение. В работе рассмотрена иерархическая игра, а также коалиционный случай этой игры с условием того, что технологическая матрица каждого игрока зависит от коалиции, которой этот игрок принадлежит, и сделаны предположения о нахождении оптимальной коалиционной структуры.

Литература

1. Петросян Л.А., Зенкевич Н.А., Семина Е.А. Теория игр М.: Высшая школа, 1998. 302 с.
2. Aumann R., Dreze J. Cooperative games with coalition structure // Int. J. Game theory, 1974. V. 4. P. 217–237.
3. Shapley L. S. A Value for n-person games. //Annals of Mathematical Studies. V. 28. P. 307–317
4. Петросян Л.А., Мамкина С.И. Игры с переменным коалиционным разбиением // Вестник СПбГУ. Сер. 1, 2003. Вып. 3. С. 60–69.

Григорьева К.В., Рединских Н.Д.
 Санкт-Петербургский государственный университет

Абсолютное равновесие в одной игре

Рассматривается многошаговая игра, в которой все ситуации являются абсолютными равновесиями.

Пусть задан конечный древовидный граф $\Gamma = (Z, L)$, где Z – множество вершин графа, а L – точечно-множественное отображение, заданное на множестве Z : $L(z) \subset Z$, $z \in Z$. Конечный древовидный граф с начальной вершиной z_0 будем обозначать через $\Gamma(z_0)$.

Рассмотрим многошаговую игру

$$G(z_0) = \langle N, \{U_i\}_{i \in N}, \{K_i\}_{i \in N} \rangle,$$

на древовидном графе $\Gamma(z_0)$, где $N = \{1, \dots, n\}$ – множество игроков, U_i – множество стратегий u_i игрока i , $u(z_0) = \{u_1(z_0), \dots, u_n(z_0)\}$ – ситуация в игре, $u_i \in U_i$, $U = \prod_{i=1}^n U_i$ – множество ситуаций в игре, $K_i(u(z_0))$ – функция выигрыша игрока i , $i = \overline{1, n}$, определенная на множестве ситуаций U .

Пусть ситуации $u(z_0)$ соответствует путь $p_l = \{z_0, \dots, z_k, \dots, z_l\}$, где $z_k \in L(z_{k-1})$, $L(z_l) = \emptyset$. Выигрыш $H_i(z_l)$ в окончательной позиции z_l соответствующего пути p_l определяется следующим образом:

$$H_i(x_l) = K_i(u(z_0)), \quad i = \overline{1, n}.$$

Напомним следующие понятия [1, 2].

Определение 1. Подыгрой $G(z)$ игры $G(z_0)$ называется многошаговая игра $G(z)$ с начальной вершиной $z \in Z$.

Определение 2. Ситуация $\bar{u}(z) = (\bar{u}_1(z), \dots, \bar{u}_i(z), \dots, \bar{u}_n(z))$ называется равновесной по Нэшу (NE – Nash Equilibrium), если

$$K_i(\bar{u}(z)) \geq K_i(\bar{u}(z) \parallel u_i(z)) \quad \forall i \in N \quad \forall u_i(z) \in U_i,$$

где $(\bar{u}(z) \parallel u_i(z)) = (\bar{u}_1(z), \dots, u_i(z), \dots, \bar{u}_n(z))$.

Определение 3. Ситуация NE $\bar{u}(z_0)$ называется ситуацией абсолютного NE в игре G , если для любого $z \in Z$ ситуация $\bar{u}(z) = (\bar{u}_1(z), \dots, \bar{u}_n(z))$ является ситуацией NE в подыгре $G(z)$.

Имеет место следующая основная

Теорема. В любой многошаговой игре с полной информацией на конечном древовидном графе существует ситуация абсолютного NE.

Пример. Рассмотрим игру в позиционной форме (см. рис. 1). Пусть $a > b$. Очевидно, что в этой игре на древовидном графе каждой пути p_l , $l = \overline{1, 8}$, соответствует ситуация абсолютного NE. Поэтому любая ситуация в этой игре является ситуацией абсолютного NE.

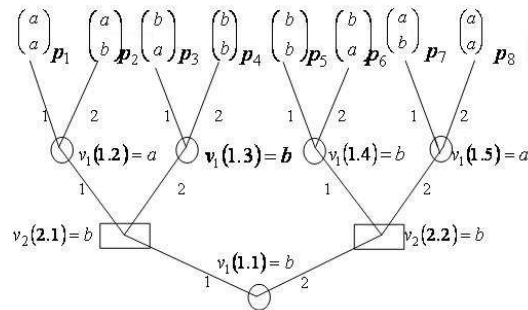


Рис. 1. Игра G

Гарантированные выигрыши в каждой позиции этой игры равны (см. рис. 1):

$$\begin{aligned} v_1(1.1) &= b, & v_1(1.2) &= a, & v_1(1.3) &= b, & v_1(1.4) &= b, & v_1(1.5) &= a, \\ v_2(1.1) &= b, & v_2(1.2) &= b, & v_2(1.3) &= b, & v_2(1.4) &= b, & v_2(1.5) &= b, \\ v_1(2.1) &= b, & v_1(2.2) &= b, \\ v_2(2.1) &= b, & v_2(2.2) &= b. \end{aligned}$$

Обозначим игру, изображенную на рис. 1, через G и расширим игру G на один шаг назад. Тогда получим игру \tilde{G} (см. рис. 2).

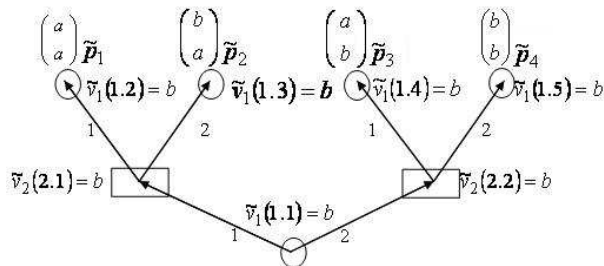


Рис. 2. Игра \tilde{G}

Очевидно, что в этой игре при таком выборе стратегий в соответствующих играх G (см. рис. 2) все пути \tilde{p}_i , $i = \overline{1, 4}$, являются равновесными. Поясним, что гарантированные выигрыши в окончательных позициях в игре \tilde{G} на рис. 2 являются гарантированными выигрышами в начальной позиции в игре G (см. рис. 1).

Также, если в соответствующих играх G будут выбраны стратегии, которые приводят к одинаковым выигрышам во всех четырех играх G , то снова получаем, что *все* пути в игре \tilde{G} равновесны, а поскольку в любой ситуации реализуется какой-либо путь в игре, то *любая* ситуация в игре \tilde{G} является ситуацией абсолютного НЕ.

Заметим также, что пары антагонистических игр (см. рис. 3–6), соответствующие играм, представленным на рис. 1, 2, дают закономерность, отраженную в приведенном ниже утверждении 1.

Из рис. 3–6 следует, что *все* стратегии игрока 1 в игре G_1 – *равновесные*, и *все* стратегии игрока 2 в игре G_2 – *равновесные*.

Следует отметить, что в конечношаговой игре со структурой, представленной на рис. 1, для выигрышей игроков выполняется

$$H_i(x_i) \geq v_i(z_0), \quad i = 1, 2.$$

Утверждение 1. *Для того, чтобы в многошаговой игре G двух игроков все исходы были абсолютными равновесиями, необходимо и достаточно, чтобы во вспомогательных антагонистических играх G_1 и G_2 все стратегии игроков 1 и 2 соответственно были равновесными.*

Доказательство. Доказательство следует из алгоритма построения пути p_i , образованного оптимальными стратегиями u_{11}, u_{22} в соответствующих вспомогательных антагонистических играх G_1, G_2 в стратегиях наказания.

Необходимость. Предположим, что в игре G все пути p_i – НЕ и пусть в играх G_1, G_2 хотя бы одна стратегия u_{11} или u_{22} не оптимальна, тогда существует ситуация в игре G_1 или G_2 – не НЕ, следовательно, существует путь p_i в игре G , не являющийся НЕ, что будет противоречием.

Достаточность. Пусть G_1, G_2 такие, что все стратегии u_{11}, u_{22} – НЕ. Игра G_1 – игра первого игрока против второго, G_2 – игра второго игрока против первого. Их стратегии u_{11} и u_{22} соответственно определяют путь p_i , о котором договариваются игроки, а u_{12} и u_{21} определяют стратегии наказания игроков второго и первого соответственно, в случае отклонения от пути p_i . Предположим, что

$u(z_0)$ в игре G не является абсолютным равновесием. Но ситуации $u(z_0)$ соответствует какой-либо путь \tilde{p}_l , образованный соответствующими стратегиями u_{11} и u_{22} в играх G_1 и G_2 . Однако все ситуации (u_{11}, u_{22}) являются НЕ по предположению. Следовательно, получено противоречие.

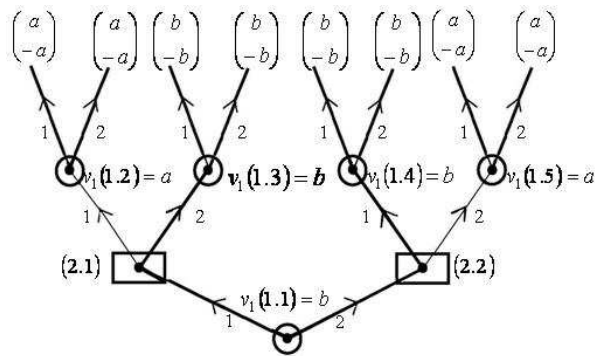


Рис. 3. Игра G_1

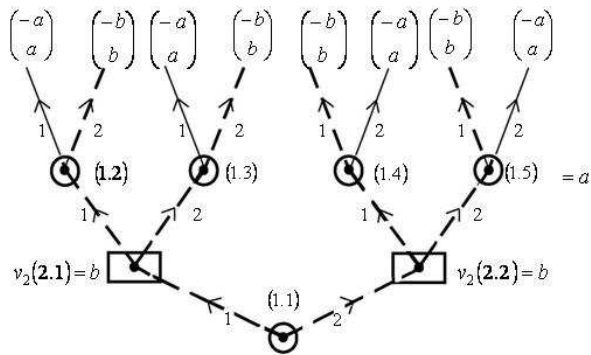


Рис. 4. Игра G_2

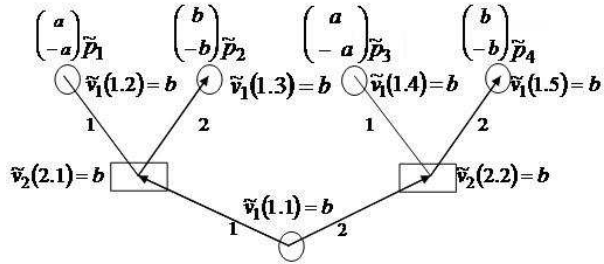


Рис. 5. Игра \tilde{G}_1

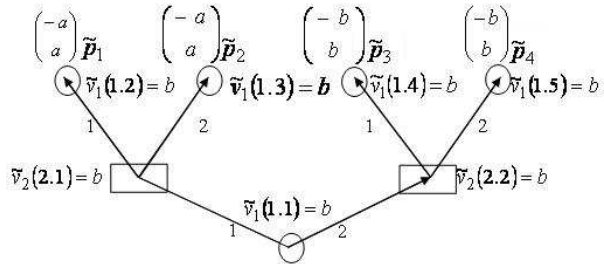


Рис. 6. Игра \tilde{G}_2

Заключение. Подводя итог проведённым рассуждениям можно утверждать, что в любой конечношаговой игре со структурой, представленной на рис. 1, все ситуации являются ситуациями абсолютного NE.

Литература

1. Григорьева К.В. Кооперативные игры и игры в позиционной форме. Часть 2. Учебное пособие. СПб.: СПбГАСУ, 2009. 134 с.
2. Петросян Л.А., Зенкевич Н.А., Семина Е.А. Теория игр. М.: Высшая школа, 1998. 304 с.

Громов И.А., Слобожанин Н.М.
Санкт-Петербургский государственный университет

О распределении пятипараметрического ресурса

1. Введение. Пусть имеется конечная совокупность объектов, каждый из которых может обладать одним или несколькими свойствами из множества свойств $\{\varepsilon_1, \dots, \varepsilon_p\}$. Будем моделировать объект со свойствами как вектор-строку нулей и единиц, причём единица на j -ом месте будет означать, что объект обладает ε_j свойством.

Введём ряд определений [1–3].

Определение 1. Слоем толщины d будем называть конечное множество $(0,1)$ -векторов таких, что сумма их одноименных компонент равна d .

Определение 2. Слой будем называть неразложимым, если он не может быть представлен в виде объединения двух слоёв.

Определение 3. Разбиение слоя на слои, каждый из которых является неразложимым, будем называть остовным разбиением.

Определение 4. Универсальным вектором будем называть вектор, целиком состоящий из единиц.

Определение 5. Предуниверсальным вектором будем называть вектор, у которого одна компонента равна 0, а все остальные равны 1.

Определение 6. Вектор длины четыре будем называть двойкой, если две его компоненты равны 1, а две другие равны 0.

В работах [1–3] были доказаны необходимые и достаточные условия существования разложения слоя толщиной d на слои заданной толщины. Целью данной работы является изучение неразложимых слоёв для трёх-, четырёх- и пятипараметрического ресурса.

2. Неразложимые слои четырёхпараметрического ресурса. Рассмотрим случай, когда исходных свойств четыре, т.е. $p = 4$. Будем искать всевозможные неразложимые слои. В [3] было доказано, что любой слой толщины $d \geq 4$ будет разложим, поэтому достаточно ограничиться слоями толщиной до 4.

Справедливы следующие леммы.

Лемма 1. Если в слое толщины $d \geq 1$ есть универсальный вектор, то из этого слоя можно выделить слой толщины 1.

Лемма 2. Если в слое толщины $d \geq 3$ есть предуниверсальные векторы четырёх типов, то из этого слоя можно выделить слой толщины 3.

Лемма 3. Если в слое толщины $d \geq 3$ есть предуниверсальные векторы трёх типов, то из этого слоя можно выделить слой толщины 1 или 2.

Лемма 4. Если в слое толщины $d \geq 2$ есть предуниверсальные векторы только двух типов, то из этого слоя можно выделить слой толщины 1 или 2.

Лемма 5. Если в слое толщины $d \geq 3$ есть предуниверсальный векторы только одного типа, то из этого слоя можно выделить слой толщины 1, 2 или 3.

Лемма 6. Из слоя толщины $d \geq 2$, в реализации которого нет предуниверсальных векторов, можно выделить слой толщины 1 или 2.

Перечислим все неразложимые слои:

1. Один слой вида: 1 1 1 1

2. Один слой, в котором есть предуниверсальные векторы четырёх типов:

$$\begin{array}{cccc} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{array}$$

3. Слои, в которых есть предуниверсальные векторы двух типов. Количество таких слоёв определяется типом предуниверсальных векторов, их будет $C_4^2 = 6$. Например,

$$\begin{array}{cccc} 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{array}$$

4. Слои, в которых есть предуниверсальный вектор только одного типа:

а) слой толщины 1. Их количество, равное четырём, определяется количеством разных предуниверсальных слоёв. Например,

$$\begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \end{array}$$

b) слои толщины 2. Существует четыре различных типа предуниверсальных векторов, для каждого из которых имеется ровно три пары двоек с фиксированной единицей. Таким образом, слоёв такого вида будет 12. Например,

$$\begin{array}{cccc} 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 \end{array}$$

с) слои толщины 3. Количество слоёв определяется количеством различных типов предуниверсального вектора. Таких слоёв будет 4.

$$\begin{array}{cccc} 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \end{array}$$

5. Слои, в которых нет ни одного предуниверсального вектора:

а) один слой вида

$$\begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}$$

б) слои толщины 1, состоящие из двоек. Нетрудно показать, что их количество равно $C_4^2/2 = 3$. Например,

$$\begin{array}{cccc} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{array}$$

с) слои толщины 1, в которых есть одна двойка. Их будет столько, сколькими способами можно расставить две единицы в четырёхбитовой строке, т.е. $C_4^2 = 6$. Например,

$$\begin{array}{cccc} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}$$

d) слои толщины 2, в которых есть три двойки. Например,

$$\begin{array}{cccc} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \end{array}$$

Количество слоёв определяется позицией единицы для двух первых строк. Таких слоёв будет 4.

Всего неразложимых слоёв в случае четырёхпараметрического ресурса будет 42. Исходя из лемм 1–6 нетрудно показать, что это будут все неразложимые слои для четырёхпараметрического ресурса. Данный аналитический результат полностью подтверждается экспериментально с помощью программы, которая ищет все неразложимые слои.

3. Нарушение закономерности неразложимости для пятипараметрического ресурса.

Лемма 7. Пусть $\{x_1, \dots, x_n\}$ слой толщины d , тогда инверсный слой $\{\neg x_1, \dots, \neg x_n\}$ – слой толщины $n - d$.

Доказательство. На i -ом месте у $\{x_1, \dots, x_n\}$ ровно $n - d$ нулей, а это значит, что инверсный слой будет иметь ровно $n - d$ единиц на i -ом месте.

Лемма 8. Если слой разложим, то инверсный слой также разложимым.

Доказательство. Рассмотрим некоторое разложение исходного слоя $\{x_1, \dots, x_n\}$ на слои $\{s_1, \dots, s_k\}$. Тогда нетрудно показать, что множество слоёв $\{\neg s_1, \dots, \neg s_k\}$ будет разложением слоя $\{\neg x_1, \dots, \neg x_n\}$.

Лемма 9. Для пятипараметрического ресурса существует неразложимый слой толщины 5.

Доказательство. Построим неразложимый слой толщины 2:

$$\begin{array}{cccccc} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \end{array}$$

Поскольку данный слой неразложим, то и инверсный слой будет неразложимым по лемме 8

$$\begin{array}{cccccc}
 1 & 1 & 1 & 1 & 0 \\
 1 & 1 & 1 & 1 & 0 \\
 1 & 1 & 1 & 0 & 1 \\
 1 & 1 & 1 & 0 & 1 \\
 1 & 0 & 0 & 1 & 1 \\
 0 & 1 & 0 & 1 & 1 \\
 0 & 0 & 1 & 1 & 1
 \end{array}$$

Таким образом, для случая пятипараметрического ресурса нарушается закономерность разложимости любого слоя с толщиной, большей или равной количеству параметров.

4. Построение неразложимых слоёв для пятипараметрического ресурса. Построим все неразложимые слои с толщиной до шести включительно для пятипараметрического ресурса. Каждую битовую строку в слое можно рассматривать как целое число. Таким образом, слой толщины d , состоящий из k -битовых строк, можно рассматривать как разбиение числа $S = d(2^p - 1)$ на $d \leq k \leq pd$ слагаемых, каждое из которых не превосходит $2^p - 1$. Сначала построим все неразложимые слои толщины $d = 1$, после чего будем строить неразложимые слои толщины $d = 2$ и т.д.

Шаг 0. Будем начинать с разбиения $S = \lfloor (d \frac{2^p - 1}{2^p - 2}) \rfloor (2^p - 2) + (d(2^p - 1) \bmod (2^p - 2))$. Если это не слой, то переходим к шагу 1 и продолжаем строить разбиения, иначе переходим к шагу 2.

Шаг 1. Пусть теперь есть некоторое разбиение числа $S = s_1 + \dots + s_k$ и пусть $i_0 = \max\{i | s_i > 1\}$. Обозначим $A = s_{i_0} + k - i_0$, тогда $A = \lfloor \frac{A}{s_{i_0} - 1} \rfloor (s_{i_0} - 1) + (A \bmod (s_{i_0} - 1))$. Далее, если окажется, что $\lfloor \frac{A}{s_{i_0} - 1} \rfloor > d$, то $A = d(s_{i_0} - 1) + (\lfloor \frac{A}{s_{i_0} - 1} \rfloor - d)(s_{i_0} - 1) + (A \bmod (s_{i_0} - 1))$. Тогда $S = s_1 + \dots + s_{i_0 - 1} + A = s_1 + \dots + s_{i_0 - 1} + \tilde{s}_{i_0} + \dots + \tilde{s}_m$.

Шаг 2. Проверяем, является ли полученное разбиение слоем, после чего делаем проверку на неразложимость, при условии, что неразложимые слои толщины, меньшей d , известны.

Таким образом, мы пройдем по всем слоям без повторов, а, поскольку в каждом слое строки будут упорядочены, то можно будет легко организовать проверку неразложимости.

В результате было получено, что неразложимых слоёв толщины

меньше 6 будет 1881. Кроме того, оказалось, что все слои с толщиной 6 и 7 разложимы. Таким образом, выдвигается следующая

Гипотеза. *Для пятипараметрического ресурса все слои с толщиной большей, чем 5, разложимы.*

5. Результаты. Получены все неразложимые слои для четырёхпараметрического ресурса аналитически и подтверждены практически с помощью программы. Было показано, что для пятипараметрического ресурса нарушается закономерность разложимости слоя с толщиной, большей либо равной, чем количество параметров. Реализована программа, которая находит неразложимые слои до заданной толщины.

Литература

1. Котина С.О., Слобожанин Н.М. Эффективный алгоритм распределения трёхпараметрического ресурса // Процессы управления и устойчивость: Труды 36-й научной конференции аспирантов и студентов / Под ред. Н.В. Смирнова, В.Н. Старкова. СПб.: Изд-во СПбГУ, 2005. С. 503–506.
2. Слобожанин Н.М., Чумак Л.И. Эффективный алгоритм распределения четырёхпараметрического ресурса // Процессы управления и устойчивость: Труды 36-й научной конференции аспирантов и студентов / Под ред. Н.В. Смирнова, В.Н. Старкова. СПб.: Изд-во СПбГУ, 2005. С. 560–564.
3. Слобожанин Н.М. Разбиение совокупности объектов на классы по заданным свойствам // Вопросы оптимального управления и исследования операций. Иркутск: Изд-во ИГУ, 1988. С. 130–142.

Захаров А.О.

Санкт-Петербургский государственный университет

Сужение множества Парето на основе нечеткой замкнутой информации¹

Рекомендовано к публикации профессором Ногиным В.Д.

1. Введение. Многокритериальные задачи принятия решений являются актуальными на сегодняшний день и занимают неотъемлемую часть современной жизни, поскольку почти любую проблему можно рассматривать с различных сторон. Желания человека (лица, принимающего решение – ЛПР) чрезвычайно разнообразны по отношению к предметам выбора и к описывающим их критериям. В связи с чем, возникает необходимость в нахождении «оптимального», в определенном смысле, варианта, который в наибольшей степени удовлетворял бы потребностям ЛПР.

При выявлении предпочтений ЛПР часто оказывается, что оно уверено в своих суждениях лишь в определенной степени. Чтобы учесть этот факт разработан аксиоматический подход сужения множества Парето, обобщенный на случай нечеткого отношения предпочтения ЛПР [1, 2]. Эта методология заключается в построении оценки (нечеткого множества) для нечеткого множества выбираемых векторов на основе дополнительной информации, выявленной у ЛПР. В качестве такой информации в данной работе будем рассматривать нечеткую замкнутую информацию.

2. Многокритериальная задача с нечетким отношением предпочтения. Сначала рассмотрим понятия нечеткого множества и нечеткого отношения [3], а потом перейдем к постановке многокритериальной задачи с нечетким отношением предпочтения.

Пусть X – некоторое множество, которое называется универсальным. Тогда нечеткое множество A в X определяется как множество упорядоченных пар $\{(x, \lambda_A(x))\}$ для любого $x \in X$, где $\lambda_A(\cdot)$ называется функцией принадлежности и принимает значения в интервале $[0, 1]$. Причем значение $\lambda_A(x)$ указывает степень принадлежности элемента x множеству A .

¹Работа выполнена при финансовой поддержке РФФИ, грант № 11-07-00449

Нечетким бинарным отношением в X называется нечеткое множество в декартовом произведении $X \times X$.

Введем множество (четкое) возможных решений X , из которого ЛПР делает свой выбор. Цели ЛПР отражает заданный числовой векторный критерий f , отображающий множество возможных решений X в пространство \mathbb{R}^m , которое называется критериальным. При этом образ $Y = f(X)$ есть множество возможных векторов (оценок). Выбор ЛПР из этого множества представляет собой нечеткое множество выбираемых векторов $C(Y)$ с функцией принадлежности $\lambda_Y^C(\cdot)$.

Для выявления вкусов и предпочтений у ЛПР вводится нечеткое бинарное отношение предпочтения на Y . Оно задается при помощи функции принадлежности $\mu_Y: Y \times Y \rightarrow [0, 1]$ такой, что соотношение $\mu_Y(y', y'') = \mu^*$ означает выбор ЛПР вектора y' вместо y'' со степенью уверенности μ^* . При этом часто, имея ввиду нечеткое отношение с функцией принадлежности μ_Y , будем говорить просто о нечетком отношении μ_Y .

Следуя аксиоматическому подходу [4], не будем стараться искать напрямую нечеткое множество выбираемых векторов $C(Y)$, а предпримем попытки его оценки, показав, таким образом, некую «нечеткую» границу, за которой не существует «наилучших», по взглядам ЛПР, решений.

Таким образом, сформулирована многокритериальная задача с нечетким отношением предпочтения в терминах векторов (оценок) (Y, μ_Y) , как это было сделано в [1, 2].

Аксиоматический подход базируется на «разумных» аксиомах нечеткого выбора [1, 2]. Они соответствуют аксиомам многокритериального выбора с четким отношением предпочтения. Однако, здесь отношение предпочтения ЛПР и множество выбираемых векторов (решений) являются нечеткими. Это аксиомы недоминирования, продолжимости, согласованности и инвариантности.

Фундаментальную роль играет нечеткий принцип Эджворта – Парето [1], согласно которому для всех $y \in Y$ имеет место $\lambda_Y^C(y) \leq \lambda_Y^P(y)$, где $\lambda_Y^P(\cdot)$ – функция принадлежности множества Парето (четкого) $P(Y) = \{y \in Y \mid \nexists y^* \in Y : y^* \geq y\}$. При этом его функция принадлежности принимает значение 1, если $y \in P(Y)$, и 0 – в противном случае.

Данный принцип очерчивает для ЛПР область его нечеткого выбора, при условии соблюдения вышеупомянутых «разумных» аксиом. Справедливо стремление упростить выбор, сужая область ком-

промиссов. Однако, это можно сделать лишь при использовании дополнительной нечеткой информации о вкусах и предпочтениях ЛППР.

3. Замкнутая информация с нечетким отношением предпочтения. В работе [5] рассматривалась четкая замкнутая информация об отношении предпочтения ЛППР, с помощью которой производилось сужение множества Парето. Теперь перенесем данный случай на многокритериальный выбор с нечетким отношением предпочтения, в связи с чем введем следующее определение.

Пусть $I = \{1, \dots, m\}$ и $I_k \subseteq I$, $I_k = \{i_1, \dots, i_k\}$, $k \geq 2$.

Определение 1. Будем говорить, что задана *нечеткая замкнутая информация* со степенями уверенности $\mu_1, \dots, \mu_k \in (0, 1]$ и с положительными скалярными параметрами $w_{i_1}^{(1)}, w_{i_2}^{(1)}, w_{i_2}^{(2)}, w_{i_3}^{(2)}, \dots, w_{i_{k-1}}^{(k-1)}, w_{i_k}^{(k-1)}, w_{i_k}^{(k)}, w_{i_1}^{(k)}$, если для векторов $y^{(1)}, \dots, y^{(k)} \in \mathbb{R}^m$ с компонентами

$$\begin{aligned} y_{i_1}^{(1)} &= w_{i_1}^{(1)}, & y_{i_2}^{(1)} &= -w_{i_2}^{(1)}, & y_s^{(1)} &= 0 \quad \forall s \in I \setminus \{i_1, i_2\}; \\ y_{i_2}^{(2)} &= w_{i_2}^{(2)}, & y_{i_3}^{(2)} &= -w_{i_3}^{(2)}, & y_s^{(2)} &= 0 \quad \forall s \in I \setminus \{i_2, i_3\}; \\ &\dots & & & & \\ y_{i_k}^{(k)} &= w_{i_k}^{(k)}, & y_{i_1}^{(k)} &= -w_{i_1}^{(k)}, & y_s^{(k)} &= 0 \quad \forall s \in I \setminus \{i_k, i_1\} \end{aligned} \quad (1)$$

справедливы равенства $\mu(y^{(1)}, 0_m) = \mu_1, \dots, \mu(y^{(k)}, 0_m) = \mu_k$.

Поскольку замкнутая информация задается при помощи набора векторов, необходимо поставить вопрос о непротиворечивости такой информации. В [2] определено понятие непротиворечивости для информации о нечетком отношении предпочтения и получено общее условие непротиворечивости.

Набор векторов $y^{(i)} \in N^m$, $i = \overline{1, k}$, $N^m = \mathbb{R}^m \setminus (\mathbb{R}_+^m \cup \mathbb{R}_-^m \cup \{0_m\})$, в совокупности с набором чисел $\mu_1, \dots, \mu_k \in (0, 1]$ называют непротиворечивым, если найдется нечеткое бинарное отношение предпочтения с функцией принадлежности $\mu(\cdot, \cdot)$, удовлетворяющее «разумным» аксиомам нечеткого выбора (аксиомам 2–4 [1, 2]), и такое, что $\mu(y^{(i)}, 0_m) = \mu_i$, $i = \overline{1, k}$.

В соответствии с [2] обозначим $\mu_{11}, \dots, \mu_{1k_1}, \dots, \mu_{l1}, \dots, \mu_{lk_l}$ такую перестановку чисел μ_1, \dots, μ_k , что

$$1 \geq \mu_{11} = \dots = \mu_{1k_1} > \dots > \mu_{l1} = \dots = \mu_{lk_l} > 0,$$

где $k = k_1 + \dots + k_l$, $1 \leq l \leq k$. Введем четкие конусы K_h ,

$h \in \{1, \dots, l\}$, порожденные единичными векторами e^1, \dots, e^m пространства \mathbb{R}^m и векторами $y^{(i)}$, $i \in \{1, \dots, k\}$, такими, что соответствующие числа μ_i удовлетворяют неравенству $\mu_i \geq \mu_{h1}$.

Введем матрицу

$$W = \begin{pmatrix} w_{i_1}^{(1)} & 0 & \dots & 0 & -w_{i_1}^{(k)} \\ -w_{i_2}^{(1)} & w_{i_2}^{(2)} & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & w_{i_{k-1}}^{(k-1)} & 0 \\ 0 & 0 & \dots & -w_{i_k}^{(k-1)} & w_{i_k}^{(k)} \end{pmatrix}.$$

Нечеткая информация, введенная в определении 1, непротиворечива только в том случае, когда $|W| > 0$. Это справедливо, поскольку первое условие теоремы 4.1 из [2] выполнено тогда и только тогда, когда $|W| > 0$. Второе условие этой теоремы также будет выполнено, так как для замкнутой нечеткой информации всегда можно построить такие конусы K_h , что каждый из них при $h \in \{2, \dots, l\}$ не будет содержать ни одного вектора $y^{(j)}$, $j \in \{1, \dots, k\}$, которому соответствует число μ_j такое, что $\mu_j < \mu_{h1}$.

4. Сужение множества Парето. Пусть задана нечеткая замкнутая информация об относительной важности трех критериев со степенями уверенности μ_1, μ_2, μ_3 , т.е. для векторов (1), где $k = 3$, $i_1 = i, i_2 = j, i_3 = l$, справедливо $\mu(y^{(1)}, 0_m) = \mu_1, \mu(y^{(2)}, 0_m) = \mu_2, \mu(y^{(3)}, 0_m) = \mu_3$. При этом выполнены неравенства $\mu_1 \geq \mu_2 \geq \mu_3$. Обозначим такую информацию через (I).

Рассмотрим функцию принадлежности

$$\lambda_Y^M(y) = 1 - \sup_{z \in Y} \zeta(z, y) \quad \forall y \in Y, \quad (2)$$

где

$$\zeta(z, y) = \begin{cases} 1, & \text{если } z - y \in \mathbb{R}_+^m, \\ \mu_1, & \text{если } \bar{z} - \bar{y} \in \mathbb{R}_+^m, z - y \notin \mathbb{R}_+^m, \\ \mu_2, & \text{если } \tilde{z} - \tilde{y} \in \mathbb{R}_+^m, \bar{z} - \bar{y} \in \mathbb{R}_+^m, \\ \mu_3, & \text{если } \hat{z} - \hat{y} \in \mathbb{R}_+^m, \tilde{z} - \tilde{y} \in \mathbb{R}_+^m, \\ 0, & \text{в остальных случаях,} \end{cases} \quad \forall y, z \in Y, y \neq z,$$

а векторы \bar{a} , \tilde{a} , \hat{a} , где $a = (a_1, \dots, a_m) \in \{y, z\}$, выглядят следующим образом (e^i, e^j, e^l – единичные векторы пространства \mathbb{R}^m):

$$\begin{aligned}\bar{a} &= a + (w_j^{(1)} a_i + (w_i^{(1)} - 1) a_j) e^j, \quad \tilde{a} = a + (w_j^{(1)} a_i + (w_i^{(1)} - 1) a_j) e^j + \\ &\quad + (w_l^{(2)} w_j^{(1)} a_i + w_l^{(2)} w_i^{(1)} a_j + (w_j^{(2)} w_i^{(1)} - 1) a_l) e^l, \\ \hat{a} &= a + ((w_j^{(2)} w_l^{(3)} - 1) a_i + w_l^{(2)} w_i^{(3)} a_j + w_j^{(2)} w_i^{(3)} a_l) e^i + (w_j^{(1)} w_l^{(3)} a_i + \\ &\quad + (w_l^{(3)} w_i^{(1)} - 1) a_j + w_j^{(1)} w_i^{(3)} a_l) e^j + (w_l^{(2)} w_j^{(1)} a_i + w_l^{(2)} w_i^{(1)} a_j + \\ &\quad + (w_j^{(2)} w_i^{(1)} - 1) a_l) e^l.\end{aligned}$$

Теорема 1. Пусть задана непротиворечивая ($|W| > 0$) нечеткая замкнутая информация (I). Тогда для любой функции принадлежности $\lambda_Y^C(\cdot)$ нечеткого множества выбираемых векторов $C(Y)$ справедливо неравенство $\lambda_Y^C(y) \leq \lambda_Y^M(y) \leq \lambda_Y^P(y)$ для всех $y \in Y$, где функция $\lambda_Y^M(\cdot)$ определяется по (2).

Заметим, что под ситуацию (I) также попадают случаи, когда $\mu_3 \geq \mu_1 \geq \mu_2$ и $\mu_2 \geq \mu_3 \geq \mu_1$. В этом можно легко убедиться переименованием критериев.

Теперь пусть также задана нечеткая замкнутая информация об относительной важности трех критериев, как и в начале данного пункта, но справедливы неравенства $\mu_1 > \mu_3 > \mu_2$. Обозначим такую ситуацию через (II).

Как и ранее, рассмотрим функцию принадлежности $\lambda_Y^M(\cdot)$, определенную в (2), где нечеткое отношение

$$\zeta(z, y) = \begin{cases} 1, & \text{если } z - y \in \mathbb{R}_+^m, \\ \mu_1, & \text{если } \bar{z} - \bar{y} \in \mathbb{R}_+^m, \quad z - y \notin \mathbb{R}_+^m, \\ \mu_3, & \text{если } \tilde{z} - \tilde{y} \in \mathbb{R}_+^m, \quad \bar{z} - \bar{y} \in \mathbb{R}_+^m, \quad \forall y, z \in Y, y \neq z, \\ \mu_2, & \text{если } \hat{z} - \hat{y} \in \mathbb{R}_+^m, \quad \tilde{z} - \tilde{y} \in \mathbb{R}_+^m, \\ 0, & \text{в остальных случаях,} \end{cases} \quad (3)$$

причем векторы \bar{b} , \tilde{b} , \hat{b} , где $b = (b_1, \dots, b_m) \in \{y, z\}$, таковы, что $\bar{b} = \bar{a}$, $\hat{b} = \hat{a}$ и

$$\begin{aligned}\tilde{b} &= b + (w_i^{(3)} b_l + (w_l^{(3)} - 1) b_i) e^i + \\ &\quad + (w_j^{(1)} w_l^{(3)} b_i + (w_i^{(1)} w_l^{(3)} - 1) b_j + w_j^{(1)} w_i^{(3)} b_l) e^j.\end{aligned}$$

Теорема 2. Пусть задана непротиворечивая ($|W| > 0$) нечеткая замкнутая информация (II). Тогда для любой функции принадлежности $\lambda_Y^C(\cdot)$ нечеткого множества выбираемых векторов $C(Y)$ справедливо неравенство $\lambda_Y^C(y) \leq \lambda_Y^M(y) \leq \lambda_Y^P(y)$ для всех $y \in Y$, где функция $\lambda_Y^M(\cdot)$ определяется по (2), а нечеткое отношение ζ по (3).

Нетрудно получить, что двойные неравенства $\mu_2 > \mu_1 > \mu_3$ и $\mu_3 > \mu_2 > \mu_1$ при соответствующих переименованиях критериев можно привести к двойному неравенству $\mu_1 > \mu_3 > \mu_2$, т.е. прийти к ситуации (II).

Таким образом, рассмотрены все возможные неравенства, связывающие числа μ_1, μ_2, μ_3 , поскольку число возможных перестановок (μ_1, μ_2, μ_3) равно 6.

Конструктивно результаты теорем 1, 2 говорят о том, что для сужения множества Парето, используя нечеткую замкнутую информацию, необходимо решить четыре многокритериальных задачи (с четким отношением предпочтения).

Литература

1. Ногин В.Д. Принцип Эджворта – Парето и относительная важность критериев в случае нечеткого отношения предпочтения // Журнал вычислительной математики и математической физики, 2003. Т. 43. № 11. С. 1676–1686.
2. Богданова А.В., Ногин В.Д. Сужение множества Парето на основе простейших наборов нечёткой информации об относительной важности критериев // Вестник СПбГУ. Сер. 10, 2007. Вып. 2. С. 3–17.
3. Zadeh L.A. Fuzzy sets // Information and Control, 1965. Vol. 8. P. 338–353.
4. Ногин В.Д. Принятие решений в многокритериальной среде: количественный подход. М.: Физматлит, 2005. 176 с.
5. Захаров А.О. Сужение множества Парето на основе замкнутой информации об отношении предпочтения ЛПР // Вестник СПбГУ. Сер. 10, 2009. Вып. 4. С. 69–83.

Земцова В.Н.

Санкт-Петербургский государственный университет

Нечеткая суммарно-разностная функция выбора

Рекомендовано к публикации профессором Ногиным В.Д.

Введение и постановка задачи. Моделирование процессов принятия решений в экономике и социологии базируется на понятиях отношения предпочтения и функции выбора [2]. Учет в моделях принятия решения нечеткости предпочтений, присущих человеческому мышлению, позволяет не только упорядочить альтернативы в порядке убывания значимости для ЛПП, но и сузить множество рассматриваемых альтернатив. Для этого требуется построить нечеткие функции выбора, базирующиеся на нечетких отношениях предпочтения, которые в литературе определяются как рациональные функции выбора.

Поставим задачу: рассмотреть новый класс нечетких функций выбора, изучить их свойства. Важно проверить, удовлетворяет ли введенная нечеткая функция выбора аксиомам наследования (H^*), согласия (C^*), константности выбора (K^*) и отбрасывания (O^*) [1]. В связи с тем, что наибольший интерес представляют рациональные функции выбора, возникает вопрос: существуют ли среди функций рассматриваемого класса иррациональные. В случае положительного ответа требуется определить условия, которым должна удовлетворять нечеткая функция выбора, чтобы полученное множество совпадало с нечетким множеством недоминируемых альтернатив, другими словами, критерий рационализуемости.

Определение нечеткой суммарно-разностной функции выбора. Пусть имеется некоторое множество Ω объектов произвольной природы и задано некоторое универсальное множество $A \subset \Omega$. Через 2^A обозначим множество всех подмножеств A .

Определение 1. Пусть (R, μ_R) – нечеткое бинарное отношение на A [2]. Определим для $\forall X \in 2^A$

$$h_R^C(X) = \left\{ x \in X \mid \sum_{z \in X} [\mu_R(x, z) - \mu_R(z, x)] \geq \sum_{z \in X} [\mu_R(y, z) - \mu_R(z, y)], \forall y \in X \right\}.$$

Будем называть h_R^C нечеткой суммарно-разностной (НСР) функцией выбора на (R, μ_R) .

Замечание. Эквивалентным образом нечеткая суммарно-разностная функция выбора на (R, μ_R) может быть записана:

$$x^* = \arg \max_{x \in X} h_R(X)(x), \quad h_R^C(X) = h_R(X)(x^*).$$

Определение 2. Функцию выбора C на A будем называть нечеткой суммарно-разностной рационализуемой (НСР-рационализуемой), если существует нечеткое бинарное отношение (R, μ_R) такое, что C – нечеткая суммарно-разностная функция выбора на (R, μ_R) , т.е. $C = h_R^C$. В частности, матричное представление $M(R, \mu_R)$ носит название нечеткой суммарно-разностной реализации C .

Пример. Пусть $A = \{x_1, x_2, x_3, x_4\}$ и нечеткое бинарное отношение (R, μ_R) на A задается матрицей

$$\begin{pmatrix} 1,00 & 0,59 & 0,35 & 0,29 \\ 0,60 & 1,00 & 0,24 & 0,27 \\ 0,40 & 0,20 & 1,00 & 0,30 \\ 0,30 & 0,28 & 0,20 & 1,00 \end{pmatrix}.$$

Покажем, что данное матричное представление является реализацией нечеткой суммарно-разностной функции выбора. Поставим задачу построить по заданной матрице выявленных предпочтений НСР-функцию выбора. Для этого нужно рассмотреть всевозможные подмножества множества A и осуществить выбор из каждого из них, согласно определению НСР-функции выбора. В результате несложных вычислений получаем, что данному матричному представлению соответствует НСР-функция выбора C на A . Её реализация на трехэлементных подмножествах представлена в таблице.

Таблица. Выбор на трехэлементных подмножествах

A	x_1, x_2, x_3	x_1, x_2, x_4	x_1, x_3, x_4	x_2, x_3, x_4
$C(A)$	x_2	x_4	x_3	x_3
$C(A)(x)$	0,05	0,02	0,15	0,06

Свойства нечеткой суммарно-разностной функции выбора. Здесь будем предполагать, что нечеткое бинарное отношение удовлетворяет условиям рефлексивности и связности [2], функция выбора C является нечеткой суммарно-разностной.

1. Для любых $x, y \in A : \mu_R(x, y) > \mu_R(y, x) \Rightarrow C(\{x, y\}) = \{x\}$.
2. Для $\forall x \in A \forall y \in A \setminus \{x\} : \mu_R(y, x) > \mu_R(x, y) \Rightarrow x \notin C(A)$.

Замечание. Выполнение свойства 1 означает, что из аксиомы наследования (H^*) следует свойство 2, но это отнюдь не означает, что из свойства 2 следует выполнение аксиомы (H^*). Получаем, что НСР-функция выбора удовлетворяет более слабому условию, чем аксиома наследования.

3. Если $\exists x \in A : \mu_R(x, y) \geq \mu_R(y, x), \forall y \in A \setminus \{x\} \Rightarrow x \notin C(A)$.

Замечание. Выполнение свойства 1 означает, что из аксиомы согласия (C^*) следует свойство:

$$\forall x \in A : \mu_R(x, y) \geq \mu_R(y, x), \quad \forall y \in A \setminus \{x\} \Rightarrow x \in C(A),$$

т.е. если указанное выше свойство не выполняется, то и аксиома (C^*) не будет иметь место. Заметим, что отрицание этого свойства как раз и есть выполнение свойства 3. По доказанному нечеткая суммарно-разностная функция выбора удовлетворяет 3. Следовательно, свойство 3 может быть переформулировано в свойство:

4. Функция выбора C не удовлетворяет аксиоме согласия (C^*).

5. Существуют нечеткие суммарно-разностные иррационализуемые функции выбора.

Таким образом, можно сделать вывод, что нечеткая суммарно-разностная функция выбора C обладает следующими свойствами:

- удовлетворяет более слабому условию, чем аксиома наследования (H^*);
- не удовлетворяет аксиоме согласия (C^*);
- существуют иррациональные функции выбора C .

С одной стороны, рационализуемость в обычном смысле следует из выполнения аксиом (H^*) и (C^*), которые в данном случае не выполняются. С другой стороны, свойство 5 говорит о том, что существуют иррациональные функции выбора C . Возникает вопрос, как тогда охарактеризовать НСР-рационализуемость.

Критерий рационализуемости. Обозначим $d(X)$ – множество индексов всех элементов, которые принадлежат X . Определим сле-

дующую функцию

$$\delta_{ij}^{lk}(X) = \begin{cases} 1 & i \neq j : i = l \text{ и } j \in d(X), j \neq k; \\ 1 & i \neq j : i \in d(X), i \neq l \text{ и } j = k; \\ -1 & i \neq j : i = k \text{ и } j \in d(X), j \neq l; \\ -1 & i \neq j : i \in d(X), i \neq k \text{ и } j = l; \\ 2 & i \neq j : i = l \text{ и } j = k; \\ -2 & i \neq j : i = k \text{ и } j = l; \\ 0 & \text{в остальных случаях;} \end{cases}$$

для любого $X \in 2^A$ и для любых $l, k \in d(X)$, $l \neq k$ и $i, j = 1, \dots, n$.
Рассмотрим вектор $u \in \mathbb{R}^{n \times n}$:

$$u = (u_{11}, \dots, u_{1n}; \dots; u_{n1}, \dots, u_{nn})$$

и функцию выбора C , заданную на A . Введем в рассмотрение два множества

$$U[C(X)] = \bigcup_{l \in d(C(X))} \bigcup_{k \in d(X) \setminus \{l\}} \{u \in \mathbb{R}^{n \times n} : u_{ij} = \delta_{ij}^{lk}(X), i, j = \overline{1, n}\},$$

$$U^+[C(X)] = \bigcup_{l \in d(C(X))} \bigcup_{k \in d(X) \setminus d(C(X))} \{u \in \mathbb{R}^{n \times n} : u_{ij} = \delta_{ij}^{lk}(X), i, j = \overline{1, n}\},$$

$$U[C] = \bigcup_{X \in 2^A} U[C(X)], \quad U^+[C] = \bigcup_{X \in 2^A} U^+[C(X)].$$

Имеет место следующая

Теорема. Для того чтобы функция выбора C на A была НСР-рационализуема, необходимо и достаточно, чтобы существовал такой вектор $w \in \mathbb{R}^{n \times n}$, что

$$wi^T \geq 0, \forall u \in U[C] \quad \text{и} \quad wi^T > 0, \forall u \in U^+[C].$$

Доказательство. Необходимость. Если C – НСР-рационализуемая функция выбора, то существует такое нечеткое бинарное отношение $(R, \mu_R) : C$ – НСР-функция выбора на (R, μ_R) . Пусть $w \in \mathbb{R}^{n \times n} : w_{ij} = \mu_R(x_i, x_j)$, $i, j = 1, \dots, n$. Распишем произведение в левой части неравенства, которое требуется доказать.

$$\begin{aligned}
wu^T &= \sum_{i,j=1}^n w_{ij}u_{ij} = \sum_{z \in A} [\mu_R(x, z) - \mu_R(y, z) - \mu_R(z, x) - \mu_R(z, y)] = \\
&= \sum_{z \in A} [\mu_R(x, z) - \mu_R(z, x)] - \sum_{z \in A} [\mu_R(y, z) - \mu_R(z, y)].
\end{aligned}$$

Из определения нечеткой суммарно-разностной функции выбора следует, что последнее выражение неотрицательно.

Достаточность. Возьмем произвольный вектор $w \in \mathbb{R}^{n \times n}$, удовлетворяющий условиям теоремы. Обозначим $\bar{w} = \max_{i,j=1,\dots,n} w_{ij}$ и $\underline{w} = \min_{i,j=1,\dots,n} w_{ij}$. Если $\bar{w} = \underline{w}$, то условие в определении нечеткой суммарно-разностной функции выбора выполняется автоматически. Рассмотрим случай, когда $\bar{w} \neq \underline{w}$. Обозначим матрицу

$$W^* = \begin{bmatrix} w_{11}^* & \cdots & w_{1n}^* \\ \vdots & \ddots & \vdots \\ w_{n1}^* & \cdots & w_{nn}^* \end{bmatrix}, \quad w_{ij}^* = \frac{w_{ij} - \underline{w}}{\bar{w} - \underline{w}}, \quad w_{ij}^* \in [0, 1].$$

Следовательно, существует нечеткое бинарное отношение (R, μ_R) такое, что C является нечеткой суммарно-разностной на (R, μ_R) .

Теорема доказана.

Следствие 1. *Для того чтобы функция выбора C на A была НСР-рационализуема, необходимо и достаточно, чтобы существовал такой вектор $w \in \mathbb{R}^{n \times n}$, что*

$$wu^T = 0, \quad \forall u \in U[C] \setminus U^+[C] \quad \text{и} \quad wu^T > 0, \quad \forall u \in U^+[C].$$

Следствие 2. *Для того чтобы однозначная функция выбора C на A была НСР-рационализуема, необходимо и достаточно, чтобы существовал такой вектор $w \in \mathbb{R}^{n \times n}$, что*

$$wu^T > 0, \quad \forall u \in U^+[C].$$

Несложно заметить, что доказательство теоремы носит конструктивный характер. В заключение приведем алгоритм построения НСР-реализации произвольной функции выбора C .

1. Пусть $i, j = 1, \dots, n$. Введите начальные значения w_{ij} .
2. Введите $u \in U[C]$ и вычислите $v = \sum_{i,j=1,\dots,n} w_{ij}u_{ij}$.
3. Если $v = 0, u \in U[C] \setminus U^+[C]$ и $v > 0, u \in U^+[C]$, то перейти на шаг 5; в противном случае перейти на следующий шаг 4.
4. Заменить значения w_{ij} на $w_{ij} + \eta u_{ij}$, где $0 < \eta < 1$.
5. Возвращаться на шаг 2 до тех пор, пока значение v не установится для всех входных значений u , т.е. $v = 0, u \in U[C] \setminus U^+[C]$ и $v > 0, u \in U^+[C]$.
6. Обозначим $w_{ij}^* = \frac{w_{ij} - \underline{w}}{\bar{w} - \underline{w}}$, где $\bar{w} = \max_{i,j=1,\dots,n} w_{ij}$ и $\underline{w} = \min_{i,j=1,\dots,n} w_{ij}$.
7. Вывести НСР-реализацию $C(W^*)$ и прекратить процесс.

Предложенная в данной работе рациональная функция выбора базируется на нечетком бинарном отношении, которое удовлетворяет условиям рефлексивности и связности, что позволяет расширить область решения проблемы выбора. На практике данная задача свелась к выбору, который легко реализовать. Отметим, что выбор может быть неоднозначным.

Литература

1. Земцова В.Н. Обобщение теорем Сена и Шварца // Процессы управления и устойчивость: Труды 41-й международной научной конференции аспирантов и студентов / Под ред. Н.В. Смирнова, Г.Ш. Тамасяна. СПб.: Издат. Дом С.-Петербур. гос. ун-та, 2010. С. 593–598.
2. Орловский С.А. Проблемы принятия решений при нечеткой исходной информации. М.: Наука, 1981. 208 с.
3. Xiao Luo. A new approach to sum – fuzzy rational choice functions // Fussy Sets and Systems, 2002. Vol. 127. P. 241–246.

Зенович О.С., Малафеев О.А.
 Санкт-Петербургский государственный университет

Оптимизационная модель формирования портфеля инвестора на фондовом рынке

Рассматривается модель рынка ценных бумаг, на котором обращаются корпоративные акции и облигации. Пусть $|I|$ – конечное множество агентов, оперирующих на рынке и оказывающих своими действиями влияние на его состояние.

Пусть, для простоты, вначале у каждого агента $i \in I$ имеются два типа ценных бумаг – облигации и акции, цены которых изменяются во времени в зависимости от управлений агентов u_i^t , выбираемых ими в каждый из моментов $t \in T = \{0, 1, \dots\}$. Множество управлений агента i в момент t обозначим через U_i^t . Набор управлений агентов $u^t = (u_1^t, \dots, u_{|I|}^t)$ назовем управлением рынка в момент t .

Агенты, выбирая свои управления в момент t , обладают информацией обо всей истории функционирования рынка $\{u^{t'}\}_{t'=0}^t$.

Состояние рынка ценных бумаг в момент t описывается наборами цен на бумаги всех агентов $i \in I$ – облигации и акции, которые обозначаются соответственно b_i^t и s_i^t . Цена b_i^t меняется в момент $t \in T$ по правилу $b_i^{t+1} = c_i^{t+1}(u^t)b_i^t$, где $c_i^t \geq 1$. Цена s_i^t изменяется следующим образом $s_i^{t+1} = \xi_i^{t+1}s_i^t$, где ξ_i^{t+1} – неизвестная априори величина со значением из компакта K_i^{t+1} с границами: верхней $\bar{K}_i^{t+1}(u^t)$ и нижней $\underline{K}_i^{t+1}(u^t)$, которые зависят от выбора u^t . Полагаем, что $0 < \underline{K}_i^{t+1} < c_i^{t+1} < \bar{K}_i^{t+1}$. Для практических приложений представляют интерес два следующих случая:

- 1) $K_i^{t+1} = \{\underline{K}_i^{t+1}, \bar{K}_i^{t+1}\}$, т.е. K_i^{t+1} состоит из двух точек;
 - 2) $K_i^{t+1} = [\underline{K}_i^{t+1}, \bar{K}_i^{t+1}]$, т.е. K_i^{t+1} – замкнутый интервал.
- Положим $K^t = (K_1^t, \dots, K_{|I|}^t)$.

Будем далее полагать, что динамика изменения множества K^{t+1} определяется следующим образом $K^{t+1} = F(K^t, u^t)$, где F – оператор, сопоставляющий всякому набору отрезков K^t и набору допустимых управлений u^t набор отрезков K^{t+1} , при $t \in T = \{0, 1, \dots\}$.

Инвестор $i \in I$, выбирая на каждом шаге t управление u_i^t , также ведёт свой портфель по следующей схеме. Если X_i^t – капитал портфеля в момент t , то инвестор выбирает количество ν_i^{t+1} акций так, что

сумму X_i^t можно представить в виде $X_i^t = \nu_i^{t+1} s_i^t + (X_i^t - \nu_i^{t+1} s_i^t)$, где выражение в скобках соответствует той части капитала, которая хранится в облигациях (на банковском счёте). Здесь $\nu_i^t \in R^1$, что означает допустимую продажу на срок без покрытия и одалживания. В момент $t+1$ становится известным значение переменной ξ^{t+1} , и капитал агента i становится равным $X_i^{t+1} = \nu_i^{t+1} \xi_i^{t+1} s_i^t + (X_i^t - \nu_i^{t+1} s_i^t) c_i^{t+1}$, если не принимать во внимание комиссионные трейдера, связанные с осуществлением сделок купли-продажи (transaction costs). Существуют различные виды комиссионных. Будем здесь рассматривать комиссионные, пропорциональные объёму сделки. Они представляются в следующем виде $Tr_C = k |\nu_i^{t+1} - \nu_i^t| s_i^{t+1}$, где k – это константа, принимающая значения между 0 и 1.

Таким образом, капитал агента i в момент времени $t+1$ с учётом пропорциональных издержек принимает следующий вид

$$X_i^{t+1} = \nu_i^{t+1} \xi_i^{t+1} s_i^t + (X_i^t - \nu_i^{t+1} s_i^t) c_i^{t+1} - k |\nu_i^{t+1} - \nu_i^t| \xi_i^{t+1} s_i^t.$$

Стратегией φ_i агента i в конечношаговом процессе назовем выбор в каждый момент времени $t \in T$ управления u_i^t и числа ν_i^t на основании информации о предыстории процесса изменения рынка, и в частности, реализовавшихся стратегий всех агентов вплоть до момента $t-1$, а также $X^0, X^1, X^2, \dots, X^{t-1}, s^0, s^1, s^2, \dots, s^{t-1}$.

Агент i , продавая опцион по цене $C_i = X_i^0$, стремится вести свой портфель таким образом, чтобы уплатить покупателю в конечный момент $\bar{t} = |T|$ премию $f_i(s_i^{\bar{t}})$, зависящую от цены акции $s_i^{\bar{t}}$. Для европейского колл-опциона (call option), согласно которому покупатель имеет право купить акции по заранее условленной цене P_i , в момент \bar{t} премия имеет вид $f_i(s_i^{\bar{t}}) = \max(0, s_i^{\bar{t}} - P_i)$. Отсюда получаем, что доход агента i равен величине $X_i^{\bar{t}} - f_i(s_i^{\bar{t}})$.

Далее удобно будет оперировать с количеством бумаг (облигаций) $Y_i^t = X_i^t / b_i^t$ в качестве капитала, который будем называть относительным, так что задача агента $i \in I$ о максимизации дохода $X_i^{\bar{t}} - f_i(s_i^{\bar{t}})$ эквивалентна задаче о максимизации $Y_i^{\bar{t}} - f_i(s_i^{\bar{t}}) / b_i^{\bar{t}}$.

Рассмотрим последний этап многошагового процесса. В начале последнего этапа (или в конце предпоследнего этапа) у агента $i \in I$ имеется относительный капитал $Y_i^{\bar{t}-1} = X_i^{\bar{t}-1} / b_i^{\bar{t}-1}$.

Поэтому относительный капитал агента $i \in I$ на последнем этапе

будет выражаться следующим образом

$$\begin{aligned}
Y_i^{\bar{t}}(\nu_i^{\bar{t}}, \xi_i^{\bar{t}}) - \frac{f_i(\xi_i^{\bar{t}} s_i^{\bar{t}-1})}{b_i^{\bar{t}}} &= \left[\nu_i^{\bar{t}} \xi_i^{\bar{t}} s_i^{\bar{t}-1} + (X_i^{\bar{t}-1} - \nu_i^{\bar{t}} s_i^{\bar{t}-1}) c_i^{\bar{t}} - \right. \\
&\quad \left. - k |\nu_i^{\bar{t}} - \nu_i^{\bar{t}-1} | \xi_i^{\bar{t}} s_i^{\bar{t}-1} - f_i(\xi_i^{\bar{t}} s_i^{\bar{t}-1}) \right] / b_i^{\bar{t}} = Y_i^{\bar{t}-1} + \\
&+ \left[s_i^{\bar{t}-1} / b_i^{\bar{t}-1} c_i^{\bar{t}}(u^{\bar{t}-1}) \right] (\nu_i^{\bar{t}}(\xi_i^{\bar{t}} - c_i^{\bar{t}}(u^{\bar{t}-1})) - k |\nu_i^{\bar{t}} - \nu_i^{\bar{t}-1} | \xi_i^{\bar{t}}) - \frac{f_i}{c_i^{\bar{t}} b_i^{\bar{t}-1}}.
\end{aligned}$$

При условии, что все агенты выбрали конкретные управления $u^{\bar{t}-1}$, гарантированный доход агента $i \in I$ на последнем шаге равен величине

$$\begin{aligned}
Y_i^{\bar{t}-1} + \frac{1}{c_i^{\bar{t}}(u^{\bar{t}-1}) b_i^{\bar{t}-1}} \max_{\nu_i^{\bar{t}}} \min_{\xi_i^{\bar{t}} \in K_i^{\bar{t}}} \left[\nu_i^{\bar{t}} s_i^{\bar{t}-1} (\xi_i^{\bar{t}} - c_i^{\bar{t}}(u^{\bar{t}-1})) - \right. \\
\left. - s_i^{\bar{t}-1} k |\nu_i^{\bar{t}} - \nu_i^{\bar{t}-1} | \xi_i^{\bar{t}} - f_i(\xi_i^{\bar{t}} s_i^{\bar{t}-1}) \right]. \tag{1}
\end{aligned}$$

Обозначим теперь через $Val_U(\Gamma_H) = Val(\Gamma_H, U)$ оператор значения бескоалиционной игры n лиц с множествами стратегий агентов U_1, \dots, U_n ($U = \prod_I U_i$) и их функциями выигрыша H_1, \dots, H_n , ($H = (H_1, \dots, H_n)$), определенными на произведении $U = \prod_I U_i$.

Допустим, что значение этого оператора единственно. Например, это может быть байесово равновесие, выбранное по схеме, предложенной в [1]. Или, в более общем виде, может быть взята однозначная ветвь многозначной функции значений функции выигрыша в равновесных ситуациях (аналогично [2]) с выполнением условия непрерывности этой функции (допущение регулярности).

Положив также

$$\begin{aligned}
(D_i^{\bar{t}} f_i)(s_i^{\bar{t}}) = D_i^{\bar{t}} f_i(s_i^{\bar{t}}, u^{\bar{t}-1}) &= \frac{1}{c_i^{\bar{t}}(u^{\bar{t}-1})} \min_{\nu_i^{\bar{t}}} \max_{\xi_i^{\bar{t}} \in K_i^{\bar{t}}} \left[\nu_i^{\bar{t}} s_i^{\bar{t}-1} (\xi_i^{\bar{t}} - \right. \\
&\quad \left. - c_i^{\bar{t}}(u^{\bar{t}-1})) - s_i^{\bar{t}-1} k |\nu_i^{\bar{t}} - \nu_i^{\bar{t}-1} | \xi_i^{\bar{t}} - f_i(\xi_i^{\bar{t}} s_i^{\bar{t}-1}, u^{\bar{t}-1}) \right], \tag{2}
\end{aligned}$$

получим выражение для (1)

$$Y_i^{\bar{t}-1} - \frac{1}{b_i^{\bar{t}-1}} D_i^{\bar{t}} f_i(s_i^{\bar{t}}, u^{\bar{t}-1}). \tag{3}$$

Можно в стандартной биномиальной модели несколько ослабить допущение, полагая f_i нестрого выпуклыми и неубывающими (аналогично [3]).

Тогда в (2) максимум достигается на концах $K_i^{\bar{t}}$, так что (2) принимает вид

$$D_i^{\bar{t}} f_i(s_i^{\bar{t}}, u^{\bar{t}-1}) = \frac{1}{c_i^{\bar{t}}(u^{\bar{t}-1})} \min_{\nu_i^{\bar{t}}} \max \left\{ f_i(\overline{K}_i^{\bar{t}} s_i^{\bar{t}-1}) - \right. \\ \left. - \nu_i^{\bar{t}} s_i^{\bar{t}-1} (\overline{K}_i^{\bar{t}} - c_i^{\bar{t}}(u^{\bar{t}-1})) + s_i^{\bar{t}-1} k |\nu_i^{\bar{t}} - \nu_i^{\bar{t}-1}| \overline{K}_i^{\bar{t}}, f_i(\underline{K}_i^{\bar{t}} s_i^{\bar{t}-1}) - \right. \\ \left. - \nu_i^{\bar{t}} s_i^{\bar{t}-1} (\underline{K}_i^{\bar{t}} - c_i^{\bar{t}}(u^{\bar{t}-1})) + s_i^{\bar{t}-1} k |\nu_i^{\bar{t}} - \nu_i^{\bar{t}-1}| \underline{K}_i^{\bar{t}} \right\}. \quad (4)$$

Таким образом, при $\nu_i^{\bar{t}} \geq \nu_{i,e}^{\bar{t}}$ (здесь рассматривается случай, когда $\nu_i^{\bar{t}} > \nu_i^{\bar{t}-1}$) второй член под знаком максимума в (4) доставляет максимум всему выражению.

Здесь

$$\nu_{i,e}^{\bar{t}-1} = \left[\frac{f_i(\overline{K}_i^{\bar{t}} s_i^{\bar{t}-1}) - f_i(\underline{K}_i^{\bar{t}} s_i^{\bar{t}-1}) - s_i^{\bar{t}-1} k \nu_i^{\bar{t}-1} (\overline{K}_i^{\bar{t}} - \underline{K}_i^{\bar{t}})}{s_i^{\bar{t}-1} [(1-k)(\overline{K}_i^{\bar{t}} - \underline{K}_i^{\bar{t}})]} \right].$$

Отсюда получаем, что минимум в (4) доставляется значением $\nu_i^{\bar{t}} = \nu_{i,e}^{\bar{t}}$. Следовательно

$$D_i^{\bar{t}} f_i(s_i^{\bar{t}}, u^{\bar{t}-1}) = \frac{1}{c_i^{\bar{t}}(u^{\bar{t}-1})} \left[\frac{c_i^{\bar{t}}(u^{\bar{t}-1}) - \underline{K}_i^{\bar{t}}}{(1-k)(\overline{K}_i^{\bar{t}} - \underline{K}_i^{\bar{t}})} f_i(\overline{K}_i^{\bar{t}} s_i^{\bar{t}-1}) + \right. \\ \left. + \frac{(\overline{K}_i^{\bar{t}} - c_i^{\bar{t}}(u^{\bar{t}-1}) - k(\overline{K}_i^{\bar{t}} - \underline{K}_i^{\bar{t}}))}{(1-k)(\overline{K}_i^{\bar{t}} - \underline{K}_i^{\bar{t}})} f_i(\underline{K}_i^{\bar{t}} s_i^{\bar{t}-1}) + \frac{(\overline{K}_i^{\bar{t}} - c_i^{\bar{t}}(u^{\bar{t}-1}))}{(1-k)} \times \right. \\ \left. \times s_i^{\bar{t}-1} k \nu_i^{\bar{t}-1} + \frac{(f_i(\overline{K}_i^{\bar{t}} s_i^{\bar{t}-1}) - f_i(\underline{K}_i^{\bar{t}} s_i^{\bar{t}-1}) - \nu_i^{\bar{t}-1} s_i^{\bar{t}-1} (\overline{K}_i^{\bar{t}} - \underline{K}_i^{\bar{t}}))}{(1-k)(\overline{K}_i^{\bar{t}} - \underline{K}_i^{\bar{t}})} k \underline{K}_i^{\bar{t}} \right].$$

Используя метод динамического программирования, можно вычислить гарантированную величину относительного дохода к моменту \bar{t} при фиксированных значениях управления u^t на каждом шаге.

Пример. Рассматривается конечношаговая игра $G = \langle A, \Gamma_{x_0}, H \rangle$, где $A = (A_1, A_2)$ – инвесторы, Γ_{x_0} – конечный древовидный граф игры G с начальной вершиной $x_0 \in X$, где X – множество вершин графа Γ , $H = (H_1, H_2)$ – функции выигрышей инвесторов A_1 и A_2 в игре G . Данная игра $G = \langle A, \Gamma_{x_0}, H \rangle$ протекает на бирже на интервале времени $[0, T]$. Будем считать, что в игре G фиксировано разбиение $\tau = \{0 = t_0 < t_1 < \dots < t_n = T\}$ интервала $[0, T]$ на отрезки длины $\Delta t_k = t_k - t_{k-1}$, $k = 1, \dots, n$. Инвестор A_i , $i = 1, 2$ на каждом шаге t_k , $k = 1, \dots, n$, переформирует свой портфель $P_i(\nu_i(t_k), \mu_i(t_k))$, где $\nu_i(t_k)$ – количество одноименных акций на интервале времени Δt_k , $k = 1, \dots, n$, в портфеле P_i , а $\mu_i(t_k)$ – количество облигаций инвестора A_i , $i = 1, 2$, на интервале Δt_k , $k = 1, \dots, n$. При этом инвестор распределяет текущие средства $D_i(t)$ между акциями (рисковый актив) и облигациями (безрисковый актив), имея целью максимизировать наличный капитал $D_i(T)$, оцениваемый стоимостью акций и облигаций в портфеле P_i , в конечный момент времени T . Инвесторы A_1 и A_2 на каждом шаге t_k , $k = 1, \dots, n$, выбирают альтернативы $\alpha_{j_i}(t_k)$, где $j_i = j_i^1, \dots, j_i^l$, l – конечное число, $k = 1, \dots, n$, $i = 1, 2$, по очереди, имея информацию обо всей истории процесса. Для простоты изложения в примере будем считать, что $n = 3, l = 2$. Тогда данную игру G можно представить в виде древовидного графа, представленного на рис. 1, где в вершинах, обозначенных кружками, ходит игрок A_1 , а в квадратах – A_2 . В скобках указаны выигрыши инвесторов A_i , $i = 1, 2$, в соответствии с выбираемыми стратегиями формирования портфеля P_i .

Опишем понятную процедуру динамического программирования отыскания ситуации равновесия в данном примере.

Обозначим через S_i , $i = 1, 2$, стратегии инвесторов A_i , $i = 1, 2$, как функции, определённые на информационных множествах I_i^{np} , где $i = 1, 2$, $n = 0, \dots, 3$, $p = 1, \dots, 8$. Тогда в соответствии с понятной процедурой динамического программирования, рассматривая подыгры $G(I_2^{31}), G(I_2^{32}), G(I_2^{33}), G(I_2^{34}), G(I_2^{35}), G(I_2^{36}), G(I_2^{37})$ и $G(I_2^{38})$, получим

$$S_2(I_2^{31}) = \alpha_2, \quad H_1(S_2(I_2^{31})) = 4, \quad H_2(S_2(I_2^{31})) = 2;$$

$$S_2(I_2^{32}) = \alpha_1, \quad H_1(S_2(I_2^{32})) = 3, \quad H_2(S_2(I_2^{32})) = 6;$$

$$S_2(I_2^{33}) = \alpha_2, \quad H_1(\cdot) = 2, \quad H_2(\cdot) = 8; \quad S_2(I_2^{34}) = \alpha_1, \quad H_1(\cdot) = 3, \quad H_2(\cdot) = 4;$$

$$S_2(I_2^{35}) = \alpha_2, \quad H_1(\cdot) = 6, \quad H_2(\cdot) = 9; \quad S_2(I_2^{36}) = \alpha_2, \quad H_1(\cdot) = 8, \quad H_2(\cdot) = 3;$$

$$S_2(I_2^{37}) = \alpha_1, H_1(\cdot) = 5, H_2(\cdot) = 6; S_2(I_2^{38}) = \alpha_1, H_1(\cdot) = 1, H_2(\cdot) = 7.$$

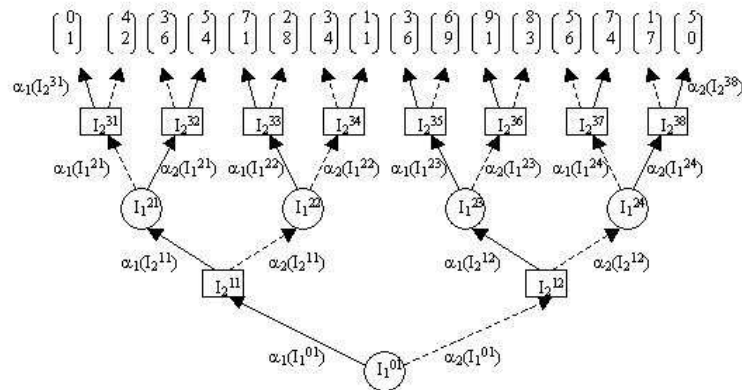


Рис. 1. Древоидный граф игры G

Далее в подыграх $G(I_1^{21})$, $G(I_1^{22})$, $G(I_1^{23})$ и $G(I_1^{24})$:

$$S_1(I_1^{21}) = \alpha_1, H_1(\cdot) = 4, H_2(\cdot) = 2; S_1(I_1^{22}) = \alpha_2, H_1(\cdot) = 3, H_2(\cdot) = 4;$$

$$S_1(I_1^{23}) = \alpha_2, H_1(\cdot) = 8, H_2(\cdot) = 3; S_1(I_1^{24}) = \alpha_1, H_1(\cdot) = 5, H_2(\cdot) = 6.$$

Затем рассматриваем подыгры $G(I_2^{11})$ и $G(I_2^{12})$:

$$S_2(I_2^{11}) = \alpha_2, H_1(\cdot) = 3, H_2(\cdot) = 4; S_2(I_2^{12}) = \alpha_2, H_1(\cdot) = 5, H_2(\cdot) = 6.$$

Наконец решаем игру $G = G(I_1^{01})$. Окончательно получаем $S_1(I_1^{01}) = \alpha_2$, $H_1(S_1(I_1^{01})) = 5$, $H_2(\cdot) = 6$. В результате имеем ситуацию абсолютного равновесия Курно – Нэша (u_1^*, u_2^*) , в которой $u_1^* = (2, 2, 2, 1, 1)$, $u_2^* = (1, 2, 2, 2, 2, 1, 2, 2, 1, 1)$, и игра приводит к выигрышу $(5, 6)$. На рис. 1 стратегии и путь обозначены пунктиром.

Литература

1. Харшаньи Дж., Зельтен Р. Общая теория выбора равновесия в играх. СПб.: Изд-во «Экономическая школа», 2001, 405 с.
2. Малафеев О.А. Управляемые конфликтные системы. СПб.: Изд-во СПбГУ, 2000, 276 с.
3. Шарп У., Александер Г., Бейли Дж. Инвестиции. М.: Инфра-М, 2010, 1028 с.

Зубарева М.Л., Губар Е.А.
Санкт-Петербургский государственный университет

Применение метода Кларка – Райта для построения инкассаторских маршрутов в сети банкоматов

1. Постановка задачи. Предположим, что существует сеть из нескольких банкоматов и банка. Банкоматы обслуживаются некоторым количеством инкассаторских бригад, которые базируются в банке. Требуется оптимизировать процесс обслуживания банкоматов, т.е. составить оптимальный маршрут их объезда таким образом, чтобы издержки были минимальны, а банкоматы заполнены наличностью [1].

Такая задача может быть сформулирована, как задача маршрутизации транспортных средств с ограничениями на вместимость (Capacitated Vehicle Routing Problem, CVRP) [2], которая является известной задачей целочисленного программирования.

В данной работе рассмотрим построение инкассаторских маршрутов приближенным методом Кларка – Райта и сравним эти маршруты с точным решением, полученным по методу ветвей и границ.

2. Задача маршрутизации транспортных средств с ограничениями на вместимость. Опишем представление задачи CVRP для построения инкассаторских маршрутов в сети банкоматов [3]. Рассматривается полный неориентированный граф $G = (V, E)$, где $V = \{0, \dots, n\}$ – это множество вершин, а E – это множество неориентированных рёбер. Вершина 0 обозначает банк, другие вершины представляют собой банкоматы. Расстояния между пунктами i и j обозначаются как $c_{ij} = c_{ji}$. Имеется M одинаковых инкассаторских машин, каждая из которых имеет вместимость $Q > 0$. Каждый банкомат имеет целочисленное значение вместимости q_i – количество денег, которое необходимо загрузить в банкомат, $0 < q_i \leq Q$. Каждый банкомат может быть обслужен только одной бригадой и ни одна бригада не может обслужить несколько банкоматов, если их общая вместимость превышает вместимость этой машины. Задача состоит в нахождении M маршрутов, в которых начальные и конечные точки совпадают с базовым пунктом, вместимость каждой машины не превышает заданного Q , а сумма издержек на маршруте минимальна.

Пусть переменная x_{ij} показывает, сколько раз инкассаторская машина проследует от пункта i до пункта j . В данном случае $x_{ij} = x_{ji}$, так как задача неориентированная. Пусть множество $V_c = V \setminus \{0\}$ обозначает множество банкоматов без банка и дано некоторое подмножество банкоматов $S \subseteq V_c$. Обозначим $r(S)$ – минимальное количество машин, которое требуется для обслуживания банкоматов из S . Тогда формулировка задачи целочисленного программирования будет следующей:

$$\min \sum_{i < j} c_{ij} x_{ij}, \quad (1)$$

$$\sum_{i < j} x_{ij} + \sum_{i > j} x_{ji} = 2, \quad i \in V_c, \quad (2)$$

$$\sum_{i \in S} \sum_{j \in V \setminus S} x_{ij} \geq 2r(S), \quad S \subseteq V_c, \quad |S| > 1, \quad (3)$$

$$\sum_{j \in V_c} x_{0j} = 2M, \quad (4)$$

$$x_{ij} \in \{0, 1\}, \quad i \in V_c, \quad j \in V_c, \quad i < j, \quad (5)$$

$$x_{0j} \in \{0, 1, 2\}, \quad j \in V_c. \quad (6)$$

Ограничения (2) гарантируют, что инкассаторской бригадой банкомат обслуживается единожды. Ограничения (3) означают, что $r(S)$ бригад проходят через S банкоматов, где S – это любое подмножество банкоматов. Неравенства (3) остаются правомерными, если $r(S)$ заменить нижней границей этих величин $\lceil \sum_{i \in S} q_i / Q \rceil$, где $\lceil y \rceil$ означает наименьшее целое, большее, чем y . Ограничения (4) выражают то, что M бригад должны покинуть центр и M бригад должны вернуться обратно. Ограничения (5) и (6) являются ограничениями на целочисленность переменных. Переменная $x_{0j} = 2$ соответствует маршруту, содержащему только один банкомат j .

3. Методы решения CVRP. Существуют различные методы решения CVRP, одни из них позволяют найти точное решение задачи, другие – приближенное. Среди точных методов наиболее известны:

- метод ветвей и границ;
- метод отсекающих плоскостей.

Методами, дающими приближенное решение, являются следующие эвристические методы:

- метод Кларка – Райта (последовательный и параллельный);
- генетические алгоритмы;
- алгоритмы на основе муравьиных колоний.

Так как CVRP относится к классу NP-трудных задач, то применение точных методов становится затруднительным при больших размерностях задачи. Эвристические же методы дают приближенное решение, но достигают его сравнительно быстро, при этом получаемые решения достаточно точны для требуемых целей.

Рассмотрим реализацию эвристического метода Кларка – Райта на примере конкретной сети банкоматов г. Санкт-Петербурга.

4. Метод Кларка – Райта и его применение. Алгоритм Кларка – Райта – это наиболее известный и широко распространенный эвристический метод для решения задачи маршрутизации транспорта [4]. В его основе лежит процесс слияния мелких маршрутов в более крупные, проводимый до тех пор, пока есть возможность уменьшить суммарные издержки. Центральную роль в этом алгоритме играет понятие выигрыша S_{ij} – это снижение суммарных издержек на маршрутах $0 - i - 0$ и $0 - j - 0$, получаемое при их объединении в маршрут $0 - i - j - 0$:

$$S_{ij} = c_{i0} + c_{0j} - c_{ij}.$$

Существуют две разновидности рассматриваемого алгоритма: последовательный алгоритм и параллельный. В последовательном алгоритме только один маршрут строится за один шаг, в то время как в параллельном одновременно могут строиться сразу несколько маршрутов.

За начальное решение принимается совокупность маршрутов, в которой к каждому пункту следует отдельная машина, т.е. количество таких маршрутов совпадает с количеством получателей. На первом шаге алгоритма подсчитываются S_{ij} для всех пар пунктов, и все пары пунктов сортируются в порядке убывания выигрышей S_{ij} (учитываются только положительные выигрыши). Затем рассматривается пара (i, j) с максимальным значением выигрыша, которая включается в один маршрут, если выполнены следующие условия:

1. Пара (i, j) не рассматривалась на предыдущих шагах алгоритма.
2. Пункты i и j не входят в состав одного и того же маршрута.

3. Пункты i и j являются начальным и/или конечным пунктом тех маршрутов, в состав которых они входят.
4. Сумма q_i на получаемом маршруте не превышает вместимость машины Q .

Если хотя бы одно из условий не выполняется, переходим к следующему значению S_{ij} . Шаги повторяются до окончательного составления маршрута. В последовательном алгоритме требуется несколько пробеганий по отсортированному списку выигрышей, в параллельном – только одно.

Рассмотрим применение данного метода для составления оптимальных маршрутов инкассаторских бригад в некоторой сети банкоматов г. Санкт-Петербурга. Для примера рассмотрим заявку, согласно которой вскоре потребуют загрузки банкоматы на следующих станциях метро: 2 – Технологический институт, 3 – Московские ворота, 4 – Ломоносовская, 5 – Василеостровская, 6 – Проспект Большевиков, 7 – Площадь Ленина, 8 – Нарвская, 9 – Чкаловская и 10 – Сенная Площадь. Банк располагается по адресу Невский пр., д. 160. Предположим, что в каждый банкомат требуется загрузить $q_i = 3$ кассеты, а емкость инкассаторской машины $Q = 12$. Расстояния между банкоматами и банком приведены в таблице 1.

Таблица 1. Расстояния между банкоматами и банком, м

	Банк	2	3	4	5	6	7	8	9	10
Банк	0	3250	6530	9000	5005	10007	6680	7810	7650	3940
2	3250	0	2930	10000	4870	13500	5480	3860	6770	1280
3	6530	2930	0	10120	7940	13070	10610	5410	9180	4050
4	9000	10000	10120	0	13690	6000	11900	14500	15100	10540
5	5005	4870	7940	13690	0	15300	5990	5970	2750	4030
6	10007	13500	13070	6000	15300	0	11100	14560	14600	10480
7	6680	5480	10610	11900	5990	11100	0	9070	4690	6500
8	7810	3860	5410	14500	5970	14560	9070	0	8300	4670
9	7650	6770	9180	15100	2750	14600	4690	8300	0	5010
10	3940	1280	4050	10540	4030	10480	6500	4670	5010	0

Для приведенных расстояний подсчитываются выигрыши S_{ij} , представленные в таблице 2. Затем выигрыши сортируются по убыванию.

Таблица 2. Таблица выигрышей S_{ij} , м

	2	3	4	5	6	7	8	9	10
2	0	6850	2250	3385	-243	4450	7200	4130	5910
3	6850	0	5410	3595	3467	2600	8930	5000	6420
4	2250	5410	0	315	13007	3780	2310	1550	2400
5	3385	3595	315	0	-288	5695	6845	9905	4915
6	-243	3467	13070	-288	0	5587	3257	3057	3467
7	4450	2600	3780	5695	5587	0	5420	9640	4120
8	7200	8930	2310	6845	3257	5420	0	7160	7080
9	4130	5000	1550	9905	3057	9640	7160	0	6580
10	5910	6420	2400	4915	3467	4120	7080	6580	0

Проследим действие последовательного алгоритма. Сначала рассматривается пара банкоматов (4, 6) с максимальным значением $S_{4,6} = 13007$. Данные банкоматы удовлетворяют условиям 1–4, значит, они могут быть включены в один маршрут: $0 - 4 - 6 - 0$, где 0 обозначает банк. Следующая пара банкоматов (5, 9) не рассматривается, так как при их объединении мы получили бы уже два маршрута ($0 - 4 - 6 - 0$ и $0 - 5 - 9 - 0$), а в последовательном алгоритме два маршрута одновременно строиться не могут. Далее рассматривается пара банкоматов (6, 7), при объединении которых получаем маршрут $0 - 4 - 6 - 7 - 0$. Аналогичным образом продолжается применение алгоритма и для других пар. В результате получаем решение, состоящее из трех маршрутов: $0 - 4 - 6 - 7 - 8 - 0$, $0 - 3 - 10 - 5 - 9 - 0$ и $0 - 2 - 0$. Суммарная длина маршрута равна 72690 м.

Рассмотрим параллельный алгоритм. Аналогично первому случаю, пара банкоматов (4, 6) включается в один маршрут. Пара (5, 9), удовлетворяющая условиям 1–4, также объединяется в одном маршруте. Следующей рассматривается пара (7, 9) и, таким образом, получаем маршрут $0 - 5 - 9 - 7 - 0$. Затем пары банкоматов (3, 8) и (2, 8) включаются в маршрут $0 - 2 - 3 - 8 - 0$. Пара (8, 9) не рассматривается, так как нарушается условие 3. Продолжая аналогичным образом применение алгоритма, получаем в итоге три маршрута: $0 - 4 - 6 - 0$, $0 - 5 - 9 - 7 - 0$ и $0 - 2 - 3 - 8 - 10 - 0$. Суммарная длина маршрута при этом равна 64332 м.

При помощи метода ветвей и границ подсчитано и точное решение данной задачи. Ограничения (2)–(6) целочисленной формулировки CVRP были заданы в пакете Maple и с помощью функции LPSolve, реализующей метод ветвей и границ, получено решение: $0 - 4 - 6 - 0$, $0 - 5 - 9 - 7 - 0$ и $0 - 2 - 3 - 8 - 10 - 0$ с суммарной длиной маршрута 64332 м. Построенные маршруты даны на рис. 1.

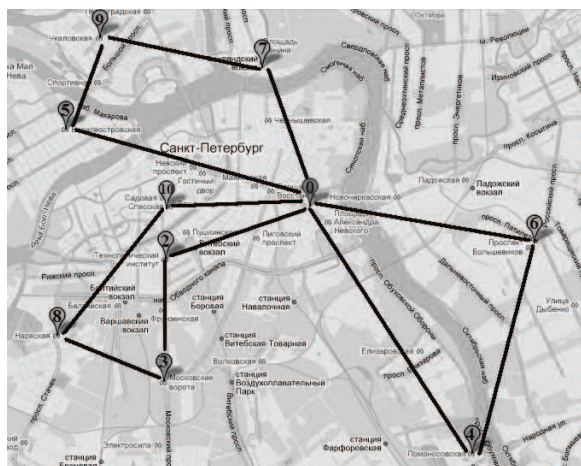


Рис. 1. Оптимальные маршруты инкассаторских бригад

5. Заключение. В вышеизложенном примере мы получили, что решение, вычисленное с помощью параллельного алгоритма Кларка – Райта, совпадает с точным решением для рассматриваемой конкретной задачи. Последовательный же алгоритм Кларка – Райта приводит к решению, которое отличается от оптимального на 11%.

Литература

1. Журнал ПЛАС – Прогнозирование остатков наличных денежных средств в АТМ.
http://www.plusworld.ru/experts-position/page1_2796.php
2. Ralphs T.K., Kopman L., Pulleyblank W.R., Trotter L.E. On the capacitated vehicle routing problem // *Mathematical Programming*, 2003. P. 343–359.
3. Bodin L., Maniezzo V., Mingozzi A. Street routing and scheduling problems // *Handbook of Transportation Science* / Ed. by R.W. Hall. New York: Kluwer Academic Publishers, 2003. P. 413–449.
4. Clarke G., Wright J.W. Scheduling of vehicles from a central depot to a number of delivery points // *Operations Research*, 1964. Vol. 12, № 4. P. 568–581.

Иванченко Е. В., Колбин В. В.
Санкт-Петербургский государственный университет

Особенности моделей ценообразования европейского и американского типов опционов в условиях нечетких данных

Введение. Под опционом понимают оплаченное право, но не обязательство, купить или продать некоторые базовые активы на заранее определенных условиях при наступлении срока или до него.

В качестве базовых активов могут служить как классические финансовые инструменты (акции, облигации и т.д.) и различные товары (нефть, золото, зерно), так и некоторые индексы (например, индексы фондового рынка) и валютные курсы. Различают опционы на право покупки (option call) и на право продажи (option put).

Опцион, который может быть реализован только в строго оговоренный в контракте день, называют европейским. Если опцион может быть исполнен в любой момент на протяжении всего срока действия контракта, то такой опцион называют американским.

Как правило, из-за большой степени риска, опционы рассматривают только на малых промежутках времени (сроком до одного года). В этой статье будет предложен метод рассмотрения опционов на больших промежутках времени, основанный на теории нечетких множеств.

Модель Блэка – Шоулза (BS) предназначена для моделирования процесса ценообразования европейского типа опционов. Данная модель подробно рассмотрена в [1]. Дифференциальное уравнение Блэка – Шоулза – Мертона с краевыми условиями, соответственно, для call и put опционов имеет вид:

$$\frac{1}{2}\sigma^2 S^2 \frac{\partial^2 f}{\partial S^2} + bS \frac{\partial f}{\partial S} - rf + \frac{\partial f}{\partial t} = 0, \quad (1)$$

$$f = \max(S - E, 0), \quad (2)$$

$$f = \max(E - S, 0), \quad (3)$$

Здесь f – цена опциона, S – текущая цена актива, E – цена исполнения, σ – волатильность относительных изменений, b – инвестиционный параметр, r – непрерывная безрисковая процентная ставка.

Решая (1) с краевыми условиями (2) и (3), получим

$$c = Se^{-qt}N(d_1) - Ee^{-rt}N(d_2),$$

$$p = Ee^{-rt}N(-d_2) - Se^{-qt}N(-d_1),$$

где c – цена опциона call, p – цена опциона put,

$$d_1 = \frac{\ln\left(\frac{S}{E}\right) + \left(r + \frac{\sigma^2}{2}\right)t}{\sigma\sqrt{t}},$$

$$d_2 = \frac{\ln\left(\frac{S}{E}\right) + \left(r - \frac{\sigma^2}{2}\right)t}{\sigma\sqrt{t}} = d_1 - \sigma\sqrt{t}.$$

Модель Барона – Адези – Уэйли (BAW) предназначена для моделирования процесса ценообразования американского типа опционов. Эта модель является модификацией модели BS и основывается на дифференциальном уравнении Блэка – Шоулза – Мертона. Данная модель подробно рассмотрена в [2].

Для американского опциона call выражение для цены опциона имеет следующий вид:

$$C(S, T) = c(S, T) + A_2 \left(\frac{S}{S^*}\right)^{\beta_2}, \quad S < S^*,$$

$$C(S, T) = S - E, \quad S \geq S^*,$$

где C – цена американского опциона call, c – цена европейского опциона call, S^* – критическое значение курса акций.

$$A_2 = \frac{(1 - e^{(b-r)T}N(d_1(S^*)))}{\beta_2} S^*,$$

$$\beta_2 = -\frac{\left(N - 1 + \sqrt{(N - 1)^2 + \frac{4M}{K}}\right)}{2}, \quad (4)$$

$$S^* - E = c(S^*, T) + \frac{(1 - e^{(b-r)T} N(d_1(S^*)))}{\beta_2} S^*,$$

Для американского опциона put выражение для цены опциона имеет следующий вид:

$$P(S, T) = p(S, T) + A_1 \left(\frac{S}{S^{**}}\right)^{\beta_1}, \quad S > S^{**},$$

$$P(S, T) = E - S, \quad S \leq S^{**},$$

где P – цена американского опциона put, p – цена европейского опциона put, S^{**} – критическое значение курса акций.

$$A_1 = -\frac{(1 - e^{(b-r)T} N(-d_1(S^{**})))}{\beta_1} S^{**},$$

$$\beta_1 = -\frac{\left(N - 1 - \sqrt{(N - 1)^2 + \frac{4M}{K}}\right)}{2}, \quad (5)$$

$$E - S^{**} = p(S^{**}, T) - \frac{(1 - e^{(b-r)T} N(-d_1(S^{**})))}{\beta_1} S^{**}.$$

Критические значения курса акций S^* и S^{**} находят, применяя к выражениям (4) и (5) численный метод Ньютона – Рафсона.

Сравнение моделей. В основе модели BS лежит предположение о том, что рынок опционов подчиняется нормальному закону распределения. На деле это не так. Следовательно, с точки зрения рисковости, модель BS является существенно более сложной, чем модель BAW. Однако, при незначительной волатильности (колебаниях) цены актива, а также при небольшом сроке действия опциона (обычно до полугода), результаты рассмотренных выше моделей практически не отличаются. Если имеется портфель из подлежащих опционам активов, то здесь удобнее применить модель ценообразования

американских опционов, в том числе модель BAW. В этом случае, указанная модель применяется для того, чтобы не требовать синхронизации срока жизни портфеля на подлежащих опционам активах и сроков соответствующих опционных контрактов. К тому же, вероятностное моделирование американского типа опционов, по своей трудоемкости, является более простым, нежели вероятностное моделирование европейского типа опционов.

Нечетко-множественный подход. Покажем применение приведенных в статье моделей ценообразования опционов с учетом временного фактора. В качестве подлежащего опциону актива используется жилищно-коммунальное хозяйство города Санкт-Петербурга, рассчитанное на двадцать лет вперед с предположениями о его ежегодном внутреннем десятипроцентном росте и тринадцатипроцентном рыночном росте. По своей структуре актив ЖКХ является портфелем, поскольку его составляющие не принадлежат городу (разные собственники). Актив ЖКХ формируется в двадцатилетнем объеме и представляется на бирже в качестве опциона. Соответственно, возникает проблема рассмотрения финансовых рисков как с точки зрения владельца актива (города), которому принадлежит опцион put на указанный актив, так и потенциального покупателя (инвестора), который покупает опцион call на тот же актив.

На формирование цены актива ЖКХ оказывают влияние ряд факторов, одним из которых является погода. Например, если за один год наблюдалась холодная зима и теплое лето, а на другой – была противоположная ситуация, то, естественно, годовые объемы активов ЖКХ, в рассмотренных случаях, не будут одинаковыми, причем, может случиться и так, что они будут существенно отличаться друг от друга. А таких случаев – двадцать. При этом создается рискованная ситуация, связанная с неопределенностью, а также изменяются параметры, характеризующие опцион с точки зрения рисковости. На основании теории нечетких множеств можно варьировать цену актива ЖКХ (представленного в двадцатилетнем объеме) в зависимости от погоды и, тем самым, получать данные, которые будут являться более точными по сравнению с обычной оценкой.

Поскольку указанные опционы рассматриваются на двадцатилетнем временном промежутке, то в этом случае следует применить теорию нечетких множеств. В качестве носителя будет выступать актив ЖКХ в двадцатилетнем объеме. Стоимость этого актива будет выступать в роли нечеткого множества. На стоимость актива ЖКХ оказывает фактор времени года. Предлагается разбить весь год на четыре интервала (зима, весна, лето, осень) и рассматривать указанный актив на каждом интервале в течение двадцати лет. Причем на каждом интервале будет введена функция принадлежности, отражающая степень условности, с которой носитель принадлежит нечеткому множеству.

Соответственно, на основе изменений параметров чувствительности цен опционов, а также некоторых других показателей можно анализировать сложившуюся ситуацию и в дальнейшем сделать прогноз относительно поведения рассмотренного актива ЖКХ.

Заключение. Рассмотренный в данной статье метод будет использован в дипломной работе. На его основе будут получены результаты, согласно которым будут составлены функционалы оптимизации, отражающие поведение потенциального инвестора с точки зрения его психологии (рискофоб, рисконейтрал и рискофил).

Литература

1. Hull J.K. Options, futures and other derivatives. Prentice-Hall, Fifth edition, 2003. 700 p.
2. Barone-Adesi G., Whaley R. Efficient analytic approximation of american option values // The Journal of Finance, 1987. Vol. 42. P. 301–320.

Ильина А.В., Петросян Л.А.
Санкт-Петербургский государственный университет

Условие Д.В.К. Янга для динамических игр с нетрансферабельными выигрышами

1. Введение. Впервые условие устойчивости против иррационального поведения игроков, или условие Д.В.К. Янга, было введено в статье [1]. Под «иррациональностью» поведения понимается отказ игрока от кооперативного соглашения. В статье [1] предложено условие, в соответствии с которым иррациональные действия игрока в какой-то момент времени не приведут к тому, что его результирующий выигрыш будет меньше, чем выигрыш в случае изначального отсутствия кооперации. Д.В.К. Янг рассмотрел динамические игры с трансферабельными выигрышами. В данной статье рассмотрено условие устойчивости против иррационального поведения для динамических игр с нетрансферабельными выигрышами.

2. Постановка задачи. Рассмотрим неантагонистическую дифференциальную игру $\Gamma(x_0, [t_0, +\infty))$ двух лиц из начального состояния x_0 продолжительности $[t_0, +\infty)$. Состояние игры изменяется в соответствии с системой дифференциальных уравнений

$$\dot{x}(t) = f[t, x(t), u_1(t), u_2(t)], \quad t \in [t_0, +\infty), \quad x(t_0) = x_0. \quad (1)$$

Мгновенный выигрыш игрока $i \in \{1, 2\}$ в момент времени $t \in [t_0, +\infty]$ обозначим через $g_i[t, x(t), u_1(t), u_2(t)]$. Выигрыши являются нетрансферабельными, т.е. они не могут быть измерены в единой шкале, и перераспределения выигрыша невозможно. Выигрыш игрока i определяется по формуле

$$K_i(x_0, [t_0, +\infty), u_1(\cdot), u_2(\cdot)) = \int_{t_0}^{+\infty} e^{-\rho(t-t_0)} g_i[t, x(t), u_1(t), u_2(t)] dt,$$

где ρ – ставка дисконтирования. Предположим, что игроки решили кооперироваться. Тогда будем считать, что они выбирают пару оптимальных управлений $u^*(t) = (u_1^*(t), u_2^*(t))$, которые определяют

Парето-оптимальный исход в игре $\Gamma(x_0, [t_0, +\infty))$

$$\lambda_i(x_0, [t_0, +\infty)) = \int_{t_0}^{+\infty} e^{-\rho(t-t_0)} g_i[t, x^*(t), u_1^*(t), u_2^*(t)] dt,$$

$x^*(t)$ – Парето-оптимальная траектория, найденная путем подстановки оптимальных управлений $u^*(t) = (u_1^*(t), u_2^*(t))$ в (1) и решения соответствующего дифференциального уравнения. В случае отсутствия кооперации игроки будут выбирать равновесие по Нэшу $\bar{u}(t) = (\bar{u}_1(t), \bar{u}_2(t))$, которое вычисляется с помощью уравнения Айзекса – Беллмана. Выигрыш i -го игрока в равновесии по Нэшу вдоль некооперативной траектории $\bar{x}(t)$ равен

$$w_i(x_0, [t_0, +\infty)) = \int_{t_0}^{+\infty} e^{-\rho(t-t_0)} g_i[t, \bar{x}(t), \bar{u}_1(t), \bar{u}_2(t)] dt, \quad i = 1, 2.$$

3. Динамическая устойчивость. Важнейшим свойством кооперативного решения является динамическая устойчивость решения. Динамическая устойчивость гарантирует сохранение оптимальности первоначально выбранного решения в каждый момент времени вдоль оптимальной траектории. В играх с трансферабельными выигрышами динамическая устойчивость обеспечивается с помощью введения процедуры распределения дележа. Для того, чтобы добиться динамической устойчивости решения игры с нетрансферабельными выигрышами необходимо ввести понятие процедуры распределения выигрыша [2].

Определение 1. Функция $\beta_i(t)$ называется *процедурой распределения выигрыша* (ПРВ), если она удовлетворяет равенству

$$\lambda_i(x_0, [t_0, +\infty)) = \int_{t_0}^{+\infty} e^{-\rho(t-t_0)} \beta_i(t) dt, \quad i = 1, 2.$$

Определение 2. Парето-оптимальное решение называется динамически устойчивым, если существует ПРВ такая, что выполняется условие индивидуальной рациональности

$$\int_t^{+\infty} e^{-\rho(\tau-t_0)} \beta_i(\tau) d\tau \geq e^{-\rho(t-t_0)} w_i(x^*(t), [t, +\infty)), \quad i = 1, 2,$$

где $w_i(x^*(t), [t, +\infty))$ – выигрыш i -го игрока в некотором равновесии по Нэшу с начальным состоянием на кооперативной траектории $x^*(t)$.

Теорема 1. *Парето-оптимальное решение является динамически устойчивым, если функции $w_i(x^*(t), [t, +\infty))$ являются дифференцируемыми. Тогда существует неотрицательная функция $\eta_i(t)$, $t \in [t_0, +\infty)$ такая, что выполнено условие*

$$\lambda_i(x_0, [t_0, +\infty)) - w_i(x_0, [t_0, +\infty)) = \int_{t_0}^{\infty} e^{-\rho(\tau-t_0)} \eta_i(\tau) d\tau, \quad i = 1, 2.$$

В этом случае ПРВ $\beta_i(t)$ может быть вычислена по формуле

$$\beta_i(t) = \eta_i(t) + \rho w_i(x^*(t), [t, +\infty)) - \frac{d}{dt} w_i(x^*(t), [t, +\infty)), \quad i = 1, 2. \quad (2)$$

4. Условие Д.В.К. Янга. Пусть в момент времени $t \in [t_0, +\infty)$ иррациональные действия игроков привели к отказу от кооперативной схемы. Запишем условие устойчивости против иррационального поведения

$$w_i(x_0, [t_0, +\infty)) \leq \int_{t_0}^t e^{-\rho(\tau-t_0)} \beta_i(\tau) d\tau + e^{-\rho(t-t_0)} w_i(x^*(t), [t, +\infty)),$$

где $\beta_i(t)$ – динамически устойчивая ПРВ. Теперь запишем достаточное условие выполнения условия Д.В.К. Янга.

Теорема 2. *Для того, чтобы условие Д.В.К. Янга было выполнено для игрока i на всем отрезке $[t_0, +\infty)$, достаточно, чтобы выполнялось условие*

$$\beta_i(t) \geq -\frac{d}{dt} w_i(x^*(t), [t, +\infty)) + \rho w_i(x^*(t), [t, +\infty)).$$

Покажем теперь, что условие Д.В.К. Янга выполняется для динамически-устойчивого (состоятельного во времени) Парето-оптимального решения. Если Парето-оптимальное решение является состоятельным во времени, то ПРВ можно вычислять по формуле (2). Перепишем формулу (2) в виде

$$\beta_i(t) - \eta_i(t) = \rho w_i(x^*(t), [t, +\infty)) - \frac{d}{dt} w_i(x^*(t), [t, +\infty)), \quad i = 1, 2,$$

где $\eta_i(t)$ – неотрицательная функция, т.е. $\eta_i(t) \geq 0$. Тогда достаточное условие Д.В.К. Янга выполнено для любой ПРВ, удовлетворяющей равенству (2).

5. Пример. В качестве демонстрации техники вычисления ПРВ и проверки условия Д.В.К. Янга приведем модель сокращения вредных выбросов в атмосферу. В качестве Парето-оптимальных решений рассмотрим выигрыши игроков, которые реализуют максимальный суммарный выигрыш.

5.1. Постановка задачи. Пусть I – это множество игроков (стран), участвующих в игре сокращения вредных выбросов, $I = \{1, 2\}$.

Выбросы игрока $i \in I$ в момент времени $t \in [0, +\infty)$ обозначим $u_i(t)$. Пусть $x(t)$ – динамика накопления загрязнения за время t , которая определяется дифференциальным уравнением

$$\dot{x}(t) = \sum_{i \in I} u_i(t) - \delta x(t), \quad x(0) = x_0, \quad (3)$$

где δ – коэффициент, характеризующий долю поглощенного загрязнения. Пусть $C_i(u_i)$ – издержки на уменьшение выбросов, которые несет страна i , если она удерживает выбросы на некотором допустимом уровне u_i :

$$C_i(u_i(t)) = \frac{\gamma}{2} (u_i(t) - \bar{u}_i)^2, \quad 0 \leq u_i(t) \leq \bar{u}_i, \quad \gamma > 0.$$

Пусть $D_i(x)$ – издержки возмещения ущерба от загрязнения.

$$D_i(x) = \pi x(t), \quad \pi > 0.$$

Каждый игрок стремится минимизировать суммарные издержки

$$K_i(x_0, u, [0; +\infty)) = \int_0^{+\infty} e^{-\rho t} (C_i(u_i(t)) + D_i(x(t))) dt, \quad (4)$$

где $u = (u_1, \dots, u_n)$ – это ситуация в игре, а ρ – ставка дисконтирования. Будем считать, в отличие от [3], что выигрыши нетрансферабельные.

5.2. Решение задачи. Неравенства, введенные в пункте 3, изменятся на противоположные, так как в рассматриваемом Парето-оптимальном решении игроки стремятся минимизировать суммарные издержки. Условие Д.В.К. Янга для задачи сокращения вредных выбросов:

$$w_i(x_0, [0, +\infty)) \geq \int_0^t e^{-\rho\tau} \beta_i(\tau) d\tau + e^{-\rho t} w_i(x^*(t), [t, +\infty)),$$

где $w_i(x_0, [0, +\infty))$ – выигрыш игрока i за время $[0, +\infty)$ вдоль некооперативной траектории в некотором равновесии по Нэшу. Функция $\beta_i(\tau)$ – динамически-устойчивая ПРВ. Пусть $w_i(x^*(t), [t, +\infty))$ – выигрыш игрока i за время $[0, +\infty)$ с начальным состоянием на кооперативной траектории в некотором равновесии по Нэшу.

Решая уравнение Гамильтона – Якоби – Беллмана [3], найдем

$$w_i(x_0, [0, +\infty)) = \frac{\pi}{\rho(\rho + \delta)} \left(\frac{\pi}{2\gamma(\rho + \delta)} + \sum_{i \in I} \bar{u}_i - \frac{2\pi}{\gamma(\rho + \delta)} + \rho x_0 \right).$$

Выигрыш игрока i (компонента минимального суммарного выигрыша) определяется следующим образом

$$\lambda_i(x_0, [0, +\infty)) = \frac{\pi}{\rho(\rho + \delta)} \left(\sum_{i \in I} \bar{u}_i - \frac{2\pi}{\gamma(\rho + \delta)} + \rho x_0 \right).$$

Соответствующие кооперативные Парето-оптимальные стратегии равны $u_i^I = \bar{u}_i - \frac{2\pi}{\gamma(\rho + \delta)}$, а Парето-оптимальная кооперативная траектория имеет вид

$$x^I(t) = e^{-\delta t} x_0 + \frac{1}{\delta} \sum_{i \in I} u_i^I (1 - e^{-\delta t}).$$

Для того, чтобы вычислить $w_i(x^*(t), [t, +\infty))$, необходимо подставить соответствующие равновесные по Нэшу стратегии и начальное состояние на Парето-оптимальной кооперативной траектории в функцию выигрыша (4). Получаем

$$w_i(x^*(t), [t, +\infty)) = -\frac{\pi^2(8\rho + 3\delta)}{2\delta\rho\gamma(\rho + \delta)^2} + \frac{\pi e^{-\delta t} x_0}{\rho + \delta} + \frac{\pi \sum_{i \in I} \bar{u}_i}{\rho\delta} - \frac{\pi \sum_{i \in I} u_i^I e^{-\delta t}}{\delta(\rho + \delta)}.$$

Теперь по теореме 1 найдем динамически устойчивую ПРВ. Так как перед нами стоит задача минимизации суммарного выигрыша, то сначала вычислим неположительную функцию $\eta_i(t)$, $t \in [0, +\infty)$.

$$\lambda_i(x_0, [0, +\infty)) - w_i(x_0, [0, +\infty)) = -\frac{\pi^2}{2\rho\gamma(\rho + \delta)^2} = \int_0^{+\infty} e^{-\rho\tau} \eta_i(\tau) d\tau.$$

За неположительную функцию возьмем $\eta_i(t) = -\frac{\pi^2}{2\gamma(\rho + \delta)^2}$. Теперь подставим все полученные значения в формулу (2) и получим выражение для динамически устойчивой ПРВ

$$\beta_i(t) = -\frac{\pi^2(4\rho + 2\delta)}{\delta\gamma(\rho + \delta)^2} + \pi e^{-\delta t} x_0 + \frac{\pi}{\delta} \left(\sum_{i \in I} \bar{u}_i - \sum_{i \in I} u_i^I e^{-\delta t} \right). \quad (5)$$

Используя ПРВ (5), можно убедиться, что условие Д.В.К. Янга для задачи сокращения вредных выбросов выполнено.

Литература

1. Yeung D.W.K. An irrational-behavior-proofness condition in cooperative differential games // Intern. J. of Game Theory Rew., 2006. Vol. 8. P. 739–744.
2. Petrosjan L.A. The Time-Consistency problem in nonlinear dynamics // RBCM, J. of Brazilian Soc. of Mechanical Sciences, 1997. Vol. 19, № 2. P. 291–303.
3. Petrosyan L., Zaccour G. Time-consistent Shapley value allocation of pollution cost reduction // Journal of Economic Dynamics and Control, 2003. Vol. 27. P. 381–398.

Канторович А.В.

Санкт-Петербургский государственный университет

Теоретико-игровая модель взаимодействия и кооперации в цепи поставок

Рекомендовано к публикации доцентом Зенкевичем Н.А.

1. Введение. В данной работе исследована цепь поставок, которая организована следующим образом: фирма-производитель (производитель) производит некоторый продукт, который он затем продает оптом розничному торговцу (продавцу). Производитель выбирает цену, по которой он реализует товар продавцу, и объем поставки. Продавец определяет цену для конечного потребителя и величину затрат, связанных с маркетингом. Исследуемая модель цепи поставок является развитием модели, рассмотренной в [1]. Близкие задачи исследовались в работах [2–9]. Однако у рассматриваемой модели имеется ряд отличительных особенностей. Спрос в исследуемой модели не является постоянным, в отличие от [2–4]. В моделях, изученных в работах [5–7], спрос также не является постоянным, но там не учитывается взаимодействие производителя и продавца. Еще одна отличительная особенность модели состоит в том, что размер поставки определяет производитель (традиционно его устанавливает покупатель). Это характерно для областей промышленности, где производится узкоспециализированная продукция и цена производства высока (см. [8]). И наконец, в отличие от модели [9], исследованный вид функции спроса позволяет обобщить полученные результаты на случай нескольких продавцов.

2. Постановка задачи. Будем считать, что объем годового спроса зависит от цены $P \geq 0$, устанавливаемой продавцом для конечного потребителя, и расходов продавца на маркетинг $M \geq 0$ следующим образом: $D = k(h - P)^a M^b$, где a и b – коэффициенты эластичности спроса по цене и затратам продавца на маркетинг соответственно ($0 < b < 1$, $b + 1 < a$). Здесь h – максимальная цена, при которой конечный потребитель будет покупать товар, и $k > 0$ – параметр масштаба для функции спроса. В модели дефицит отсутствует, поэтому количество r произведенной за день продукции превышает дневной спрос d . Без ограничения общности будем считать, что $r = ud$, $u \geq 1$.

Данную модель можно рассматривать как игру двух лиц. Производитель (игрок 1) устанавливает цену $V \geq 0$ продукта для конечного потребителя, а также размер поставки $Q \geq 0$. Поэтому стратегия производителя представляет собой вектор (V, Q) . В качестве выигрыша производителя рассмотрим его чистую прибыль, т.е. доход от продаж за вычетом затрат. Затраты включают в себя расходы на производство, установку и хранение. Поэтому выигрыш производителя может быть записан в виде

$$\begin{aligned} \Pi_p(V, Q) &= DV - DC_p - DA_p Q^{-1} - 0,5iC_p Q u^{-1} = \\ &= k(h - P)^a M^b V - k(h - P)^a M^b C_p - \\ &\quad - k(h - P)^a M^b A_p Q^{-1} - 0,5iC_p Q u^{-1}. \end{aligned}$$

Здесь C_p – удельная цена производства (включая закупку материалов), A_p – цена работ по установке (обработке заказа) для производителя, i – доля затрат от C_p на хранение одной единицы в год. Поэтому средние расходы на хранение равны $0,5iC_p Q u^{-1}$.

Продавец (игрок 2) выбирает цену для конечного потребителя $P \geq 0$ и объем расходов на маркетинг $M \geq 0$. Поэтому стратегия продавца – вектор (P, M) . Выигрыш продавца также представляет его чистую прибыль, равную доходу от продаж за вычетом расходов на закупку товара и маркетинг, цены заказа и расходов на хранение, т.е.

$$\begin{aligned} \Pi_s(P, M) &= DP - DV - DM - DA_s Q^{-1} - 0,5iVQ = \\ &= k(h - P)^a M^b P - k(h - P)^a M^b V - k(h - P)^a M^{b+1} - \\ &\quad - k(h - P)^a M^b A_s Q^{-1} - 0,5iVQ. \end{aligned}$$

Здесь A_s – цена заказа для продавца, i – доля затрат от V на хранение одной единицы в год. Средние расходы на хранение равны $0,5iVQ$.

В итоге, построена бесконечная игра двух лиц в нормальной форме $\Gamma = \langle N, \{(V, Q)\}, \{(P, M)\}, \Pi_p, \Pi_s \rangle$, где $N = \{1, 2\}$ – множество игроков.

3. Некооперативный случай. Можно показать, что в игре Γ нет равновесия по Нэшу в чистых стратегиях. Поэтому рассмотрим игру Штакельберга. На первом шаге производитель (игрок 1, лидер) первым выбирает свою стратегию. На втором шаге продавец (игрок 2, последователь) выбирает свою стратегию с учетом выбора лидера.

Приравняв первые частные производные $\Pi_s(P, M)$ нулю, получим систему уравнений для нахождения стационарных точек функции выигрыша продавца. При $D > 0$ эта система имеет единственное решение $P^*(V, Q), M^*(V, Q)$, где

$$P^*(V, Q) = \frac{b(h+1) + a(V + A_s Q^{-1})}{a+b+1},$$

$$M^*(V, Q) = \frac{b(h-V - A_s Q^{-1})}{a+b+1}.$$

Поскольку второй дифференциал функции $\Pi_s(P, M)$ является отрицательно-определенной квадратичной формой, стратегия $(P^*(V, Q), M^*(V, Q))$ является наилучшим ответом на стратегию (V, Q) производителя.

Нахождение оптимальной стратегии производителя сводится к максимизации его функции выигрыша при условии, что продавец выбирает стратегию $(P^*(V, Q), M^*(V, Q))$, т.е. к решению задачи

$$\max_{V, Q} (k(h-P)^a M^b V - k(h-P)^a M^b C_p - k(h-P)^a M^b A_p Q^{-1} - 0,5iC_p Q u^{-1})$$

при условии, что

$$P = P^*(V, Q) = \frac{b(h+1) + a(V + A_s Q^{-1})}{a+b+1},$$

$$M = M^*(V, Q) = \frac{b(h-V - A_s Q^{-1})}{a+b+1}.$$

Данная задача имеет единственное решение. Однако найти его удастся только численно.

4. Кооперативный случай. Рассмотрим задачу максимизации суммарного выигрыша игроков

$$Z = \Pi_p + \Pi_s = DP - DM - DQ^{-1}(A_p + A_s) - 0,5iQ(V + C_p u^{-1}).$$

Поскольку последнее слагаемое всегда положительно, то максимум Z очевидно достигается при наименьшем V , удовлетворяющем предположениям задачи, т.е. $V = 0$. Условия первого порядка относительно

но P , M и Q приводят к соотношениям

$$Q = \sqrt{\frac{2k(h-P)^a M^b u(A_p + A_s)}{iC_p}},$$

$$P = \frac{(b+1)h + aC_p + (A_p + A_s)aQ^{-1}}{a+b+1},$$

$$M = \frac{b(h - C_p - (A_p + A_s)Q^{-1})}{a+b+1}.$$

Обозначим $Z^* = \max_{P,M,Q,V} Z$. Построим характеристическую функцию $v(s)$ следующим образом: $v(\emptyset) = v(\{1\}) = v(\{2\}) = 0$, $v(\{1, 2\}) = Z^*$. В этой игре вектор Шепли равен $sh = (0,5Z^*; 0,5Z^*)$. Далее размеры выигрышей игроков могут быть найдены численно.

5. Сравнение моделей. Сравнив решения в кооперативном и некооперативном случаях, видим, что соответствующие оптимальные цены и расходы на маркетинг связаны следующими соотношениями: $P_c^* = P_{nc}^* - \frac{a(V - C_p - A_p Q^{-1})}{a+b+1}$; $M_c^* = M_n^* + \frac{a(V - C_p - A_p Q^{-1})}{a+b+1}$. Здесь P_c^* и M_c^* соответствуют кооперативному случаю, а P_{nc}^* и M_{nc}^* – некооперативному. Разумно предположить, что выполнено неравенство $V > C_p + A_p Q^{-1}$, так как в противном случае продавец работает в убыток. Поэтому $P_c^* < P_{nc}^*$, а $M_c^* > M_{nc}^*$.

6. Численные примеры. Численные расчеты проводились для следующих значений параметров: $a = 1,7$; $b = 0,15$; $k = 1$; $u = 1,1$; $h = 40$; $i = 0,1$; $A_s = 40$; $A_p = 140$; $C_p = 1,5$. В таблицах 1–4 приведена зависимость решений от параметров a и b . Во всех примерах с ростом значений параметров растут и выигрыши игроков. Выгоднее для обоих игроков оказывается кооперативный вариант. Как в кооперативном, так и в некооперативном случаях оптимальная цена P^* и величина расходов на маркетинг M^* убывают с ростом a и возрастают с ростом b . Также ведет себя и оптимальная цена V^* в некооперативном случае (в случае кооперации $V^* = 0$). Зависимость оптимального размера поставки Q^* от b в кооперативной модели не прослеживается. Во всех остальных случаях Q^* возрастает с ростом параметров.

Таблица 1. Зависимость решения от параметра a . Некооперативный случай

a	1,7	1,9	2,1	2,3	2,5
P^*	25,3	24,0	22,8	21,8	20,9
M^*	1,3	1,3	1,2	1,2	1,1
V^*	15,2	14,2	13,4	12,7	12,1
Q^*	516,4	729,8	1031,9	1461,5	2072,3
D^*	100,1	201,9	403,5	809,6	1626,5
Π_p^*	482,9	1186,2	2606,7	5471,8	11170,0
Π_s^*	1319,9	2475,8	4677,1	8890,1	16989,3

Таблица 2. Зависимость решения от параметра a . Кооперативный случай

a	1,7	1,9	2,1	2,3	2,5
P^*	17,2	16,1	15,2	14,4	13,7
M^*	2,0	1,9	1,8	1,7	1,6
V^*	0	0	0	0	0
Q^*	748,9	920,1	1269,5	2160,2	3035,3
D^*	188,9	339,9	611,9	1713,0	2804,9
Π_p^*	1493,8	2831,5	5401,1	10360,7	19957,3
Π_s^*	1493,8	2831,5	5401,1	10360,7	19957,3

Таблица 3. Зависимость решения от параметра b . Некооперативный случай

b	0,15	0,2	0,25	0,3	0,35
P^*	25,3	25,3	25,4	25,6	25,7
M^*	1,3	1,7	2,1	2,5	2,9
V^*	15,2	14,9	14,7	14,5	14,3
Q^*	516,4	531,9	549,7	570,9	595,3
D^*	100,1	107,9	114,5	123,4	134,1
Π_p^*	482,9	529,0	575,0	632,2	701,3
Π_s^*	1319,9	1372,4	1444,5	1535,1	1644,3

Таблица 4. Зависимость решения от параметра b . Кооперативный случай

b	0,15	0,2	0,25	0,3	0,35
P^*	17,2	17,6	17,9	18,3	18,7
M^*	2,0	2,6	3,2	3,8	4,4
V^*	0	0	0	0	0
Q^*	748,9	629,6	805,2	189,3	833,6
D^*	188,9	243,1	220,2	242,8	267,8
Π_p^*	1493,8	1557,7	1647,0	1756,2	1886,9
Π_s^*	1493,8	1557,7	1647,0	1756,2	1886,9

Литература

1. Yang Shan-Lin, Zhou Yong Wu. Two-echelon supply chain models: Considering duopolistic retailers' different competitive behaviors // International Journal of Production Economics, 2006. Vol. 103. P. 104–116.
2. Chan Chi Kin, Kingsman Brian G. Coordination in a single-vendor multibuyer supply chain by synchronizing delivery and production cycles // Transportation Research, 2007. Part E, Vol. 43. P. 90–111.
3. Heuvel Wilco van den, Borm Peter, Hamers Herbert. Economic lot-sizing games // European Journal of Operational Research, 2007. Vol. 176. P. 1117–1130.
4. Dai Tinglong, Qi Xiangtong. An acquisition policy for a multi-supplier system with a finite-time horizon // Computers and Operations Research, 2007. Vol. 34. P. 2758–2773.
5. Jung Hoon, Cerry M. Klein. Optimal inventory policies under decreasing cost functions via geometric programming // European Journal of Operational Research, 2001. Vol. 132. P. 628–642.
6. Jung Hoon, Cerry M. Klein. Optimal inventory policies for an economic order quantity model with decreasing cost functions // European Journal of Operational Research, 2005. Vol. 165, № 1. P. 108–126.
7. Sajadi Seyed J., Orouge Maryam, Aryanezhad Mir B. Optimal production and marketing planning // Computational Optimization and Applications, 2005. Vol. 30, № 2. P. 195–203.
8. Kelle Peter, Al-khateeb Faisal, Miller Pam Anders. Partnership and negotiation support by joint optimal ordering-setup policies for JIT // International Journal of Production Economics, 2003. Vol. 81, 82. P. 431–441.
9. Esmaili M., Aryanezhad Mir B., Zeepongsekul P. A game theory approach in seller-buyer supply chain // European Journal of Operational Research, 2009. Vol. 195. P. 442–448.

Карпов М.И., Петросян Л.А.
Санкт-Петербургский государственный университет

Минимизация затрат в транспортной сети

Введение. В работе рассматривается транспортная сеть с несколькими игроками. Каждому игроку соответствует пара вершин – начальная и конечная вершины. Необходимо найти траекторию перехода из начальных вершин в конечные, минимизирующую суммарные затраты на переход. Если игроки действуют по-отдельности, то оптимальный путь из начальной вершины в конечную для каждого из игроков можно найти с помощью известных алгоритмов нахождения кратчайшего пути между двумя вершинами, например, с помощью алгоритма Дейкстры или алгоритма Флойда [1]. Суммарные затраты при этом будут равны сумме затрат на переход по каждому из путей.

Введём дополнительное условие. Если какое-либо ребро входит в пути нескольких игроков, то плата за него взимается не с каждого игрока, а один раз, вне зависимости от того, сколько игроков проходят по этому ребру. В таком случае суммарные затраты на переход из начальных вершин в конечные будут равны сумме затрат на переход по каждому из ребёр, присутствующих в траектории.

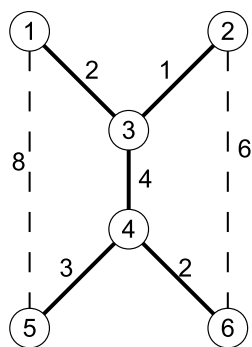


Рис. 1. Пример сети

На рис. 1 изображена сеть, состоящая из шести вершин. Участвуют два игрока. Начальная и конечная вершины для первого игрока – это вершины с номерами 1 и 5 соответственно, для второго игрока – вершины с номерами 2 и 6. Пунктиром выделена оптимальная траектория перехода из начальных вершин в конечные в случае, если игроки действуют по-отдельности. Жирной линией выделена оптимальная траектория перехода из начальных вершин в конечные в случае, если игроки действуют сообща. При этом в первом случае суммарные затраты равны 14, во втором случае – 12.

Данная работа посвящена решению задачи поиска оптимальной траектории, минимизирующей суммарные затраты в случае, когда игроки объединяют свои усилия. Найден вид уравнения, позволяющего найти минимальные суммарные затраты. Представлен метод последовательных приближений, позволяющий решить это уравнение и восстановить оптимальную траекторию, обеспечивающую минимальные суммарные затраты. Метод основан на принципе оптимальности Беллмана.

1. Постановка задачи. Рассмотрим сеть

$$G(X, D, C), \quad (1)$$

где $X = \{1, \dots, x, \dots, n\}$ – множество вершин, пронумерованных от 1 до n , D – множество ребер, $C = \{c_{xy}\}_{n \times n}$, $x, y \in X$ – матрица затрат. Если для некоторой пары вершин $x_1, x_2 \in G$ не существует ребро (x_1, x_2) , дополним сеть G ребром (x_1, x_2) , имеющим вес $c_{x_1 x_2} = \sum_{(x,y) \in D} c_{xy} + 1$. Зададим множество игроков

$$L = \{1, \dots, i, \dots, l\}.$$

Определение. Состоянием z сети (1) будем называть набор объектов вида $z = (M_1, \dots, M_x, \dots, M_n)$, где $M_x \subset L$ – подмножество игроков, находящихся в вершине $x \in X$ в состоянии z , такое, что $\bigcup_x M_x = L$; для любых $x, y \in X : x \neq y$, пересечение $M_x \cap M_y = \emptyset$.

Множество всевозможных состояний сети обозначим через Ω : $\Omega = \{z\}$.

Определение. Местоположением $z^{(i)}$ игрока $i \in L$, соответствующим состоянию $z = (M_1, \dots, M_x, \dots, M_n)$, назовем вершину $x \in X$ такую, что $i \in M_x$.

Замечание. В некоторых случаях состояние системы удобно представлять в виде вектора $z = (z^{(1)}, \dots, z^{(i)}, \dots, z^{(l)})$, где $z^{(i)}$ – местоположение i -го игрока в состоянии z .

Определение. Траекторией перехода из некоторого состояния z_1 в некоторое состояние z_m будем называть некоторую последовательность состояний $z_1, \dots, z_k, \dots, z_m$ и обозначать $P(z_1, z_m) = (z_1, \dots, z_k, \dots, z_m)$.

Всякая траектория $P(z_1, z_m) = (z_1, \dots, z_k, \dots, z_m)$ задает l последовательностей местоположений игроков: $z_1^{(i)}, \dots, z_k^{(i)}, \dots, z_m^{(i)}$, $i = \overline{1, l}$.

Определение. Путем перехода игрока $i \in L$ из местоположения $z_1^{(i)}$ в местоположение $z_m^{(i)}$, соответствующим траектории

$P(z_1, z_m) = (z_1, \dots, z_k, \dots, z_m)$, будем называть последовательность местоположений $z_1^{(i)}, \dots, z_k^{(i)}, \dots, z_m^{(i)}$ этого игрока в сети и обозначать $p^{(i)}(z_1^{(i)}, z_m^{(i)}) = (z_1^{(i)}, \dots, z_k^{(i)}, \dots, z_m^{(i)})$.

Рассмотрим путь $\bar{p}^{(i)}(z_1^{(i)}, z_m^{(i)}) = (z_1^{(i)}, \dots, z_k^{(i)}, \dots, z_m^{(i)})$, соответствующий некоторой траектории $\bar{P}(z_1, z_m)$. Всякое местоположение $z_k^{(i)} \in \bar{p}^{(i)}(z_1^{(i)}, z_m^{(i)})$, $k = \overline{1, m}$, – номер вершины, в которой находится игрок i в состоянии z_k . Следовательно, каждая пара $(z_k^{(i)}, z_{k+1}^{(i)})$, $k = \overline{1, m-1}$, соседних местоположений из последовательности $(z_1^{(i)}, z_m^{(i)}) = (z_1^{(i)}, \dots, z_k^{(i)}, \dots, z_m^{(i)})$ задает ребро $(z_k^{(i)}, z_{k+1}^{(i)}) \in D$, $k = \overline{1, m-1}$ сети (1), то есть путь $\bar{p}^{(i)}(z_1^{(i)}, z_m^{(i)})$ состоит из $m-1$ ребра $(z_k^{(i)}, z_{k+1}^{(i)})$, $k = \overline{1, m-1}$. Обозначим через $E(\bar{P}(z_1, z_m))$ множество всех ребер, которые составляют пути $\bar{p}^{(i)}(z_1^{(i)}, z_m^{(i)})$, $i = \overline{1, l}$, соответствующие траектории $\bar{P}(z_1, z_m)$.

Определение. Суммарные затраты на переход из состояния z_1 в состояние z_m по траектории $\bar{P}(z_1, z_m)$ определяются по формуле

$$S(z_1, z_m) = \sum_{(x,y) \in E(\bar{P}(z_1, z_m))} c_{xy}.$$

Определение. Оптимальной траекторией перехода из состояния z_1 в состояние z_m будем называть траекторию $P^*(z_1, z_m)$, обеспечивающую минимальные суммарные затраты на переход из состояния z_1 в состояние z_m .

Сформулируем теперь задачу в терминах введенных нами определений.

Пусть задана сеть $G(X, D, C)$, $X = \{1, \dots, x, \dots, n\}$; множество игроков $L = \{1, \dots, i, \dots, l\}$; начальное состояние $z_0 \in \Omega$ и конечное состояние $z_N \in \Omega$. Необходимо найти оптимальную траекторию перехода из состояния z_0 в состояние z_N .

2. Решение задачи. Задача решается с помощью метода динамического программирования. В основе этого метода лежит принцип оптимальности, сформулированный Р. Беллманом [2].

2.1. Принцип оптимальности. Оптимальное поведение обладает тем свойством, что, каковы бы ни были первоначальное состояние и решение в начальный момент, последующие решения должны составлять оптимальное поведение относительно состояния, получающегося в результате первого решения.

Переформулируем этот принцип в терминах рассматриваемой задачи следующим образом.

Утверждение 1. *Оптимальная траектория $P^*(z_0, z_N)$ перехода из состояния z_0 в состояние z_N обладает тем свойством, что, какой бы ни была траектория перехода из состояния z_0 в некоторое состояние $z_k \in P^*(z_0, z_N)$, оставшаяся часть этой траектории, начинающаяся из состояния z_k , будет оптимальной траекторией перехода из состояния z_k в состояние z_N .*

Введём в рассмотрение функцию Беллмана $W(z)$. Значение этой функции в любой точке $z \in \Omega$ равно минимальным суммарным затратам при переходе из состояния z в конечное состояние z_N . Тогда утверждение 1 можно записать в виде функционального уравнения Беллмана

$$W(z) = \min_{z' \in \Omega} (F(z, z') + W(z')) \quad (2)$$

с граничным условием

$$W(z_N) = 0. \quad (3)$$

Здесь

$$F(z, z') = \sum_{\substack{M_x \in z, M_y \in z': \\ M_x \cap M_y \neq \emptyset}} c_{xy}$$

это затраты на переход из состояния z в состояние z' .

2.2. Метод последовательных приближений. Введем функцию $W_k(z)$, значение которой в каждой точке $z \in \Omega$ равно минимальным суммарным затратам при переходе из состояния z в состояние z_N при дополнительном ограничении, что оптимальная траектория $P_k^*(z, z_N)$, обеспечивающая это минимальное значение суммарных затрат, содержит не более k промежуточных состояний, $k = 0, 1, \dots$

Тогда получаем, что

$$\begin{aligned} W_0(z) &= F(z, z_N), \\ W_k(z) &= \min_{z' \in \Omega} (F(z, z') + W_{k-1}(z')), \quad z \in \Omega, \quad k = 1, 2, \dots \end{aligned} \quad (4)$$

Ясно, что с увеличением числа k шагов значения функций $W_k(z)$ могут разве только уменьшаться. Пусть R – количество всевозможных состояний сети G , $|\Omega| = R$. Тогда при $k \geq R$ последовательность функций $W_k(z)$ сходится: $W_R(z) = W_{R+1}(z) = W(z)$. Таким образом, значение $W_R(z_0)$ есть значение минимальных суммарных затрат $W(z_0)$ при переходе из начального состояния z_0 в конечное состояние z_N по оптимальной траектории $P^*(z_0, z_N) = (z_1^*, \dots, z_R^*)$. Для нахождения значений z_1^*, \dots, z_R^* необходимо рассмотреть процесс (4) в обратном порядке, начиная с шага $k = R$. Таким образом, задача решена.

2.3. Блок-схема алгоритма. На рис. 2 приведена блок-схема алгоритма поиска оптимальной траектории. Данная программа реализована на языке C++, сделаны вычисления для сети, состоящей из 150 вершин и 4 игроков.

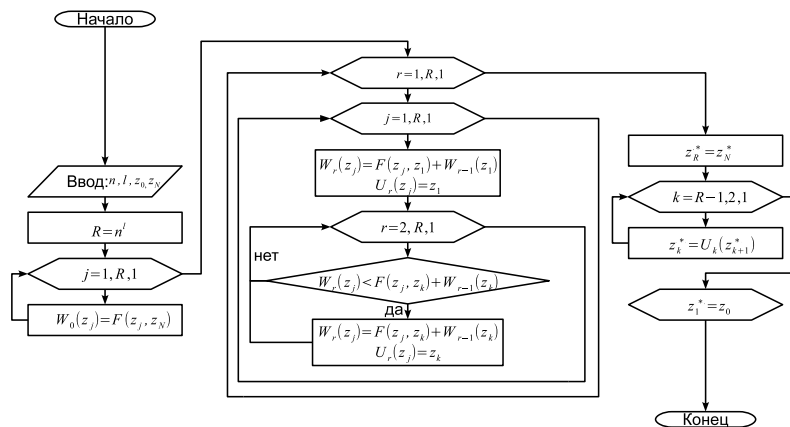


Рис. 2. Блок-схема алгоритма поиска оптимальной траектории

3. Пример. На рис. 3 приведен пример сети, состоящей из 17 вершин, в которой участвуют 3 игрока. Начальное состояние сети – $z_0 = (1, 2, 3)$, конечное состояние сети – $z_N = (15, 16, 17)$. Пунктиром выделена оптимальная траектория перехода из начальных вершин в конечные в случае, если игроки действуют по-отдельности. Жирной линией выделена оптимальная траектория перехода из начальных вершин в конечные в случае, если игроки действуют сообща. Суммарные затраты в первом случае равны 39, во втором случае – 31.

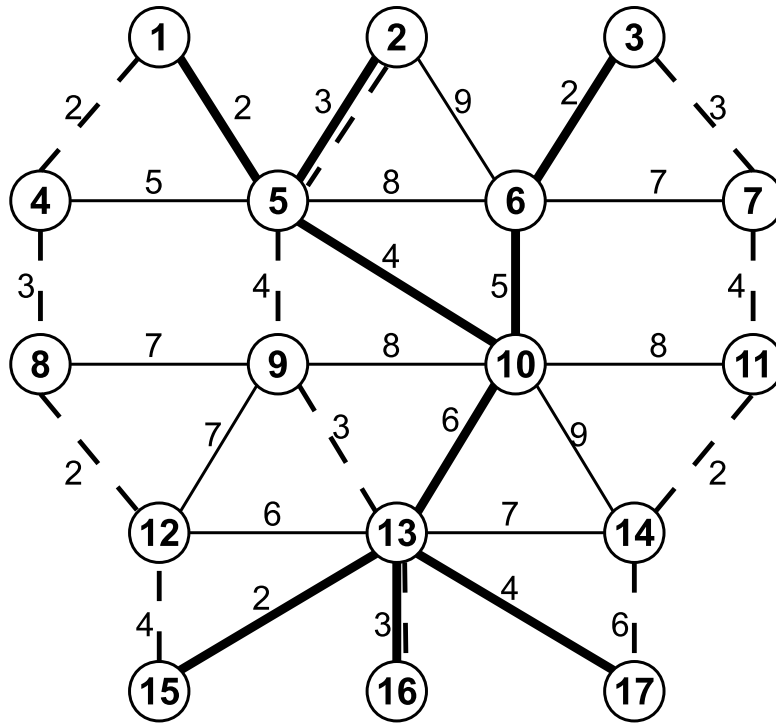


Рис. 3. Пример работы алгоритма

4. Заключение. Представленный в работе алгоритм поиска оптимальной траектории перехода из начальных вершин в конечные позволяет минимизировать суммарные затраты на переход. Вопрос разделения затрат между игроками представляет тему дальнейшего исследования.

Литература

1. Кристофидес Н. Теория графов. Алгоритмический подход. М.: Мир, 1978. 432 с.
2. Беллман Р. Динамическое программирование. М.: ИЛ, 1960. 400 с.

Кумачева С.Ш., Чугунникова Н.А.
Санкт-Петербургский государственный университет

Оценка и прогнозирование валютного риска

Неблагоприятное изменение курсов иностранных валют по отношению к валюте баланса компании может привести к серьезным потерям. Поэтому разработка методологии управления валютным риском предельно актуальна. Валютный риск [1] складывается из валютных позиций – остатков средств в иностранных валютах, которые формируют активы и пассивы в соответствующих валютах. Валютная позиция становится открытой, когда соотношение требований и обязательств нарушается.

Использование мер риска позволяет получить оценку риска предприятия в соответствии с его доходами или по сравнению с другими предприятиями. С их помощью возможно оценить максимальный убыток и сформировать политику компании по регулированию рисков.

В данной работе оценка риска проводилась с помощью следующих мер: экономический индекс риска R^e [2], операторное измерение риска R^o [3], коэффициент Шарпа Sha (мера дохода при заданном риске).

Математическая постановка задачи. Пусть X – портфель открытых валютных позиций (ОВП). x_1, \dots, x_n – ОВП (количество единиц отдельной иностранной валюты в портфеле), k_1, \dots, k_n – курсы в рублях соответствующих иностранных валют. Тогда значение портфеля будет определяться по формуле

$$X = x_1 k_1 + \dots + x_n k_n.$$

Валютный риск будет равен объему данного портфеля X .

Цель данной работы заключается в нахождении оценки и прогноза валютного риска, связанного с возможностью получения предприятием убытков.

Для решения поставленной задачи необходимо задать уравнение прибыли. Пусть X_t – значение портфеля в момент времени t . Тогда соответственно $x_{1,t}, \dots, x_{n,t}$ – ОВП, $k_{1,t}, \dots, k_{n,t}$ – курсы в рублях соответствующих иностранных валют в момент времени t . Прибыль

по открытой позиции валюты i за день определяется по формуле

$$W_{i,t} = x_{i,t}(k_{i,t} - k_{i,t-1}).$$

Тогда прибыль по всем ОВП равна

$$W_t = \sum_{i=1}^n W_{i,t} = \sum_{i=1}^n x_{i,t}(k_{i,t} - k_{i,t-1}). \quad (1)$$

Таким образом на изменение прибыли по каждой отдельной валюте влияет величина ОВП и изменение курса за день. Зависимость прибыли от изменения курсов валют составляет экономический риск.

Рассмотрим другой тип валютного риска – операционный. Операционный риск связан с совершенными собственными операциями по покупке и продаже иностранной валюты.

Для определения прибыли от совершения собственных операций необходимо вычислить обменный курс валют

$$k_{ij,t} = k_{i,t}/k_{j,t},$$

где i – валюта, в которой совершается сделка, j – валюта, в которую переводится валюта i .

Пусть $s_{1,t}, \dots, s_{m,t}$ – суммы сделок в валюте i , $k_{s_{1,t}}, \dots, k_{s_{m,t}}$ – обменные курсы, по которым совершаются сделки. Тогда прибыль от совершения сделки при покупке валюты можно представить в виде

$$W_{s_{i_1,t}} = s_{i_1,t}(k_{ij,t} - k_{s_{i_1,t}})k_{j,t}.$$

Аналогично прибыль от совершения сделки при продаже валюты можно выразить следующим образом

$$W_{s_{i_1,t}} = s_{i_1,t}(k_{s_{i_1,t}} - k_{ij,t})k_{j,t}.$$

Прибыль по всем заключенным сделкам

$$W_{st} = \sum_{i_1=1}^m W_{s_{i_1,t}}. \quad (2)$$

Таким образом операционный риск зависит от суммы сделок и от обменных курсов валют.

Меры риска определены на области предприятий g – случайных величин с положительным математическим ожиданием. В данной работе в качестве предприятия рассматривается прибыль, определенная в уравнениях (1), (2). Для нахождения величины экономического и операторного индексов риска необходимо решить следующие уравнения:

$$Ee^{-g/R^e(g)} = 1, \quad (3)$$

$$E\left(\ln\left(1 + \frac{1}{R^o(g)}g\right)\right) = 0. \quad (4)$$

Коэффициент Шарпа определяется следующим образом:

$$Sha(g) = E(g)/\sigma(g), \quad (5)$$

где $E(g)$ – математическое ожидание g , $\sigma(g)$ – стандартное отклонение.

Применение мер риска к получению оценки. С помощью рассмотренных выше мер риска оценим экономический и операционный риски. По данным некоторой компании за каждый день в течение полугода (май 2010 – ноябрь 2010) была посчитана прибыль от совершения операций с валютой, что составляет операционный риск, и от переоценки ОВП – экономический риск. Данные представлены на рис. 1

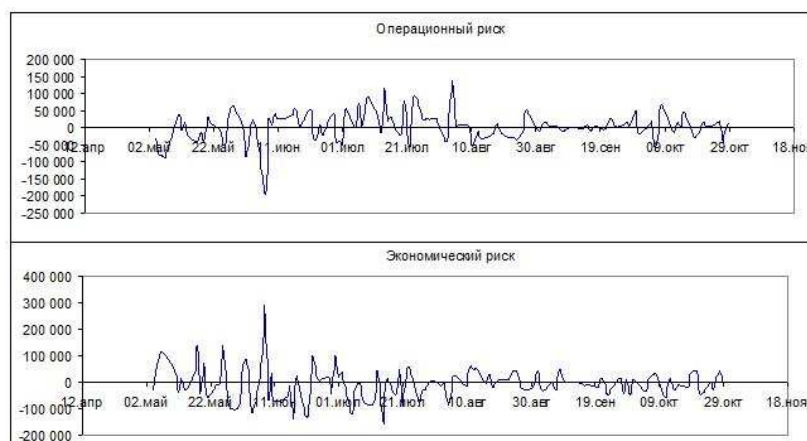


Рис. 1. Значения риска

Основные валюты, по которым возникают наибольшие прибыли или убыток, – это доллар и евро. По данным за полгода о размере ОВП, о соответствующих курсах валюты за каждый день была рассчитана прибыль от переоценки валютных позиций.

С помощью критерия серий было проведено исследование, результатом которого явилось принятие гипотезы о случайности использованной выборки для 5% уровня значимости. Прибыль от совершения операций с валютой и от переоценки ОВП исследовалась с помощью указанных выше мер риска. Для нахождения экономического и операторного индексов риска в среде MATLAB были решены соответственно уравнения (3) и (4). Коэффициент Шарпа был вычислен по формуле (5). Полученные результаты приведены в таблице 1.

Таблица 1. Оценки мер риска

	Экономический риск	Операционный риск	Переоценка EUR	Переоценка USD
R^e	466270	72257	14630	79457
R^o	473370	101350	19082	218830
Sha	0,0487	0,0874	0,4940	0,1792

Прогнозирование валютного риска с использованием временных рядов. С целью получения прогноза валютного риска был проведен анализ данных с помощью временных рядов. При исследовании значений валютного риска был выявлен логарифмический тренд. Также была рассмотрена сезонная компонента при различных периодах. Наилучший результат, обусловленный наименьшей дисперсией, был получен при периоде, равном кварталу. Был вычислен прогноз значений валютного риска через месяц и квартал. Результаты приведены на рис. 2.

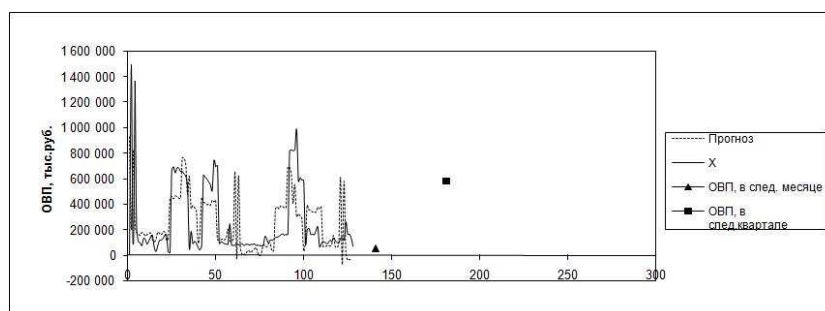


Рис. 2. Прогноз значений

Аналогично были исследованы значения валютного риска в части отдельных валют. Результаты приведены в таблице 2.

Таблица 2. Прогноз

	Валютный риск	ОВП в EUR	ОВП в USD
Прогноз через месяц	53692	299172	2251471
Прогноз через квартал	580025	384737	1202314

При оценке экономического и операционного риска с помощью рассматриваемых в работе мер был получен один и тот же результат: значение оценки экономического риска больше, чем операционного. Аналогично можно сделать вывод о том, что риск получить убыток при переоценке долларов больше, чем при переоценке евро. Следовательно, компании необходимо контролировать экономический риск, предпринимая меры по его снижению, в особенности, необходимо закрывать позиции в долларах.

Литература

1. Лобанов А.А., Чугунов А.В. Энциклопедия финансового риск-менеджмента. М.: Альпина Паблишер, 2003. 786 с.
2. Aumann R.J., Serrano R. An economic index of riskiness // Journal of Political Economy, 2008. Vol. 116, № 5. P. 810–836.
3. Foster D.P., Hart S. An operational measure of riskiness// Journal of Political Economy, 2009. Vol. 117, № 5. P. 785–814.

Мазалова А.В.

Санкт-Петербургский государственный университет

Парадокс Браесса

Рекомендовано к публикации профессором Буре В.М.

Парадокс Браесса. Эта модель была предложена Браессом в 1968 году. Рассмотрим дорожную сеть (рис. 1). Пусть 60 автомобилей движутся из пункта A в пункт B . Задержка на сегментах (C, B) и (A, D) не зависит от числа автомобилей и равна 1 часу, а на сегментах (A, C) и (D, B) она пропорциональна числу движущихся автомобилей. Нетрудно видеть, что равновесием здесь будет одинаковое распределение автомобилей по обеим веткам, т.е. по 30 автомобилей на каждой ветке. При этом каждый из автомобилей затратит на дорогу 1,5 часа.

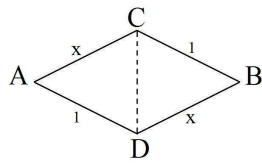


Рис. 1. Парадокс Браесса

Соединим пункты C и D скоростной дорогой, двигаясь по которой каждый автомобиль имеет 0 минут задержку. Тогда тем автомобилям, которые ранее ехали по пути (A, D, B) , выгоднее будет ехать по пути (A, C, D, B) . Это же касается автомобилей, ехавших по пути (A, C, B) . Таким образом, равновесием по Нэшу будет ситуация, в которой все автомобили едут по пути (A, C, D, B) , но при этом каждый из автомобилей затратит на дорогу 2 часа. Получилась парадоксальная ситуация: построили новую скоростную дорогу, а затраты каждого из участников увеличились [1].

Парадокс Браесса на квадрате. На единичном квадрате располагаются две фирмы в точках $(0, \frac{1}{2})$ и $(1, \frac{1}{2})$. Каждая из них задает свою цену на один и тот же товар. Пусть цены будут c_1 и c_2 соответственно. Будем считать, что цены приблизительно равны, но $c_1 \leq c_2$. По всем сторонам квадрата равномерно располагаются покупатели (внутри квадрата покупателей нет). Покупатель из каждой точки сравнивает затраты от посещения каждой из фирм, причем затраты складываются из цены на товар плюс транспортные расходы [2].

Предположим, что построена скоростная дорога, проходящая через центр квадрата, при движении по которой затраты покупателей равны нулю.

Рассмотрим случай без дороги. Пусть x характеризует границу, которая разделяет покупателей на два множества тех: кто идет в первую фирму, и тех, кто идет во вторую (рис. 2).

$$c_1 \leq c_2,$$

$$c_1 + 1 + x = c_2 + 1 - x, \quad x = \frac{c_2 - c_1}{2}.$$

Найдем функцию выигрыша для игрока I и II:

$$H_1(c_1, c_2) = c_1(2 + 2x) = c_1(2 + c_2 - c_1),$$

$$H_2(c_1, c_2) = c_1(2 - 2x) = c_2(2 - c_2 + c_1),$$

$$\frac{\partial H_1}{\partial c_1^o} = -2c_1^o + c_2^o + 2, \quad \frac{\partial H_2}{\partial c_2^o} = c_1^o - 2c_2^o + 2.$$

Равновесие по Нэшу найдем из условий

$$\frac{\partial H_1}{\partial c_1^o} = 0, \quad \frac{\partial H_2}{\partial c_2^o} = 0.$$

Отсюда следует, что $c_1^o = c_2^o = 2$.

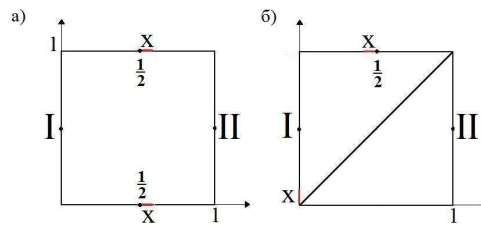


Рис. 2. Случай без дороги. Случай с дорогой

А теперь рассмотрим случай, связанный с построением дороги. Пусть дорога имеет вид, изображенный на Рис. 2. Здесь x характеризует границу, которая разделяет покупателей на тех, кому выгодно идти в I, и тех, кому выгодно идти в II.

1. Если $c_1 > c_2$, то

$$c_1 + \frac{1}{2} - x = c_2 + \frac{1}{2} + x, \quad x = \frac{c_1 - c_2}{2}.$$

Найдем функцию выигрыша для игрока I:

$$H_1(c_1, c_2) = c_1 \left(\frac{3}{2} - 2x \right) = c_1 \left(\frac{3}{2} - c_1 + c_2 \right).$$

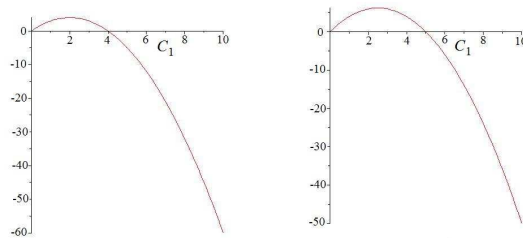


Рис. 3. Функция выигрыша фирмы I при $c_2^* = \frac{5}{2}$

2. Если $c_2 > c_1$, то функция выигрыша игрока I имеет вид

$$H_1(c_1, c_2) = c_1 \left(\frac{5}{2} + 2x \right) = c_1 \left(\frac{5}{2} - c_1 + c_2 \right).$$

3. Если $c_1 = c_2$, то $H_1(c_1, c_2) = 2c_1$.

Рассмотрим ситуацию $(c_1, c_2) = \left(\frac{5}{2}, \frac{5}{2}\right)$. По графику видно, что если фирма II выбирает стратегию $c_2^* = \frac{5}{2}$, то наилучший ответ фирмы I на эту стратегию будет $c_1^* = \frac{5}{2}$. Поэтому ситуация равновесия в данной игре это ситуация $(c_1^*, c_2^*) = \left(\frac{5}{2}, \frac{5}{2}\right)$.

Получилась парадоксальная ситуация: построили новую скоростную дорогу, а затраты каждого из покупателей увеличились.

Литература

1. Мазалов В.В. Математическая теория игр и приложения. М.: Лань, 2010. 345 с.
2. Мазалова А.В. Дуополия Хотеллинга на окружности // Процессы управления и устойчивость: Труды 40-й международной научной конференции аспирантов и студентов / Под ред. Н. В. Смирнова, Г. Ш. Тамасяна. СПб.: Издат. Дом С.-Петерб. гос. ун-та, 2009. С. 643–598.

Мельникова Е.А.

Санкт-Петербургский государственный университет

Проблема очереди в системе обслуживания: теоретико-игровая модель

Рекомендовано к публикации профессором Петросяном Л.А.

Введение. Представим стихийное бедствие в городе. Пострадало много домов, фирм. . . Все эти фирмы – агенты – одновременно обращаются в некоторую ремонтную службу. Все фирмы в итоге должны быть обслужены, причем как можно скорее, поскольку каждая минута ожидания оборачивается издержками. Так как фирмы одновременно обслуживаться не могут, их следует выстроить в очередь, а также установить денежные компенсации для тех, кто вынужден ожидать. Обслуживание каждой фирмы занимает единицу времени.

Цель работы – найти решение этой так называемой «проблемы очереди», удовлетворяющее некоторым аксиомам. Первая из них – Парето-оптимальность. В условиях задачи можно разделить эту аксиому на две: оптимальность очереди (т.е. минимальность суммарных ожидаемых издержек всех агентов) и сбалансированность (сумма всех компенсационных выплат равна 0). Вторая аксиома – аксиома справедливости – требует, чтобы агенты с равными издержками считались равноправными.

Рассмотрим задачу с точки зрения кооперативной игры. Для этого необходимо определить характеристическую функцию. Существует множество способов её введения [1]. В данной работе значение характеристической функции определяется как максимальный гарантированный суммарный выигрыш коалиции, равный сумме компенсационных платежей всем членам коалиции за вычетом издержек, при том условии, что все члены коалиции расположены в конце очереди. Поскольку удастся доказать супераддитивность такой характеристической функции, то имеет смысл использовать в качестве принципа оптимальности вектор Шепли.

Модель. Пусть N – конечное множество агентов, и $n = |N|$. Каждому агенту $i \in N$ ставятся в соответствие его позиция $\sigma_i \in \mathbb{N}_+$ (где \mathbb{N}_+ – множество строго положительных натуральных чисел) в очереди и положительный или отрицательный платеж $t_i \in \mathbb{R}$ (здесь

\mathbb{R} – множество вещественных чисел). Пусть $c_i \in \mathbb{R}_+$ (где \mathbb{R}_+ – множество неотрицательных вещественных чисел) – ожидаемые издержки агента $i \in N$. Если агент $i \in N$ получает, скажем, σ_i -ое место в очереди, то его суммарные ожидаемые издержки составляют $c_i(\sigma_i - 1)$. Агент, находящийся в начале очереди, не имеет издержек. Функция выигрыша для $i \in N$ выглядит следующим образом [2]:

$$u_i(\sigma_i, t_i, c_i) = t_i - (\sigma_i - 1)c_i.$$

Определение 1. *Проблема очереди* описывается набором $q = (N, C)$, $C \in \mathbb{R}_+^n$, где $C = (c_1, \dots, c_n)$, \mathbb{R}_+ – множество неотрицательных вещественных чисел.

Определение 2. *Распределение* – это набор $z = (\sigma, t) = (\sigma_i, t_i)_{i \in N} \in (\mathbb{N}_+ \times \mathbb{R})^n$, где для $\forall i \in N$ место i -го агента в очереди обозначено σ_i , а t_i – его компенсационный платеж.

При заданных $\sigma = (\sigma_1, \dots, \sigma_n) \in \mathbb{N}_+^n$ и $i \in N$ определим множества агентов, предшествующих и стоящих после i -го агента, через $P_i(\sigma)$ и $F_i(\sigma)$, соответственно.

Определение 3. Распределение z *допустимо* для $q = (N, C)$ тогда и только тогда, когда не существует пары агентов с совпадающей позицией в очереди, и сумма всех платежей неположительна, т.е. множество допустимых распределений имеет вид

$$\Sigma(q) = \left\{ z = (\sigma, t) \in (\mathbb{N}_+ \times \mathbb{R})^n \mid \begin{array}{l} \sum_{i \in N} t_i \leq 0, \\ \forall i, j \in N : i \neq j \Rightarrow \sigma_i \neq \sigma_j \end{array} \right\}.$$

Определение 4. Распределение z *оптимально* для $q = (N, C)$, если оно доставляет минимум суммарным ожидаемым издержкам всех агентов и платежи в сумме составляют 0, т.е. когда для всех $z' = (\sigma', t') \in \Sigma(q)$ справедливы $\sum_{i \in N} (\sigma_i - 1)c_i \leq \sum_{i \in N} (\sigma'_i - 1)c_i$ и $\sum_{i \in N} t_i = 0$.

Пусть очередь, минимизирующая суммарные ожидаемые издержки всех агентов, есть $\sigma^* = (\sigma_1^*, \dots, \sigma_n^*)$. Позиции агентов в такой очереди σ^* зависят от ожидаемых издержек самих агентов: чем выше ожидаемые издержки за единицу времени, тем меньше номер положения агента в очереди. Очередь σ^* единственна с точностью до перестановки агентов с одинаковыми ожидаемыми издержками, т.е. в очереди такие агенты стоят друг за другом и их можно менять местами.

Множество оптимальных очередей для $q = (N, C)$ обозначается $\sigma^*(q)$. Суммарные ожидаемые издержки равны для всех оптимальных очередей.

Определение 5. *Правило распределения* φ ставит в соответствие каждой проблеме $q = (N, C)$ непустое подмножество распределений $\varphi(q)$.

Организация очереди. Один из подходов решения проблемы очереди $q = (N, C)$ заключается в том, чтобы рассмотреть ее как кооперативную игру. Определим для нее характеристическую функцию.

Рассмотрим коалицию $S \subseteq N$ и соответствующую ей редуцированную задачу $q' = (S, C')$, где C' – вектор издержек агентов коалиции S . Пусть $\sigma' \in \sigma^*(q')$ – оптимальная очередь для новой задачи. Теперь вернемся к проблеме очереди q и переставим коалицию S в конец очереди. Тогда S будет стоять на местах $n - |S| + 1, \dots, n$. Поэтому суммарные издержки всех агентов коалиции S будут равны $\sum_{i \in S} (n - |S| + \sigma'_i - 1)c_i$. Иными словами, для всех $q = (N, C)$, $\forall S \subseteq N$

$$v_q(S) = - \sum_{i \in S} (\sigma'_i - 1)c_i - (n - |S|) \sum_{i \in S} c_i.$$

Утверждение. *Функция (1) является супераддитивной.*

Доказательство. Покажем супераддитивность функции $v_q(S)$, т.е. для $\forall S, T \subset N : S \cap T = \emptyset$ имеет место

$$v_q(T) + v_q(S) \leq v_q(T \cup S).$$

Рассмотрим

$$\begin{aligned} v_q(T \cup S) - (v_q(T) + v_q(S)) &= - \sum_{i \in S \cup T} (\sigma_i^1 - 1)c_i - \\ &- (n - |S| - |T|) \sum_{i \in S \cup T} c_i + \sum_{i \in S} (\sigma_i^2 - 1)c_i + (n - |S|) \sum_{i \in S} c_i + \\ &+ \sum_{i \in T} (\sigma_i^3 - 1)c_i + (n - |T|) \sum_{i \in T} c_i = - \sum_{i \in S \cup T} (\sigma_i^1 - 1)c_i - \\ &- (n - |S|) \sum_{i \in S} c_i - (n - |T|) \sum_{i \in T} c_i + \sum_{i \in S} (\sigma_i^2 - 1 + |T|)c_i + \\ &+ (n - |S|) \sum_{i \in S} c_i + \sum_{i \in T} (\sigma_i^3 - 1 + |S|)c_i + (n - |T|) \sum_{i \in T} c_i = \end{aligned}$$

$$= - \sum_{i \in S \cup T} (\sigma_i^1 - 1)c_i + \sum_{i \in S} (\sigma_i^2 - 1 + |T|)c_i + \sum_{i \in T} (\sigma_i^3 - 1 + |S|)c_i,$$

где $\sigma^1 \in \sigma^*(N, C_{S \cup T})$, $\sigma^2 \in \sigma^*(N, C_S)$, $\sigma^3 \in \sigma^*(N, C_T)$.

Так как при переходе от коалиции $S \cup T$ к S положения членов коалиции T меняются не более, чем на $|S|$, верно $\sigma_i^1 - 1 \leq \sigma_i^3 - 1 + |S|$ при всех $i \in T$.

Аналогично, $\sigma_i^1 - 1 \leq \sigma_i^2 - 1 + |T|$ для $\forall i \in S$.

Отсюда следует

$$v_q(T \cup S) - (v_q(T) + v_q(S)) \geq 0,$$

что доказывает супераддитивность функции v_q .

Маргинальный вклад агента $i \in N \setminus S$ в коалицию S , равный $v_q(S \cup \{i\}) - v_q(S)$, состоит из выплаты за ожидание непосредственно агента i , его издержек за время $(n - |S|)$ и выплаты, которую он получит от агентов, стоящих перед ним в новой (перестроенной) очереди. Действительно,

$$v_q(S \cup \{i\}) = - \sum_{j \in S \cup \{i\}} (\sigma_j - 1)c_j - (n - |S| - 1) \sum_{j \in S \cup \{i\}} c_j,$$

где $\sigma \in \sigma^*(S \cup \{i\}, C_{S \cup \{i\}})$, а $C_{S \cup \{i\}}$ – вектор издержек коалиции $S \cup \{i\}$.

Пусть $\sigma' \in \sigma^*(S, C_S)$. Рассмотрим

$$\begin{aligned} v_q(S \cup \{i\}) - v_q(S) &= \\ &= - \sum_{j \in S} (\sigma_j - 1)c_j - (\sigma_i - 1)c_i - (n - |S| - 1) \sum_{j \in S} c_j - \\ &\quad - (n - |S| - 1)c_i + \sum_{j \in S} (\sigma'_j - 1)c_j + (n - |S|) \sum_{j \in S} c_j = \\ &= - \sum_{j \in F_i(\sigma)} c_j - (\sigma_i - 1)c_i + \sum_{j \in S} c_j - (n - |S| - 1)c_i. \end{aligned}$$

Таким образом, для всех $q = (N, C)$, $S \subseteq N$, $i \in N \setminus S$

$$v_q(S \cup \{i\}) - v_q(S) = -(\sigma_i - 1)c_i - (n - |S| - 1)c_i + \sum_{j \in P_i(\sigma)} c_j,$$

где $\sigma \in \sigma(S \cup \{i\}, C_{S \cup \{i\}})$.

Для всех $q = (N, C)$ выплата i -му агенту, в соответствии с вектором Шепли [3], равна

$$SV_i(v_q) = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!(n - |S| - 1)!}{n!} [v_q(S \cup \{i\}) - v_q(S)].$$

В нашем случае

$$\begin{aligned} & SV_i(v_q) = \\ &= \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!(n - |S| - 1)!}{n!} \left[(1 - \sigma_i)c_i - (n - |S| - 1)c_i + \sum_{j \in P_i(\sigma)} c_j \right] = \\ &= - \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!(n - |S| - 1)!}{n!} \left[(n - |S| - 1)c_i + \sum_{j \in S, j \in P_i(\sigma)} (c_i - c_j) \right] = \\ &= -c_i \sum_{s=0}^{n-1} \frac{n-s-1}{n} - \sum_{j \in P_i(\sigma)} (c_i - c_j) \sum_{s=0}^{n-2} \frac{s+1}{n(n-1)} = \\ &= -\frac{n-1}{2}c_i - \sum_{j \in P_i(\sigma)} \frac{c_i - c_j}{2} = -\frac{1}{2} \left((n + \sigma_i - 2)c_i - \sum_{j \in P_i(\sigma)} c_j \right) = \\ &= -\frac{1}{2}(n + \sigma_i - 2)c_i + \sum_{j \in P_i(\sigma)} \frac{c_j}{2}, \end{aligned}$$

где $\sigma \in \sigma^*(N, C)$.

Оказывается, что существуют значения платежей t_i , при которых выигрыш агента i совпадает с его компонентой вектора Шепли. Действительно, положим

$$SV_i(v_q) = u_i(\sigma_i, t_i, c_i) = t_i - (\sigma_i - 1)c_i.$$

Тогда

$$\begin{aligned} t_i = SV_i(v_q) + (\sigma_i - 1)c_i &= \sum_{j \in P_i(\sigma)} c_j - \frac{1}{2}(n + \sigma_i - 2)c_j + \\ &+ (\sigma_i - 1)c_i = -\frac{n - \sigma_i}{2}c_i + \sum_{j \in P_i(\sigma)} c_j. \end{aligned}$$

На основании проведенных рассуждений можно сформулировать следующую теорему.

Теорема. Пусть $q = (N, C)$, распределение $z = (\sigma, t) \in \Sigma(q)$ предоставляет агентам выигрыши, соответствующий вектору Шепли для v_q . Тогда $\sigma \in \sigma^*(q)$, и для всех $i \in N$

$$t_i = -\frac{n - \sigma_i}{2} c_i + \sum_{j \in P_i(\sigma)} c_j.$$

Справедливость аксиом 1, 2 следует из определения вектора Шепли:

1. *Парето-оптимальность:* σ оптимальна для $q = (N, C)$, т.е. для всех $z' = (\sigma', t') \in \Sigma(q)$

$$\sum_{i \in N} (\sigma_i - 1) c_i \leq \sum_{i \in N} (\sigma'_i - 1) c_i,$$

и t сбалансирована, т.е. $\sum_{i \in N} t_i = 0$.

2. *Одинаковый выигрыш для агентов с равными издержками:* для любых $q = (N, C)$ и любых $z = (\sigma, t) \in \varphi(q)$, $i, j \in N$, $i \neq j$ и $c_i = c_j$ имеем $u(\sigma_i, t_i, c_i) = u(\sigma_j, t_j, c_j)$.

Таким образом, в данной статье получено решение проблемы очереди, удовлетворяющее «разумным» требованиям Парето-оптимальности и справедливости.

Литература

1. Maniquet F. A characterization of the Shapley value in queueing problems // Econ. Theory, 2003. Vol. 109. P. 90–103.
2. Kaii C., Ramaekers E. Characterizations of Pareto-efficient, fair, and strategy-proof allocation rules in queueing problems // Games and Economic Behavior, 2010. Vol. 68. P. 220–232.
3. Петросян Л.А., Зенкевич Н.А., Семина Е.А. Теория игр. М.: Высшая школа, 1998. 304 с.

Мерзлякова Ю.Д., Губар Е.А.
Санкт-Петербургский государственный университет

Оптимальная политика инкассации банкомата

Постановка задачи. Качественное обслуживание клиентов карточных платежных систем, а именно обеспечение спроса клиентов, возможно только при бесперебойном функционировании сети банкоматов, которое можно гарантировать при своевременной загрузке банкоматов денежными средствами. При этом следует учитывать не только интересы клиентов, но и банка.

В данной работе рассматривается актуальная проблема нахождения оптимального набора купюр для загрузки банкомата с целью минимизации потерь от отвлечения денежных средств при одновременном обеспечении достаточного числа денежных купюр для удовлетворения спроса клиентов. На основе статистических данных о расходе купюр в банкомате анализируется интенсивность спроса клиентов, строится аналитическая зависимость количества купюр каждого номинала в каждой из кассет банкомата, и по результатам вычислений составляется рекомендация по оптимизации текущего плана инкассации банкомата. Также проводится сравнительный анализ потерь от отвлечения денежных средств до и после оптимизации процесса загрузок банкомата.

Решение задачи. Для того, чтобы рассматривать задачу оптимальной загрузки банкомата денежными купюрами, определим основные параметры банкомата:

1. *Количество кассет в банкомате.* Будем считать, что банкомат имеет четырехкассетную конфигурацию, причем в каждую кассету загружаются купюры одного номинала. Номер кассеты обозначим через $i = \overline{1, 4}$. Максимальное количество купюр в одной кассете не может превышать 2500 штук.
2. *Номинал купюры.* Пусть в кассету i загружаются купюры номиналом N_i . Будем рассматривать купюры номиналом 100, 500, 1000 и 5000 рублей, т.е. $N_1 = 100$, $N_2 = 500$, $N_3 = 1000$, $N_4 = 5000$.
3. *Критерий необходимости загрузки банкомата* – количество купюр в одной из кассет приближается к критическому значению $k \geq 0$.

4. *Условие загрузки* – при загрузке банкомата заменяются все четыре кассеты.

Перейдем к вопросу о нахождении оптимального числа купюр в каждой кассете. Рассмотрим данные наблюдений за числом купюр в банкомате за время $t = \overline{1, T}$. Обозначим через $Y_i(t)$ количество купюр i -го номинала, находящихся в банкомате в день t . Считаем, что на следующий день, после того как количество купюр в одной из кассет уменьшится до критического значения k , производится инкассация банкомата и загрузка новых кассет.

Потери от отвлечения денежных средств, загружаемых в банкомат, можно вычислить с помощью формулы

$$C = \sum_{t=1}^T \sum_{i=1}^4 N_i Y_i(t) \frac{p}{100 \cdot 365} + mI, \quad (1)$$

где суммирование идет по дням наблюдаемого периода, m – количество инкассаций в наблюдаемый период, p – годовая процентная ставка краткосрочного кредитования, I – затраты на инкассацию, загрузку купюр в кассеты и перерасчет возвращенных купюр.

Положим $p = 8\%$ [5], затраты на одну инкассацию составляют $I = 20$ тыс. рублей, $k = 0$. Базой для вычислений послужили данные статистических наблюдений за расходом денежных средств $Y_i(t)$ в банкомате. За отчетный период 90 дней было произведено пять инкассаций. Таким образом, потери от отвлечения денежных средств составили $C = 130\,137,45$ рублей.

Рассмотрим те же данные о количестве купюр в банкомате и разработаем рекомендацию по рационализации имеющегося плана инкассации банкомата, т.е. определим оптимальное число загружаемых купюр. Пусть y_{ri} – количество загружаемых купюр номинала i при r -ой инкассации, D_{ri} – интенсивность снятия купюр i -го номинала до инкассации r , $I_i = I/4$ – затраты на загрузку одной кассеты. Тогда оптимальное число купюр, загружаемых при r -ой инкассации, будем искать, исходя из следующей задачи минимизации [3]:

$$\min C_r(y_{r1}, y_{r2}, y_{r3}, y_{r4}) = \sum_{i=1}^4 \left(N_i y_{ri} \frac{p}{100 \cdot 365} + \frac{I_i D_{ri}}{y_{ri}} \right), \quad (2)$$

при ограничениях числа купюр в кассете

$$0 \leq y_{ri} \leq 2500, \quad i = \overline{1, 4}. \quad (3)$$

Первое слагаемое под знаком суммы в формуле (2) – затраты на хранение купюр в кассете, второе – на загрузку купюр в кассеты.

Очевидно, что имеет место неравномерность снятия наличных денежных средств в течение времени, обусловленная такими факторами, как режим работы банкомата, его месторасположение, а также дни зарплаты, праздничные дни и т.д. [2]. Поэтому, по сравнению с [1], был предложен следующий способ пересчета интенсивности спроса D_i .

Рассмотрим функции $f_i(t)$, аппроксимирующие зависимость известного количества купюр $Y_i(t)$ номинала N_i в банкомате от времени. Будем искать функции $f_i(t)$ в виде $f_i(t) = a_i + b_i t$, где коэффициенты a_i , b_i находятся методом наименьших квадратов. Коэффициент b_i определяет интенсивность спроса D_i на купюры номинала i [1].

При оптимизации имеющегося плана инкассации будем придерживаться следующего алгоритма. Рассмотрим данные о количестве купюр $Y_i(t)$. Предполагаем, что при $t = 1$ была проведена инкассация, т.е. $r = 1$, количество загруженных купюр при первой инкассации оставим прежним. По известным значениям $Y_i(t)$ до инкассации $r = 2$, строим функции $f_i(t) = a_i + b_i t$ и определяем интенсивность спроса на купюры D_{2i} . Предполагая что спрос на купюры после второй инкассации не изменится, находим оптимальное количество купюр y_{2i}^* , которые рекомендуется загрузить в банкомат при второй инкассации, решая задачу (2), (3). После загрузки в банкомат купюр в количестве y_{2i}^* происходит естественный расход купюр. Получаем новые данные о количестве купюр в банкомате до инкассации $r = 3$, вновь ищем интенсивность спроса D_{3i} и предполагаем, что после третьей инкассации банкомата спрос на купюры останется прежним, вычисляем y_{3i}^* и т.д. до конца наблюдаемого периода.

Применим данную схему оптимизации к ранее рассматриваемым данным статистических наблюдений за расходом денежных средств $Y_i(t)$, $t = \overline{1, 90}$. Введем обозначение $h_i = N_i p / 36500$. При ранее введенных предположениях $h_1 = 0,02$, $h_2 = 0,11$, $h_3 = 0,22$, $h_4 = 1,10$ представляют собой стоимость хранения одной купюры номинала i в единицу времени. Результаты вычисления интенсивности спроса и оптимальное количество купюр, загружаемых в банкомат, приведены в таблице 1.

Таблица 1. Рекомендация по оптимизации имеющегося плана инкассации

Номинал	N_1	N_2	N_3	N_4
Инкассация №2				
D_{2i}	117	24	48	11
y_{2i}^*	2500	1044	1044	223
Инкассация №3				
D_{3i}	112	26	50	9
y_{3i}^*	2500	1087	1066	202
Инкассация №4				
D_{4i}	117	29	41	11
y_{4i}^*	2500	1148	965	224
Инкассация №5				
D_{5i}	115	28	47	7
y_{5i}^*	2500	1128	1033	178

После оптимизации плана инкассации были посчитаны потери от отвлечения денежных средств по формуле (1). Они составили $C = 114\,957,55$ рублей. Таким образом удалось снизить потери банка на 12%, при этом количество загружаемых купюр полностью обеспечивает ожидаемый спрос клиентов.

Заключение. В данной статье рассмотрен банкомат, в каждую кассету которого загружается разный номинал купюр. Ясно, что купюры могут быть загружены различными способами, например, две кассеты с купюрами по 100 рублей и по одной кассете с купюрами 500 и 1000 рублей. Однако, одновременное определение способа загрузки купюр и их количественного состава является сложной задачей, требующей анализа зависимости запрашиваемой клиентом денежной суммы и того набора купюр, который выдает банкомат. В свою очередь, банкомат выдает купюры в соответствии с настройками программного обеспечения и наличием необходимых для выдачи купюр.

Кроме того, возможно распространение поставленной задачи на сеть банкоматов с целью составления плана инкассаций для всей сети, учитывающего издержки как на доставку кассет до банкомата, так и потери от отвлечения денежных средств. Данная постановка задачи формулируется в [4] и носит название *Inventory Routing Problem*.

Данные задачи представляют несомненный интерес для дальнейших исследований.

Литература

1. Бояршинов М.Г., Салихова О.С. Оптимальная загрузка банкомата денежными купюрами // Вестник ПГТУ. Прикладная математика и механика, 2010. № 15. С. 128–147.
2. Мерзлякова Ю.Д., Губар Е.А. Прогнозирование денежных потоков в сети банкоматов // Процессы управления и устойчивость: Труды 41-й международной научной конференции аспирантов и студентов / Под ред. Н.В. Смирнова., Г.Ш. Тамасяна. СПб.: Издат. Дом С.-Петербур. гос ун-та, 2010. С. 671–676.
3. Таха Х.А. Введение в исследование операций. М.: Издательский дом «Вильямс», 2005. 912 с.
4. Wagner M. The optimal cash deployment strategy – modeling a network of automated teller machines // Thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in Accounting. Helsinki, Finland: Swedish School of Economics and Business Administration, 2007. P. 195–207.
5. Ставка рефинансирования ЦБ РФ. http://www.cbr.ru/print.asp?file=/statistics/credit_statistics/refinancing_rates.htm

Надеждина А.М.

Санкт-Петербургский государственный университет

Модели массового обслуживания: call-центры

Рекомендовано к публикации доцентом Свиржиным М.В.

В 2009 году одним из направлений работы территориального фонда обязательного медицинского страхования (ТФ ОМС) Санкт-Петербурга явилась работа по созданию в районах города центров записи граждан на прием к врачу. Позвонив в этот центр, пациент может записаться на прием, избегая очередей в поликлиниках. Звонки поступают в центры с 7.00 до 20.00 с различной интенсивностью. Наибольшее количество звонков приходится на период с 7.00–10.00 часов. Длительность обслуживания в среднем составляет три минуты. Если же очередной звонок поступает в систему, когда все операторы заняты, то он теряется. Исключение составляют call-центры Фрунзенского и Адмиралтейского района, в котором установлены автоответчики. Таким образом, звонок, поступивший в систему, когда все операторы заняты, не теряется, а становится в очередь (количество мест в очереди – 15), и только если очередь переполнена, пациент не дозванивается. Основной задачей данной работы является определение оптимального количества операторов, необходимых для нормальной работы городских call-центров. Данный тип исследования охватывает наиболее затратную в экономическом плане часть работы. Итак, необходимо:

1. Откорректировать, просчитать и проанализировать количество операторов, которое требуется в каждом из районных call-центров города с учетом следующих ограничений:
 - время ожидания в очереди не превышает 5 минут;
 - вероятность отказа в обслуживании не больше 0,15;
 - количество операторов (линий) в call-центре не меньше двух.
2. Рассчитать различные характеристики call-центров, по часам работы.
3. В своих подсчетах необходимо учесть пики вызовов.

Для решения задачи целесообразно использовать методы теории массового обслуживания. Система массового обслуживания (СМО) представляет собой физическую систему дискретного типа с конечным (или счетным) множеством состояний. Рассмотрим физическую систему S со счетным множеством состояний S_1, \dots, S_n, \dots . В любой момент времени t система S может быть в одном из этих состояний. Обозначим $p_k(t), k = 1, 2, \dots$ вероятность того, что в момент t система будет находиться в состоянии S_k . Очевидно, для любого t $\sum_k p_k(t) = 1$. Для СМО основным фактором, обуславливающим протекающие в ней процессы, является поток заявок. Стоит отметить, что поток заявок в нашей СМО сводится к простейшему пуассоновскому потоку. СМО в разных районах будут описываться разными моделями. В Адмиралтейском и Фрунзенском районе целесообразно использовать модель многоканальных СМО с ограниченной очередью. В остальных же районах многоканальные СМО с отказами.

Многоканальная модель массового обслуживания с отказами. Имеется n каналов (линий связи), на которые поступает поток заявок с интенсивностью λ . Поток обслуживаний каждого канала имеет интенсивность μ . Пусть S_k – состояние системы, когда в ней находится k заявок, т.е. занято k каналов. Граф состояний такой СМО приведен на рис. 1.

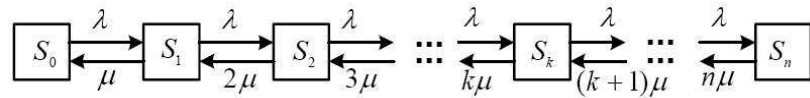


Рис. 1. Многоканальная модель массового обслуживания с отказами

Таким образом, опуская вычисления, получим следующие формулы для характеристик СМО:

1. Приведенная интенсивность потока заявок или интенсивностью нагрузки канала

$$\rho = \frac{\lambda}{\mu}.$$

2. Вероятность отказа (предельная вероятность того, что все n каналов будут заняты)

$$P_{otk} = \frac{\rho^n}{n!} p_0.$$

3. Относительная пропускная способность системы (вероятность того, что заявка будет обслужена)

$$Q = 1 - P_{otk} = 1 - \frac{\rho^n}{n!} p_0.$$

4. Абсолютная пропускная способность

$$A = \lambda Q = \lambda \left(1 - \frac{\rho^n}{n!} p_0 \right).$$

5. Среднее число занятых каналов

$$k = \rho \left(1 - \frac{\rho^n}{n!} p_0 \right).$$

Многоканальная система массового обслуживания с ограниченной очередью. Имеется n каналов обслуживания, на которые поступает пуассоновский поток заявок с интенсивностью λ . Интенсивность обслуживания заявки каждым каналом равна μ , а максимальное число мест в очереди равно m . Граф такой системы изображен на рис. 2.

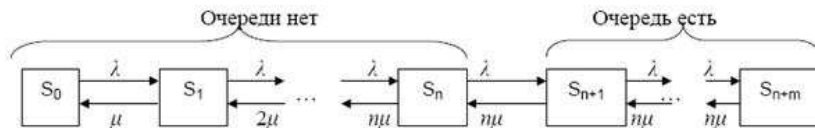


Рис. 2. Многоканальная система массового обслуживания с ограниченной очередью

Таким образом, опуская вычисления, получим следующие характеристики СМО:

1. Вероятность отказа

$$p_{otk} = p_{n+m} = \frac{\rho^{n+m}}{n^m n!} p_0.$$

2. Относительная пропускная способность

$$Q = 1 - p_{otk} = 1 - \frac{\rho^{n+m}}{n^m n!} p_0.$$

3. Абсолютная пропускная способность

$$A = \lambda Q = \lambda \left(1 - \frac{\rho^{n+m}}{n^m n!} p_0 \right).$$

4. Среднее число заявок, находящихся в очереди

$$L_{och} = \frac{\rho^{n+1}}{n n!} \frac{1 - \left(\frac{\rho}{n}\right)^m [1 + m(1 - \frac{\rho}{n})]}{\left(1 - \frac{\rho}{n}\right)^2}.$$

5. Среднее время пребывания в очереди СМО

$$t_{och} = \frac{L_{och}}{\lambda}.$$

Для определения необходимых характеристик с использованием вышеприведенных моделей теории массового обслуживания, необходимо определить следующие параметры:

1. Интенсивность входного потока λ .
2. Интенсивность обслуживания заявки μ .
3. Количество мест m в очереди (для модели СМО с ограниченной очередью).

Для решения поставленной задачи ТФ ОМС предоставил статистику по работе call-центров. Рассчитаем интенсивность λ поступления звонков на основе этих данных. Гистограмма распределения звонков для call-центра Фрунзенского района приведена на рис. 3.

Далее рассчитаем интенсивность обслуживания заявок μ в единицу времени (час). Исходя из данных, предоставленных ТФ ОМС, средняя длительность вызова составляет около трех минут. Таким образом, каждый оператор за час может обслужить $\mu = 20$ заявок.

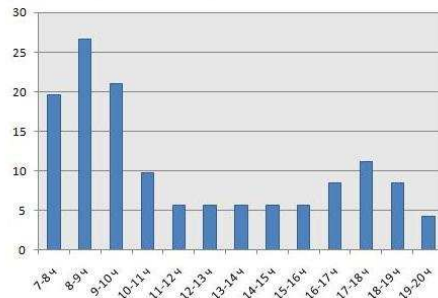


Рис. 3. Количество звонков в зависимости от времени суток

Результаты и выводы. Все вычисления выполнялись в среде Maple. Для исследований был взят предполагаемый пиковый момент загрузки call-центра с 8 до 9 часов, когда интенсивность вызовов достигает максимального значения. Определенные для такого пика числа операторов и мест в очереди являются минимальными, необходимыми для выполнения требуемых условий (вероятность отказа не больше 0,15, время ожидания в очереди не больше 5 минут), если нанять работников на весь рабочий день. В таблице приведены результаты, полученные для Фрунзенского района Санкт-Петербурга, где использовалась модель СМО с ожиданием.

Таблица. Основные характеристики СМО

Характеристики	8-9 часов
Рекомендуемое количество операторов	2
Средняя интенсивность в час-пик (λ)	27
Относительная пропускная способность (вероятность того, что заявка будет обслужена)	0,999513
Абсолютная пропускная способность (в час)	26.9869
Вероятность отказа	0.000486969
Вероятность образования очереди	0.543082
Среднее число заявок в очереди	6,62664
Среднее число заявок на обслуживании	1,34934
Среднее число заявок в системе	7,97598
Среднее время ожидания в очереди	2 мин 47 сек

Для наглядности на рис. 4. приведены графики.

Из таблицы видно, что требования к системе выполняются при количестве операторов равном 2. При таком значении n вероятность отказа равна 0,00049, что существенно ниже необходимой. То же можно сказать и про среднее время ожидания: оно составляет 2 мин. 47 сек. Однако, минимальное количество операторов в центре – 2.

Также по графику можно видеть, что при $n = 1$ характеристики системы резко ухудшаются.

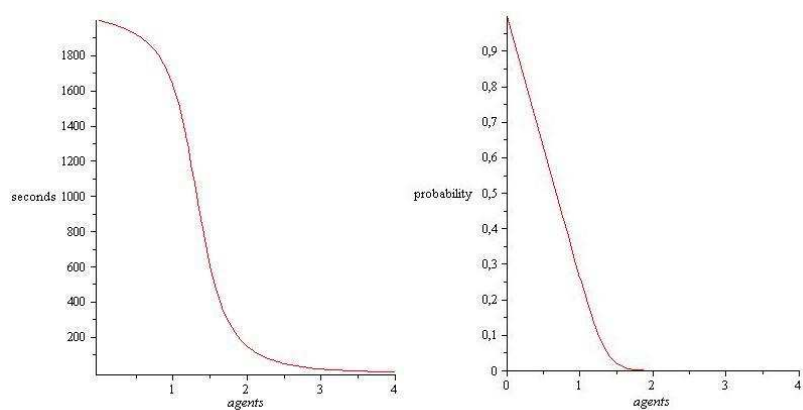


Рис. 4. Зависимость времени ожидания и вероятности отказа от количества операторов

Литература

1. Алиев Т.И. Основы моделирования дискретных систем. СПб.: СПбГУ ИТМО, 2009. 363 с.
2. Авсиевич А.В., Авсиевич Е.Н. Теория массового обслуживания: Поток требований, системы массового обслуживания. Самара: СамГАПС, 2004. 24 с.
3. Mandelbaum A., Koole G., Queueing Models of Call Centers: An Introduction // Systems Annals of Operations Research, 2002. Vol. 113. P. 41–59.

Орлов В.Г.

Санкт-Петербургский государственный университет

Сети Нэша и максимально эффективные сети

Рекомендовано к публикации профессором Петросяном Л.А.

1. Введение. В данной работе рассматривается сетевая игра n лиц. Стратегии игроков в такой игре – создавать или не создавать связи с другими игроками. От связей игроки получают прибыль. Размер прибыли зависит от игроков, участвующих в связи. Далее рассматривается возможность кооперации среди игроков, определяется характеристическая функция. Также предлагается понятие максимально эффективной сети, исследуется её связь с сетью Нэша.

2. Модель игры. Рассмотрим математическую модель сетевой игры, описанную в [1].

2.1. Постановка задачи. Пусть $N = \{1, \dots, n\}$ – множество игроков.

Определение 1. Под *стратегией* игрока i будем понимать вектор $g_i = (g_{i1}, \dots, g_{in})$, где $g_{ij} \in \{0, 1\}$ для всех $j \in N \setminus \{i\}$. Параметр $g_{ij} = 1$ тогда и только тогда, когда игрок i создал связь с игроком j , $g_{ii} = 0$.

Таким образом множество N вместе с векторами g_i образуют направленный граф (сеть), который может быть представлен в виде матрицы (g_1, \dots, g_n) .

Определение 2. Любой набор стратегий игроков $g = (g_1, \dots, g_n)$ называется *ситуацией* в игре.

Определение 3. Два игрока называются *связанными*, если хотя бы один из них создал связь с другим.

Определение 4. Игрок называется *изолированным*, если он не связан ни с одним игроком.

Определение 5. Игрок называется *крайним*, если он связан только с одним игроком.

Определение 6. *Замыканием* сети g будем называть ненаправленную сеть $\bar{g} = cl(g)$, где $\bar{g}_{ij} = \max\{g_{ij}, g_{ji}\}$ для всех $i, j \in N$.

Определение 7. *Путём* длиной k между игроками i и j будем называть последовательность игроков $(i, j_1, \dots, j_{k-1}, j)$, где $\bar{g}_{ij_1} = \bar{g}_{j_1j_2} = \dots = \bar{g}_{j_{k-1}j} = 1$.

Определение 8. Будем говорить, что игрок j *доступен* для игрока i , если между ними существует хотя бы один путь. В противном случае будем говорить, что игрок j *недоступен* для игрока i .

Определение 9. *Расстоянием* между игроками i и j будем называть длину наименьшего пути между этими игроками и обозначать $d(i, j, g)$, $d(i, j, g) \in \mathbb{Z}_+$ (множество положительных целых чисел). Если два игрока i и j не связаны, то $d(i, j, g) = \infty$, $d(i, i, g) = 0$.

Каждый игрок характеризуется двумя параметрами: весом и издержками.

Определение 10. *Вес игрока i* – это некоторый сервис, который предоставляет игрок i для других связанных с ним игроков. При оказании такого сервиса игрок i ничего не тратит. Вес обозначается v_i , где $v_i > 0$.

Определение 11. *Издержки игрока i* – это цена, которую платит игрок i за создание связи с любым другим игроком. Издержки обозначаются c_i , где $c_i > 0$.

Определение 12. Будем называть двух игроков *однородными*, если веса и издержки игроков равны.

Когда между игроками i и j есть связь, причём не важно, кто её создал, каждый игрок получает от другого прибыль, равную весу игрока. Например, если игроки i и j связаны, то игрок i получает от игрока j прибыль v_j , а игрок j получает от игрока i прибыль v_i . Также прибыль от игрока i получают другие игроки, для которых игрок i доступен, но прибыль уменьшается с увеличением расстояния между игроками.

Таким образом прибыль, которую получает игрок i от игрока j , зависит от значения v_j и от расстояния $d(i, j, g)$. Введём функцию $\Phi(v_j, d(i, j, g))$ из [1], с помощью которой можно вычислить прибыль игрока i от игрока j . Функция Φ строго возрастает по первому параметру и убывает по второму: $\mathbb{R}_+ \times \{\{0, 1, \dots, n-1\} \cup \{\infty\}\} \rightarrow \mathbb{R}$. Игрок не получает прибыли от себя или от недоступного игрока, поэтому $\Phi(v_j, 0) = \Phi(v_j, \infty) = 0$.

Общая прибыль игрока i от сети g складывается из прибылей игрока i от каждого игрока минус издержки за созданные связи. Введём $\mu_i(g) = |\{j \in N; g_{ij} = 1\}|$ – количество связей, которые создал игрок i . Теперь можно записать общую прибыль игрока i от

сети g [1]:

$$\pi_i(g) = \sum_{j \in N} \Phi(v_j, d(i, j, g)) - \mu_i(g)c_i. \quad (1)$$

2.2. Сети Нэша. Пусть D_i – множество чистых стратегий игрока i , т.е. множество всевозможных векторов g_i . Зафиксируем некоторую ситуацию $g = (g_1, \dots, g_{i-1}, g_i, g_{i+1}, \dots, g_n)$. Пусть g_{-i} – ситуация в игре, в которой игрок i использует стратегию $g_i = (0, \dots, 0)$, тогда ситуацию g , для удобства, будем обозначать $g = g_i \oplus g_{-i}$.

Определение 13. Стратегию g_i игрока i будем называть *наилучшим ответом* игрока i к ситуации g , если

$$\pi_i(g_i \oplus g_{-i}) \geq \pi_i(g'_i \oplus g_{-i})$$

для всех $g'_i \in D_i$. Пусть $BR_i(g_{-i})$ – множество всех наилучших ответов игрока i к ситуации g .

Определение 14. Сеть $g = (g_1, \dots, g_n)$ будем называть *сетью Нэша*, если $g_i \in BR_i(g_{-i})$ для каждого i , т.е. сеть образована наилучшими ответами игроков.

По существу сеть Нэша есть равновесие по Нэшу в чистых стратегиях [2].

Определение 15. Сеть Нэша будем называть *строгой сетью Нэша*, если $\pi_i(g_i \oplus g_{-i}) > \pi_i(g'_i \oplus g_{-i})$ для всех $g'_i \in D_i, g'_i \neq g_i$.

3. Кооперативный подход. Рассмотрим теперь возможность объединения игроков в коалиции.

Определение 16 [3]. Будем называть любое непустое подмножество $S \subset N$ *коалицией*.

Рассмотрим антагонистическую игру произвольной коалиции S против коалиции $N \setminus S$. Заметим, что коалиция $N \setminus S$, играя против коалиции S , с целью уменьшения выигрыша коалиции S , очевидным образом не должна создавать связи, т.е. наилучший вариант для коалиции $N \setminus S$ – использовать стратегию $(0, \dots, 0)$ всеми её игроками. Используя любую другую стратегию, коалиция $N \setminus S$ может только увеличить выигрыш коалиции S .

Исходя из этого, будем считать выигрыш коалиции S как сумму выигрышей игроков, входящих в коалицию, притом, что коалиция $N \setminus S$ использует стратегию $(0, \dots, 0)$.

Определение 17. Коалицию, содержащую всех игроков множества N , будем называть *максимальной коалицией*.

Определение 18. Сеть, которая доставляет максимальный выигрыш коалиции N , будем называть *максимально эффективной сетью* (МЭС).

Определение 19 [3]. *Характеристическая функция игры n лиц* – вещественная функция \mathcal{V} , определённая на коалициях $S \subset N$. Значение этой функции для коалиции S это максимально доступный выигрыш для этой коалиции.

Теперь, используя формулу (1), определим значение характеристической функции для каждой коалиции. Для максимальной коалиции

$$\mathcal{V}(N) = \max_g \left(\sum_{i \in N} \sum_{j \in N} \Phi(v_j, d(i, j, g)) - \mu_i(g)c_i \right).$$

Так как мы считаем выигрыш коалиции S как сумму выигрышей игроков, входящих в коалицию, с тем условием, что все игроки коалиции $N \subset S$ используют стратегию $(0, \dots, 0)$, то значение характеристической функции для коалиции S записывается аналогично:

$$\mathcal{V}(S) = \max_{g^s} \left(\sum_{i \in S} \sum_{j \in S} \Phi(v_j, d(i, j, g)) - \mu_i(g)c_i \right),$$

где $g^s = (g_1, \dots, g_n)$ такая, что $g_i = (0, \dots, 0)$ для всех $i \in N \setminus S$.

Для содержательной интерпретации числа $\mathcal{V}(S)$, как гарантированного выигрыша коалиции S , в случае, когда она действует независимо от остальных игроков, необходимо выполнение свойства супераддитивности

$$\mathcal{V}(T) + \mathcal{V}(S) \leq \mathcal{V}(T \cup S), \quad \mathcal{V}(\emptyset) = 0,$$

где $S \subset N$, $T \subset N$, $S \cap T = \emptyset$.

Утверждение 1. *Характеристическая функция удовлетворяет свойству супераддитивности.*

Рассмотрим две непересекающиеся коалиции S и T , $S \subset N$, $T \subset N$. Объединившись, эти коалиции могут придерживаться тех стратегий, которые доставляют им максимальные выигрыши, т.е. выигрыши $\mathcal{V}(S)$ и $\mathcal{V}(T)$. Таким образом, максимальный выигрыш коалиции $S \cup T$ не может быть меньше суммы максимальных выигрышей коалиций S и T , а за счёт создания новых связей он может

увеличиться. Следовательно, свойство супераддитивности выполняется.

Утверждение 2. *Существуют игры, в которых выигрыш максимальной коалиции в сети Нэша совпадает с выигрышем максимальной коалиции в МЭС, т.е. в этих играх сеть Нэша является и сетью МЭС. Также существуют игры, в которых МЭС не является сетью Нэша.*

Покажем истинность этого утверждения на примерах.

Пример 1. Рассмотрим игру шести однородных игроков. Издержки равны 24, а веса 16. Значения функции Φ приведены в таблице 1.

Таблица 1. Значения функции Φ

$d(i, j, 16)$	1	2	3	4	5
Φ	16	12	9	7	5

Теперь найдём МЭС. Пусть каждый игрок характеризуется вектором $B_i = (b_1, b_2, b_3, b_4, b_5)$, где b_j – число связей на расстоянии j у игрока i . Пусть каждая ситуация в игре характеризуется вектором $X_i = (x_1, x_2, x_3, x_4, x_5)$, где x_j – сумма j -ых компонент векторов B_i , $i = \overline{1, 6}$. Рассмотрим максимальную коалицию. Так как максимальная коалиция получает положительный выигрыш от любой связи двух игроков (каждый игрок получает от другого 16, минус издержки 24, которые платятся одним игроком), то в сети не будет изолированных игроков. Тогда $\sum_{j=1}^5 b_j = 5$. Из этого следует, что $\sum_{j=1}^5 x_j = 5 \cdot 6 = 30$. Теперь можно записать выигрыш максимальной коалиции

$$\begin{aligned} \mathcal{V}(N) &= 16x_1 + 12x_2 + 9x_3 + 7x_4 + 5x_5 - 24 \frac{1}{2}x_1 = \\ &= 4x_1 + 12x_2 + 9x_3 + 7x_4 + 5x_5. \end{aligned}$$

Для нахождения МЭС нужно решить задачу линейного программирования:

$$\left\{ \begin{array}{l} \max(\mathcal{V}) = \max(4x_1 + 12x_2 + 9x_3 + 7x_4 + 5x_5), \\ x_1 + x_2 + x_3 + x_4 + x_5 = 30, \\ x_1 \geq 0,5x_2, \quad x_1 \geq 0,33x_3, \\ x_1 \geq 0,25x_4, \quad x_1 \geq 0,2x_5. \end{array} \right.$$

Последние четыре неравенства написаны на основе рассуждений о том, что путь длины k состоит из k связей. Эта задача имеет единственное решение: $\mathcal{V} = 280, x_1 = 10, x_2 = 20, x_3 = x_4 = x_5 = 0$. Любая сеть, удовлетворяющая данному решению, будет МЭС. Все такие сети имеют одно и то же замыкание – один игрок в центре и все остальные игроки крайние, связанные с этим центральным игроком.

Сети отличаются лишь направлением связи, т.е. тем, кто создал связь. Любой вариант будет МЭС, но не любой будет сетью Нэша. Сетью Нэша будет, например, сеть, в которой все связи создали крайние игроки.

Пример 2. Рассмотрим теперь ту же игру, но с другими значениями функции Φ (см. таблицу 2).

Таблица 2. Значения функции Φ

$d(i, j, 16)$	1	2	3	4	5
Φ	16	3	2	1	0,5

Издержки, как и раньше, равны 24. В этом примере аналогично можно составить задачу линейного программирования, она также будет иметь единственное решение: $\mathcal{V} = 120, x_1 = 30, x_2 = x_3 = x_4 = x_5 = 0$. Это сети, в которых каждый игрок связан с каждым, т.е. все игроки находятся на расстоянии 1 друг от друга. В этом случае ни одна сеть МЭС не является сетью Нэша.

Литература

1. Goeree J.K., Riedl A., Ule A. In search of stars: Network formation among heterogeneous agents // Games and Economic Behavior, 2009. Vol. 67, № 2. P. 445–466.
2. Nash J.F. Equilibrium points in n -person games // Proc. Nat. Acad. Sci. USA, 1950. Vol. 36. P. 48–49.
3. Петросян Л.А., Зенкевич Н.А., Семина Е.А. Теория игр. М.: Высшая школа, 1998. 304 с.

Петров А.Г., Слобожанин Н.М.
Санкт-Петербургский государственный университет

О построении функции потенциала в задаче о кратчайшем пути

Введение. Рассмотрим задачу поиска кратчайших путей из выделенной вершины s во все остальные вершины конечного ориентированного графа $G = (L, M)$ с функцией перехода τ ($\tau : M \rightarrow Z$, Z – множество целых чисел). Одним из алгоритмов решения этой задачи является двойственный алгоритм уменьшения невязок на дугах, начинающий работу с допустимого решения двойственной задачи (потенциала) P ($P : L \rightarrow Z, P(y) - P(x) \leq \tau(x, y)$) [1], [3].

Если из вершины s существует хотя бы один путь во все остальные вершины сети, то справедливы следующие необходимые и достаточные условия оптимальности решения [6]

$$d(y) \leq d(x) + \tau(x, y), \quad \forall (x, y) \in M,$$

где $d(x)$ – длина некоторого пути из s в x .

Отсюда следует, что при отсутствии в сети циклов с отрицательным временем обхода и существовании для некоторой вершины x хотя бы одного пути во все остальные вершины, допустимым решением двойственной задачи о кратчайшем пути будет следующая функция [4]

$$P(y) = \begin{cases} 0, & y = x \\ \tau(\bar{s}(x, y)), & y \neq x, \end{cases}$$

где $\bar{s}(x, y)$ – кратчайший путь из x в y .

В [5] было показано, что от предположения о существовании для некоторой вершины путей во все остальные вершины сети можно отказаться с помощью введения дополнительных дуг с достаточно большими значениями функции перехода.

В настоящей работе предложен способ построения допустимого решения двойственной задачи для несвязной сети без изменения ее структуры. Кроме того, представлено приложение задачи о назначениях с учетом производительности.

Предварительные рассуждения. Прежде всего, заметим, что на каждой компоненте слабой связности мы вправе строить функцию потенциала независимо от других слабосвязных компонент, поэтому дальнейший анализ проведем для одной такой компоненты.

Так как в любом слабом орграфе существует остовный полумаршрут [2], можно представить компоненту слабой связности в виде объединения слабых подграфов, в каждом из которых для некоторой вершины существует путь во все остальные вершины подграфа. Вообще говоря, такое разбиение не единственно. Потребуем для наших целей, чтобы указанные подграфы были максимальными по включению.

В основе дальнейших рассуждений лежит следующее

Утверждение 1. Пусть P – допустимое решение двойственной задачи о кратчайшем пути и для некоторой вершины $y \in L$ выполнено $d^+(y) = 0$. Тогда допустимым решением также является следующая функция

$$P_1(x) = \begin{cases} P(x), & x \in L \setminus \{y\} \\ P(x) - c, & x = y, \end{cases}$$

где $c \in N$.

Доказательство. При $d^-(y) = 0$ функция потенциала может принимать в вершине $y \in L$ любое целочисленное значение. В случае, когда $d^-(y) > 0$, неравенства, в которых участвует $P(y)$, будут иметь вид $P(y) - P(z) \leq \tau(z, y)$, где $z \in \Gamma^{-1}(y)$. Тогда $P(y) - c - P(z) = P_1(y) - P_1(z) < \tau(z, y)$, что завершает доказательство.

Замечание. Проведя аналогичные рассуждения для вершины $y \in L$ такой, что $d^-(y) = 0$, получим, что допустимым решением также является следующая функция

$$P_2(x) = \begin{cases} P(x), & x \in L \setminus \{y\} \\ P(x) + c, & x = y. \end{cases}$$

Обобщением приведенного выше результата на случай, когда полустепени захода и исхода у всех вершин сети могут быть одновременно положительными, является

Утверждение 2. Пусть P – допустимое решение двойственной задачи о кратчайшем пути и L представимо в виде прямой суммы непустых подмножеств L_1 и L_2 таких, что множество начальных вершин дуг, одновременно инцидентных вершинам из L_1 и L_2 ,

принадлежит L_1 . Тогда допустимым решением также является функция

$$\tilde{P}(x) = \begin{cases} P(x) + c_1, & x \in L_1 \\ P(x) - c_2, & x \in L_2. \end{cases}$$

где $c_1, c_2 \in N$.

Доказательство. Если множество дуг, одновременно инцидентных вершинам из L_1 и L_2 , пусто, то функцию потенциала на вершинах из L_1 можно строить независимо от вершин из L_2 . Значит, к вершинам каждого из множеств в отдельности применимо следующее свойство потенциала: P – потенциал тогда и только тогда, когда $P + b$ – потенциал ($b \in Z$). Поочередно применив это свойство к L_1 и L_2 , получим функцию потенциала требуемого вида.

Рассмотрим теперь случай, когда множество дуг, одновременно инцидентных вершинам из L_1 и L_2 , не является пустым. В силу указанного ранее свойства функции потенциала ограничения двойственной задачи, соответствующие дугам, инцидентным вершинам только одного из множеств L_1 и L_2 , при изменении значений функции потенциала на произвольное целое число остаются справедливыми. Таким образом, остается проверить справедливость ограничений для дуг, одновременно инцидентных вершинам из L_1 и L_2 .

Пусть $(u, v) \in M$, $u \in L_1$, $v \in L_2$. Соответствующее неравенство будет иметь вид $P(v) - P(u) \leq \tau(u, v)$. Тогда $P(v) - c_2 - P(u) - c_1 = \tilde{P}(v) - \tilde{P}(u) < \tau(u, v)$, что завершает доказательство.

Построение функции потенциала. Будем предполагать, что в сети отсутствуют циклы с отрицательным временем обхода. Рассмотрим представление компоненты слабой связности в виде объединения слабых подграфов со свойствами, указанными выше. Возьмем в качестве исходного произвольным образом выбранный слабый подграф $G' = (L', M')$ с вершиной s' , из которой существует хотя бы один путь во все остальные вершины подграфа. Построим на G' функцию потенциала P' , соответствующую длинам кратчайших путей из s' в вершины из L' .

Либо G' совпадает с компонентой слабой связности и построение закончено, либо для G' найдется хотя бы один отличный от него подграф из разбиения такой, что существует по крайней мере одна вершина, принадлежащая обоим подграфам. Рассмотрим любой из подграфов, имеющих хотя бы одну общую вершину с G' , обозначим его $\tilde{G} = (\tilde{L}, \tilde{M})$. Удалим из \tilde{G} те общие с G' вершины, которые не яв-

ляются конечными ни для одной дуги, начальная вершина которой принадлежит \hat{G} и не является общей с G' . Построим на полученном подграфе $G'' = (L'', M'')$ функцию потенциала P'' по тому же принципу, что и в G' .

В результате проделанных операций мы получим, вообще говоря, различные значения $P'(x)$ и $P''(x)$ для $x \in D$, где $D = L' \cap L''$. Пусть $\Delta = \max_{x \in D} (P'(x) - P''(x))$. Тогда $\overline{P''} = P'' + \Delta$ будет потенциалом для G'' , причем $\overline{P''}(x) \geq P'(x)$ для $x \in D$.

Так как $d^+(x) = 0$ для $x \in D$ относительно G'' , то согласно утверждению 1 потенциалом для G'' будет следующая функция

$$\overline{\overline{P''}}(y) = \begin{cases} \overline{P''}(y), & y \in L'' \setminus D \\ P'(y), & y \in D. \end{cases}$$

Таким образом, потенциалом для графа $\hat{G} = G' \cup G''$ будет функция

$$\hat{P}(y) = \begin{cases} P'(y), & y \in L' \\ \overline{\overline{P''}}(y), & y \in L'' \setminus D. \end{cases}$$

Повторив указанную процедуру для каждого оставшегося подграфа из разбиения, принимая всякий раз в качестве исходного подграфа с построенной на нем функцией потенциала, полученный на предыдущем шаге, мы за конечное число шагов построим функцию потенциала для компоненты слабой связности в целом.

Таким образом, доказана следующая

Теорема. *Для того чтобы в конечной целочисленной сети (L, M) , τ существовал хотя бы один потенциал, необходимо и достаточно, чтобы в данной сети отсутствовали циклы с отрицательным временем обхода.*

Замечание. Приведенный способ построения функции потенциала не подходит для случая графа с множеством вершин произвольной мощности, так как величина разности значений функций потенциалов для $x \in D$ может неограниченно возрастать.

Приложение задачи о назначениях. Предположим, что кадровому агентству необходимо найти работников для обратившихся в агентство фирм. При этом каждая фирма может иметь несколько вакансий, и один и тот же работник может удовлетворять требованиям вакансий разных фирм. В случае подбора подходящего кандидата на какую-либо из вакансий агентство получает определен-

ное вознаграждение от соответствующей фирмы. Задачей кадрового агентства является максимизация прибыли.

Математическую модель задачи можно представить следующим образом. Имеется m работников I_1, \dots, I_m и n фирм J_1, \dots, J_n с вакансиями $J_{11}, \dots, J_{1,1_k}, \dots, J_{n1}, \dots, J_{n,n_k}$, где J_{jl} – l -ая вакансия j -ой фирмы. Пусть $\alpha_{i,jl} \geq 0$ – сумма, получаемая агентством в случае, когда кандидат I_i занимает вакансию J_{jl} . Нетрудно заметить, что данную задачу можно представить как задачу об оптимальных назначениях, где в качестве производительности будет выступать вознаграждение за подбор кандидата.

Литература

1. Гейл Д. Теория линейных экономических моделей. М.: ИЛ, 1963. 418 с.
2. Емеличев В.А., Мельников О.И., Сарванов В.И., Тышкевич Р.И. Лекции по теории графов М.: Наука, 1990. 384 с.
3. Йенсен П., Барнес Д. Потокное программирование. М.: Радио и связь, 1984. 392 с.
4. Петров А.Г., Слобожанин Н.М. Анализ алгоритмов целочисленного программирования // Процессы управления и устойчивость: Труды 40-й международной научной конференции аспирантов и студентов / Под ред. Н. В. Смирнова, Г. Ш. Тамасяна. СПб.: Издат. Дом С.-Петербур. гос. ун-та, 2009. С. 652–655.
5. Петров А.Г., Слобожанин Н.М. Анализ функции потенциала в задаче о кратчайшем пути // Процессы управления и устойчивость: Труды 41-й международной научной конференции аспирантов и студентов / Под ред. Н. В. Смирнова, Г. Ш. Тамасяна. СПб.: Издат. Дом С.-Петербур. гос. ун-та, 2010. С. 688–691.
6. Ahuja R.K., Magnanti T.L., Orlin J.B. Network flows: theory, algorithms, and applications. New Jersey: Prentice-Hall, 1993. 864 p.

Свищикова М.В.

Санкт-Петербургский государственный университет

Один метод прогнозирования ВВП России

Рекомендовано к публикации профессором Колбиньим В.В.

Введение. Одной из актуальных проблем является прогнозирование значений макроэкономических показателей. Эксперты во всем мире, используя различные математические методы и модели, рассматривают множество существенных экономических характеристик. Основным показателем состояния экономики в мире считается валовой внутренний продукт (ВВП) и поэтому понятно стремление специалистов построить как можно более точный метод прогнозирования изменения данного показателя. Ясно, что качество подобных методов будет выше, если они учитывают национальную специфику: особенности статистического учета и структуру экономики. В пределах России прогноз ВВП долгое время осуществлялся Министерством экономического развития и Фондом экономических исследований «Центр развития». Невысокая надежность их прогнозов побудила Российскую экономическую школу (РЭШ) совместно с ИК «Ренессанс Капитал» создать «Опережающий индикатор ВВП РенКап – РЭШ» [1]. Все три упомянутые системы прогнозирования высокочувствительны как в научном смысле, так и в практическом применении. Например, в [1] используется сглаживание выбросов, метод целевых предикторов, метод статистических факторов, фильтр Калмана, метод главных компонент [2, 3]. Текущие трудозатраты тоже впечатляют: сбор статистических данных на протяжении квартала происходит пятикратно, а общее количество факторов составило 108. В этой статье получен сопоставимый по качеству прогноз роста ВВП, но не поквартально, а на год и по значительно меньшему количеству факторов.

1. Описание метода. Здесь излагается метод прогнозирования ВВП, адаптированный для России. Отдельной проблемой прогнозирования являются метаморфозы экономики в период кризиса, когда каждый месяц приносит значительные изменения ВВП. Понятно, что адекватный прогноз на будущий год по состоянию экономики в текущем году можно получить лишь в некризисные годы. Таких в России мы выделим восемь: с 2000 г. по 2007 г.

Прогноз изменения ВВП будем производить с помощью линейной регрессии, которая в силу своей простоты наиболее устойчива к погрешностям статистических наблюдений. Обозначим ее через

$$R(\alpha, d) = \alpha_0 + \sum_{i=1}^n \alpha d_i,$$

где d_i – выбранные экономические факторы, α_i – неизвестные коэффициенты, подлежащие определению, n – количество факторов. Для удобства выкладок положим $d_0 = 1$ и построим расширенный столбец факторов $d = (d_0, \dots, d_n)^T$. Из параметров α_i создадим строку $\alpha = (\alpha_0, \dots, \alpha_n)$. Тогда $R = \alpha d$. Коэффициенты будем определять из соображений наилучшего соответствия ВВП

$$R(\alpha, d^j) \approx V^j, \quad j = \overline{1, k}, \quad (1)$$

где V^j – ВВП на j -ый год, d^j – расширенный столбец факторов на j -ый год, k – количество лет, взятых для построения регрессии.

Отдельной задачей является вопрос о том, какие факторы нужно включить в модель. Рассматривались следующие факторы: F_1 – экспорт нефти и газа (млн. долл.), F_2 – суммарный доход от экспортных пошлин на нефть и газ (млрд. долл.), F_3 – среднегодовая численность занятых в экономике (млн. чел.), F_4 – инфляция (%), F_5 – среднегодовой курс рубля к доллару (руб./долл.), F_6 – внешняя трудовая миграция (тыс. чел.), F_7 – налог на добычу полезных ископаемых (НДПИ) на нефть (руб./т.), F_8 – НДПИ на природный газ (руб./тыс. м³), F_9 – доход от НДПИ на нефть (млрд. руб), F_{10} – цена на нефть сорта Юралс (долл./барр.), F_{11} – время (годы). Среди перечисленных выше факторов есть связанные с нефтью и газом. Из них два – F_1 и F_2 – вводились во все рассмотренные комбинации, что объясняется современными особенностями национальной экономики России. Вычислительные эксперименты проводились следующим образом. В качестве начальной была взята выборка (F_1, F_2, F_3, F_4, F_5). На место каждого из факторов кроме F_1 и F_2 подставлялись по очереди остальные факторы $F_6 - F_{11}$. Таким образом, всего было рассмотрено 18 комбинаций. Лучший результат получился при использовании информации о значениях факторов (F_1, F_2, F_3, F_4, F_6). Отметим, что сюда не попал курс рубля, что согласуется с исследованием [4].

В таком виде модель (1) является описательной, так как она выражает ВВП некоторого года через параметры этого же года. Ее можно было бы использовать для прогноза ВВП за прошедший год на протяжении того времени, когда уже появилась статистическая информация о факторах, но еще нет таковой о ВВП. Однако факторы $F_1 - F_{11}$ оказываются доступными одновременно с информацией о ВВП. Более важным и интересным является прогноз прироста ВВП на будущее, поэтому предлагается связать параметры минувшего года с ВВП последующего или же с приростом ВВП в последующем году. Поскольку ВВП за минувший год известен, прогноз ВВП на последующий год получится сложением прошлогоднего ВВП и прогноза прироста ВВП на последующий год.

Эта модель строилась по данным за период с 2000 по 2006 гг., а прогноз строился на 2007 год. Этот промежуток времени был выбран специально, как относительно спокойное время между двумя кризисами.

Прогнозирующая модель имеет вид

$$R(\alpha, d^j) \approx V^{j+1} - V^j = \Delta V^j(\alpha, d^j).$$

Таким образом, регрессия R является аппроксимантом ΔV . Для того чтобы получить наилучшее приближение изменения ВВП, воспользуемся методом наименьших квадратов (МНК). Ведем критерий качества регрессии. Для этого составим [5] целевую функцию – функционал F вида

$$F(\alpha) = \sum_{j=2000}^{2005} [R(\alpha, d^j) - \Delta V^j]^2 \longrightarrow \min_{\alpha}. \quad (2)$$

Чем меньше значение F , тем лучше регрессия R . У функционала $F(\alpha)$ существует минимум. Доказано, что множество минимайзеров в (2) совпадает с множеством решения системы уравнений

$$\nabla_{\alpha} F = 0.$$

Преобразуем (2) к удобному для дифференцирования виду

$$F(\alpha) = \sum_{j=2000}^{2005} [R(\alpha, d^j) - \Delta V^j]^2 =$$

$$\begin{aligned}
&= \sum_{j=2000}^{2005} \left[R(\alpha, d^j)^2 - 2R(\alpha, d^j)\Delta V^j + (\Delta V^j)^2 \right] = \\
&= \sum_{j=2000}^{2005} \left[(\alpha d^j) ((d^j)^T (\alpha)^T) - 2\alpha d^j \Delta V^j + (\Delta V^j)^2 \right] = \\
&= \alpha \left(\sum_{j=2000}^{2005} d^j (d^j)^T \right) \alpha^T - 2\alpha \sum_{j=2000}^{2005} d^j \Delta V^j + \sum_{j=2000}^{2005} (\Delta V^j)^2.
\end{aligned}$$

После дифференцирования получаем систему

$$\frac{1}{2} \nabla F_\alpha = \alpha \sum_{j=2000}^{2005} d^j (d^j)^T - \sum_{j=2000}^{2005} d^j \Delta V^j = 0. \quad (3)$$

Решая (3) (например, методом Гаусса), получаем искомые коэффициенты α_k , $k = \overline{1, n}$. Отметим, что $d_0^j = 1$, $\forall j$, т.е. в модель введен свободный член, что повышает качество экстраполяции. Таким образом, модель является 5-факторной, но 6-параметрической.

В таблице представлены входные данные, которые использовались для расчета коэффициентов регрессии и прироста ВВП на 2007 год.

Таблица. Входные статистические данные

годы	Δ ВВП \$млрд.	Своб. член	Инфл. %	Эксп. нефти и газа (\$млн.)	Эксп. пошл. \$млрд.	Миграция тыс. чел.	Труд-ся млн. чел.
2000	46,8	1	20,2	41916	3,85	213	64,5
2001	38,5	1	18,8	42760,3	5,81	284	64,7
2002	85,3	1	15,1	45010,4	4,32	360	65,4
2003	160,4	1	12,0	59659,9	7,93	378	65,7
2004	173,1	1	11,7	80898	21,08	460	66,4
2005	226,1	1	10,9	115108,5	42,48	702	66,8
2006	291,9	1	9,1	146089,1	62,08	1014	67,2

После того как найдены коэффициенты α_k , $k = \overline{1, n}$, можно вычислить значение регрессии $R(\alpha, d^{2006})$. Тогда получим прогноз ΔV

на 2007 год. На практике данная модель прогноза ВВП России показала хорошие результаты. Так как ВВП за 2007 год известен, то сравнивая полученный прогноз прироста ВВП по данной модели с реальным приростом ВВП за 2007 год, можно вычислить относительную погрешность прогноза. Она составила 6%, это не более 21,6 млрд. долл.

В результате минимизации квадратичного отклонения были получены следующие значения коэффициентов регрессии и относительной погрешности изменения ВВП: $\alpha_1 = 355,812$, $\alpha_2 = 1,999$, $\alpha_3 = -11,285$, $\alpha_4 = 0,009$, $\alpha_5 = -14,639$, $\alpha_6 = 0,256$. Согласно им было вычислено значение регрессии на 2007 год по данным за 2006 год, и оно оказалось близко к реальному изменению ВВП в 2007 году. Относительная погрешность $\Delta V \approx 6\%$.

Заключение. Представляется перспективным введение в качестве входных данных не среднегодовые значения факторов, а поквартальные, так как увеличение числа узлов для построения регрессии должно улучшить качество предсказания ВВП.

Литература

1. Опережающий индикатор ВВП РенКап — РЭШ.
http://www.nes.ru/dataupload/files/science/Leading_GDP_Indicator_10_Sep_2010_Rus.pdf
2. Айвазян С. А., Бухштабер В. М., Енюков И. С., Мешалкин Л. Д. Прикладная статистика. Классификация и снижение размерности. М.: Финансы и статистика, 1989. 607 с.
3. Bai J., Ng S. Determining the Number of Factors in Approximate Factor Models // *Econometrica*, 2002. Vol. 70. P. 191–221.
4. Смирнов С. В. Режимы валютного курса и стабильность экономики // *Вопросы экономики*, 2010. № 1. С. 29–43.
5. Колбин В. В., Шагов А. В. Модели принятия решений. СПб.: Изд-во СПбГУ, 2002. 48 с.

Серяков И.А., Петросян Л.А.
Санкт-Петербургский государственный университет

Об условно оптимальном перемещении n игроков по непересекающимся путям

1. Введение. В статье рассматривается задача об условно оптимальном перемещении n игроков в определенную точку на графе по непересекающимся путям. Предложен алгоритм для согласованного поиска путей каждого игрока из множества N , приводящий к ситуации, где суммарные затраты на перемещение всех игроков будут приближены к минимальным. Также рассмотрена кооперативная модель поведения игроков.

2. Основные понятия. Рассмотрим граф $G(X, E)$, где $X = \{1, \dots, m, A\}$ – множество его вершин, а E – ребра, соединяющие эти вершины. Для простоты рассмотрим неориентированный граф, т.е. перемещение по ребрам возможно в обоих направлениях. Пусть (i, j) – ребро графа, соединяющее его вершины $i, j \in X$. Каждое ребро (i, j) имеет свои затраты на перемещение $\alpha(i, j)$, одинаковые для всех игроков и удовлетворяющие следующим условиям:

1. $\alpha(i, j) = \alpha(j, i)$.
2. $\alpha(i, i) = 0$.

Рассмотрим множество игроков $N = \{1, \dots, n\}$, где $n < m$, которые находятся в некоторых вершинах данного графа за исключением вершины A . Несколько игроков могут находиться в одной и той же точке графа. Все игроки из N хотят перебраться из вершины, где они находились первоначально, в вершину A с минимальными затратами на перемещение. При этом по ребру графа может переместиться только один игрок, т.е. после его перемещения данное ребро исключается из множества допустимых для перемещения ребер.

Пусть у игрока $i \in N$ имеется всего r разных путей перемещения в вершину A . Обозначим путь игрока i через

$$P_i^l = \{(j_0^l, j_1^l)_i, (j_1^l, j_2^l)_i, \dots, (j_{k-1}^l, j_k^l)_i\},$$

где $l \in \{1, \dots, r\}$, k – число ребер в данном пути, $j_0^l = i$, $j_k^l = A$. Тогда затраты игрока i на пути P_i^l

$$a(P_i^l) = \sum_{p=1}^k \alpha(j_{p-1}^l, j_p^l)_i.$$

3. Кооперативная модель поведения. Рассмотрим кооперативную модель поведения игроков. Ее целесообразность легко показать на примере (см. рис. 1).

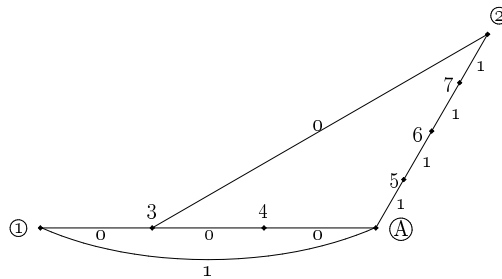


Рис. 1. Пример игры для двух игроков

Как можно увидеть из рис. 1, в игре участвуют два игрока, которые первоначально расположены в вершинах 1 и 2 графа. Рассмотрим поведение игроков в данном графе, если они действуют несогласованно [2]. Игрок 1 выберет наименее затратный путь до A : $P_1^1 = \{(1, 3), (3, 4), (4, A)\}$. Его затраты при этом будут равны нулю. Наименее затратный путь для игрока 2: $P_2^1 = \{(2, 3), (3, 4), (4, A)\}$. Его затраты при этом будут также равны нулю. Но игрок 1 уже прошел через ребра $(3, 4)$ и $(4, A)$. Следующий по наименьшим потерям путь игрока 2: $P_2^2 = \{(2, 3), (3, 1), (1, A)\}$ также нереализуем для игрока 2, так как игрок 1 уже прошел через ребро $(1, 3)$. Остается только путь $P_2^3 = \{(2, 7), (7, 6), (6, 5), (5, A)\}$. На нем игрок 2 несет потери на перемещение, равные 4.

Очевидно, что в кооперативной модели поведения игрок 1 пошел бы путем $P_1^2 = \{(1, A)\}$ с затратами 1, а игрок 2 прошел бы по пути $P_2^1 = \{(2, 3), (3, 4), (4, A)\}$ с нулевыми потерями [3]. Получилось, что суммарные затраты игроков в первом случае и во втором равны, соответственно, 4 и 1. Смысл кооперации очевиден.

3.1. Алгоритм поведения коалиции N . Предположим, что для каждого игрока $i \in N$ существует путь в вершину A , не пересекающийся с путями других игроков. Пусть каждый игрок нашел с помощью какого-либо алгоритма свой наименее затратный путь P_i^*

[5], [6], [8]. Получаем набор путей $P^* = (P_1^*, \dots, P_n^*)$. Далее следуем согласно алгоритмам 1 и 2 (см. рис. 2, 3).

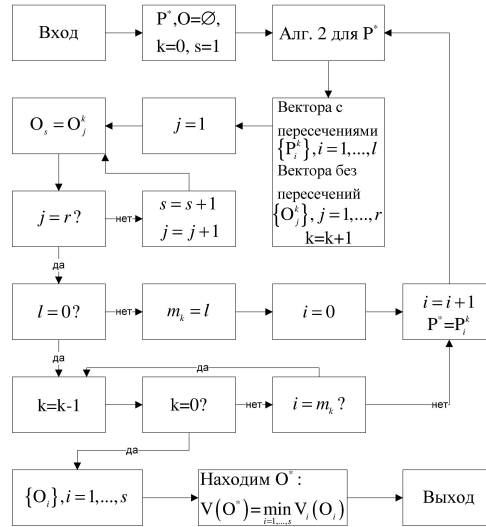


Рис. 2. Блок-схема алгоритма 1

На выходе алгоритма получаем набор путей игроков из N , который дает суммарные затраты всех игроков, которые мы считаем близкими к минимальным. Рассмотрим данный алгоритм на примере (см. рис. 4).

В силу простоты графа очевидны оптимальные пути для каждого игрока [4]: $P_1^* = \{(1, 4), (4, 7), (7, 10), (10, A)\}$; $P_2^* = \{(2, 5), (5, 8), (8, A)\}$; $P_3^* = \{(3, 6), (6, 9), (9, 8), (8, A)\}$. Суммарные затраты на перемещение в этом случае равны $V(N) = 7$. Но игрок 2 и игрок 3 имеют общее ребро в своих путях $(8, A)$. Следуя алгоритму 1, обрабатываем набор путей (P_1^*, P_2^*, P_3^*) с помощью алгоритма 2. Сперва оставляем на пути P_2^* игрока 2, а для игрока 3 ищем оптимальный путь в новом графе, с исключенным путем игрока 2. Затем оставляем на пути P_3^* игрока 3, а для игрока 2 ищем оптимальный путь в новом графе, с исключенным путем игрока 3. На выходе алгоритма 2 имеем два набора путей: P_1^1 и O_1^1 , где среди путей O_1^1 нет пересечений, а среди путей P_1^1 есть. Далее, следуя алгоритму, запоминаем набор путей O_1^1 [1]. Присваиваем $m_1 = 1$ и обрабатываем вектор путей P_1^1 с помо-

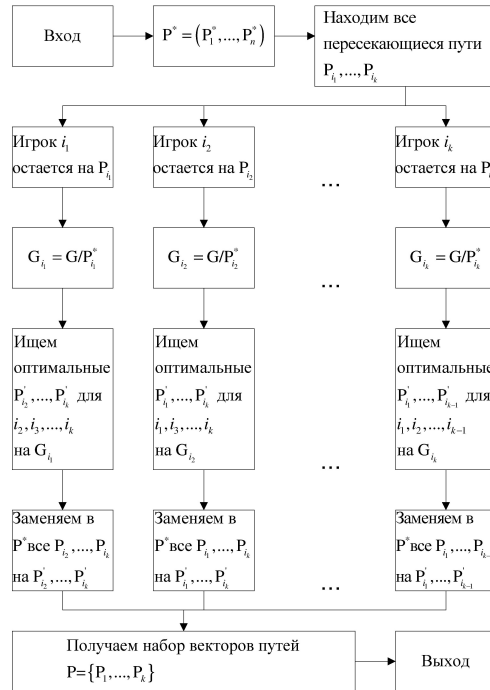


Рис. 3. Блок-схема алгоритма 2

стью алгоритма 2. Оставляем на своем пути игрока 1, а для игрока 2 ищем оптимальный путь в графе, с исключенными путями игроков 3 и 1. Затем оставляем на своем пути игрока 2, а для игрока 1 ищем оптимальный путь в графе, с исключенными путями игроков 3 и 2. В обоих случаях мы не смогли добиться того, чтобы все игроки добрались до вершины A . Таким образом на выходе алгоритма 2 имеем пустое множество. Алгоритм 1 завершился, и на выходе мы имеем единственный набор путей O_1^1 , который и будем считать условно оптимальным.

Теперь оставим предположение о том, что для каждого игрока $i \in N$ существует путь в вершину A , не пересекающийся с путями других игроков. Введем штрафные санкции на игроков, не достигших вершины A . Пусть это будет некоторая фиксированная величина t . В общем алгоритм остается прежним, за одним исключением. Теперь, когда у игрока заканчиваются варианты путей, мы записываем на него штраф t и далее продолжаем алгоритм без него.

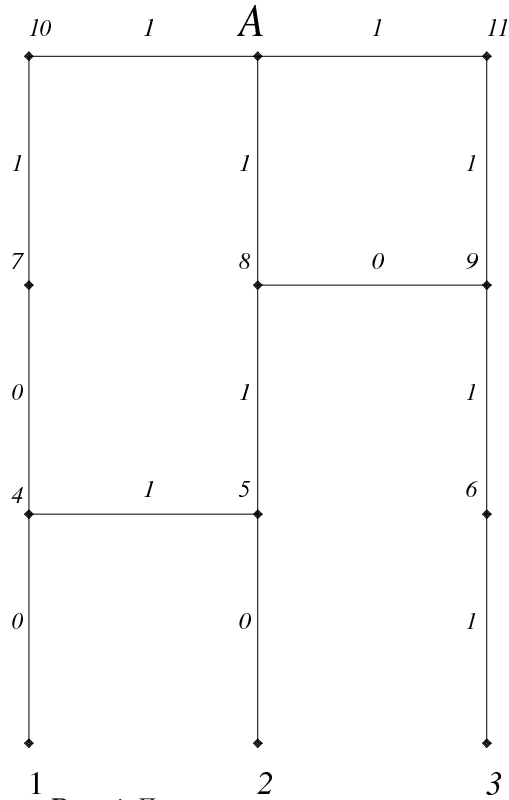


Рис. 4. Пример игры для трех игроков

3.2. Алгоритм поведения коалиции S . Рассмотрим коалицию $S \subset N$. Наихудший вариант развития событий для коалиции S , если все игроки из N/S объединятся в свою коалицию и смогут достигнуть вершины A по приведенному выше алгоритму. Пусть у них получились пути P_j^* , $j \in N/S$. Тогда, чтобы приблизить свои затраты к минимуму, коалиция S должна следовать алгоритму 1, но на графе $G(X, E / \bigcup_{j \in N/S} P_j^*)$. Запишем это:

$$V(S) = \sum_{i \in S} a(P_i^*(G(X, E / \bigcup_{i \in N/S} P_i^*))).$$

Здесь под $P_i^*(\cdot)$ понимаем путь игрока i , найденный по приведенному выше алгоритму.

4. Заключение. Итак, была сформулирована новая теоретико-игровая модель. Показана целесообразность кооперации игроков между собой для минимизации своих расходов на перемещение. Предложен алгоритм поиска путей игроков, затраты на которых наиболее близки к минимальным. Задача о поиске путей, требующих минимальные затраты на перемещение, пока остается открытой.

Литература

1. Корбут А.А., Финкельштейн Ю.Ю. Дискретное программирование. М.: Наука, 1969. 370 с.
2. Bellman R.E. On a routing problem // *Quart. Appl. Math.*, 1958. Vol. 16. P. 87–90.
3. Hu T.C. *Integer programming and network flows*. Mass.: Addison-Wesley, 1969. 452 p.
4. Balinski M.L., Quandt R.E. On an integer program for a delivery problem // *Operations Research*, 1964. Vol. 12. P. 300–304.
5. Bellmore M., Nemhauser G.L. The travelling salesman problem: A Survey // *Operations Research*, 1968. Vol. 16. P. 538–558.
6. Dantzig G.B., Fulkerson D.R., Johnson S.M. On a linear programming, combinatorial approach to the travelling salesman problem // *Operations Research*, 1959. Vol. 7. P. 58–66.
7. Land A.H., Doig A.G. An automatic method of solving discrete programming problems // *Econometrica*, 1960. Vol. 28. P. 497–520.
8. Little J.D.C., Murty K.G., Sweeney D.W. Karel C. An algorithm for the travelling salesman problem // *Operations Research*, 1963. Vol. 11. P. 972–989.

Степанова О.О.

Санкт-Петербургский государственный университет

Процесс инвестирования: применение уравнения Гамильтона – Якоби – Беллмана и метода прямых

Рекомендовано к публикации доцентом Дорофеевым Б.В.

1. Введение. В течение жизни человек зарабатывает деньги, которые может тратить по-разному. Кто-то вкладывает деньги в свое образование, кто-то тратит их на покупки, а кто-то зарабатывает на инвестировании. Последнее можно разделить на различные категории: вложения в недвижимость, в ценные бумаги, в бизнес и так далее. Ответ на вопрос, куда вложить деньги, зависит от того, насколько вы готовы рисковать. В инвестиционной деятельности существенное значение имеет фактор риска. Чем он выше тем, как правило, можно больше заработать.

В условиях рыночной экономики возможностей для инвестирования довольно много. В данной задаче мы проанализируем процесс инвестирования денег в акции. На протяжении всего процесса человек формирует свой портфель ценных бумаг с учетом того, чтобы риск был минимальным, т.е. его стратегия заключается в устранении риска. Поэтому он вкладывает все свои сбережения в акции: с риском и без риска. Момент первого вложения фиксируется ($t = 0$), а полученный процесс либо продолжается до момента, когда человек выходит на пенсию (T), либо – когда он внезапно умирает (τ). Задача заключается в оптимизации капитала человека в эти моменты, вне зависимости от его судьбы.

Для решения поставленной задачи используется два способа: решение уравнения Гамильтона – Якоби – Беллмана и метод прямых решения краевой задачи для уравнения в частных производных.

2. Постановка задачи. В работе [1] анализируется случай инвестирования капитала без страхования риска. Для этого рассматривается функция $V(x, t)$, отвечающая за оптимизацию величины вложенных средств:

$$V(x, t) = \sup_{\Theta_s \in A} E[u(x_T) | x_t = x], \quad (1)$$

где $u(x)$ – возрастающая функция капитала, которая показывает отношение к риску, A – множество стратегий вложения, Θ_s – часть капитала, вложенного в акции с риском, которая удовлетворяет условию:

$$\int_t^T \Theta_s^2 ds < \infty.$$

Функция $V(x, t)$ является решением уравнения Гамильтона – Якоби – Беллмана [2]:

$$\begin{cases} V_t'(x, t) + \max_{\Theta} \left[(\mu - r)\Theta V_x'(x, t) + \frac{1}{2}\sigma^2\Theta^2 V_{xx}''(x, t) \right] + \\ \quad + rxV_x'(x, t) = 0, \\ V(x, T) = u(x). \end{cases} \quad (2)$$

Максимум в системе (2) достигается, поэтому обозначим

$$\max_{\Theta} \left[(\mu - r)\Theta V_x'(x, t) + \frac{1}{2}\sigma^2\Theta^2 V_{xx}''(x, t) \right] = B,$$

тогда выполнено $B' = 0$, т.е.

$$(\mu - r)V_x'(x, t) + \sigma^2\Theta V_{xx}''(x, t) = 0.$$

Откуда получаем, что оптимальной является стратегия

$$\Theta_t^* = -\frac{(\mu - r)V_x'(x_t^*, t)}{\sigma^2 V_{xx}''(x_t^*, t)}, \quad (3)$$

а x_t^* – оптимальное богатство (капитал).

Подставляя (3) в систему (2), получим

$$\begin{cases} V_t'(x, t) - \frac{(\mu - r)^2}{2\sigma^2} \frac{V_x'^2(x, t)}{V_{xx}''(x, t)} + rxV_x'(x, t) = 0, \\ V(x, T) = u(x). \end{cases} \quad (4)$$

Исходя из предположения о том, что человек стремится избегать риска, полагаем, что

$$u(x) = -\frac{1}{2}e^{-\alpha x}, \quad \alpha > 0. \quad (5)$$

Используя условие (5) для системы (4), получим решение задачи нахождения оптимальной инвестиционной стратегии:

$$\begin{cases} V(x, t) = -\frac{1}{2} \exp(-\alpha x e^{r(T-t)}) \exp\left(-\frac{(\mu-r)^2}{2\sigma^2}(T-t)\right), \\ \Theta_t^* = \frac{(\mu-r) e^{-r(T-t)}}{\sigma^2 \alpha}. \end{cases} \quad (6)$$

3. Решение поставленной задачи методом прямых. Будем искать функцию $V(x, t)$, удовлетворяющую на границе прямоугольника $[0, X] \times [0, T]$ системе (4) и краевым условиям, полученным из решения системы (6):

$$\begin{cases} V(0, t) = -\frac{1}{2} \exp\left(-\frac{(\mu-r)^2}{2\sigma^2}(T-t)\right), \\ V(X, t) = -\frac{1}{2} \exp(-\alpha X e^{r(T-t)}) \exp\left(\frac{(\mu-r)^2}{2\sigma^2}(T-t)\right). \end{cases} \quad (7)$$

Воспользуемся методом прямых приближенного решения краевой задачи (4)–(7) [3]. Для этого разделим отрезок $[0, X]$ на n равных частей с шагом h , через внутренние точки деления проведем семейство параллелей $x = x_j$. На каждой такой прямой дифференциальное уравнение (4) приближенно заменим обыкновенным дифференциальным уравнением для искомых функций $V(x_j, t) = V_j(t)$. Для этого в уравнении (4) избавимся от частного дифференцирования по x с помощью формул численного дифференцирования

$$\begin{cases} \left. \frac{\partial V}{\partial x} \right|_{x=x_j} = \frac{V_{j+1}(t) - V_{j-1}(t)}{2h}, \\ \left. \frac{\partial^2 V}{\partial x^2} \right|_{x=x_j} = \frac{V_{j+1}(t) - 2V_j(t) + V_{j-1}(t)}{h^2}. \end{cases} \quad (8)$$

Тогда, подставляя выражения (8) в уравнение (4), получим систему

обыкновенных дифференциальных уравнений

$$\begin{cases} V_j'(t) - \frac{(\mu - r)^2}{2\sigma^2} \frac{V_{j+1}(t) - V_{j-1}(t)}{4(V_{j+1}(t) - 2V_j(t) + V_{j-1}(t))} + \\ \quad + rx_j V_j(t) = 0, \\ V_j(T) = u(x_j). \end{cases} \quad (9)$$

Таким образом, от линейного дифференциального уравнения (4) с частными производными мы перешли к системе (9) из $(n-1)$ обыкновенных дифференциальных уравнений с $n-1$ неизвестными функциями $V_1(t), \dots, V_{n-1}(t)$. Задача свелась к решению краевой задачи системы обыкновенных дифференциальных уравнений.

4. Численный анализ. Используя метод прямых для решения краевых задач в частных производных, получена приближенная оценка решения задачи нахождения оптимальной инвестиционной стратегии. В результате можно сделать вывод о том, что решение действительно является оптимальным и может быть использовано при выборе объема капитала, который необходимо вложить в акции. Также решение позволяет прогнозировать момент, подходящий для продажи акций.

Литература

1. Xiao Yu. Pricing a contract of linking home reversion plan and long-term care insurance via the principle of equivalent utility // Springer, 2011. Vol. 45. P. 465–475.
2. Кротов В.Ф., Лагоша Б.А., Лобанов С.М. и др. Основы теории оптимального управления. М.: Высшая школа, 1990. 430 с.
3. Демидович Б.П. и др. Численные методы анализа. М.: Наука, 1967. 368 с.

Шишкин В.В.

Санкт-Петербургский государственный университет

Трехвекторная модель параметрической устойчивости развития предприятия торговли

Рекомендовано к публикации профессором Покровским А.Н.

Эффективное антикризисное управление параметрической устойчивостью развития предприятия торговли предполагает необходимость проведения поэтапного комплексного экономического анализа торгового предприятия, определения его экономического потенциала и надежности. Для этого экономистами-аналитиками строится квадратная матрица, включающая все существенные показатели хозяйственной деятельности торгового предприятия. Разработанная нами на основе квадратной (5×5) -матрицы трехвекторная модель параметрической устойчивости развития предприятия торговли (метастабильное развитие), базирующаяся на системе обратных связей, представлена на рис. 1 и в таблице 1 следующими составляющими:

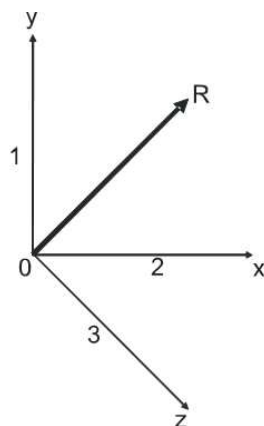


Рис. 1. Трехвекторная модель параметрической устойчивости развития торгового предприятия (метастабильное развитие)

- экономический вектор (ось oy): $K_{\text{Э}} > 1$; $J_{\text{ЭХД}} \geq 1$; $T_{J_{\text{ЭХД}}} \geq 1$; K_r ;
- финансовый вектор (ось ox): $ЗЗ \leq СОС + ЗК + ОИ$;

– вектор конкурентоустойчивости (ось oz): умеренно выраженная конкуренция;

где $K_{\text{Э}}$ – коэффициент эффективности хозяйственной деятельности; $J_{\text{ЭХД}}$ – интегральный показатель экономической эффективности хозяйственной деятельности; $T_{J_{\text{ЭХД}}}$ – темп изменения интегрального показателя экономической эффективности хозяйственной деятельности; K_r – коэффициент регулирования; $ЗЗ$ – запасы и затраты предприятия; $СОС$ – собственные оборотные средства; $ЗК$ – заемный капитал; $ОИ$ – источники, ослабляющие финансовую напряженность; $ОР$ – результирующий вектор параметрической устойчивости развития, характеризующий метастабильное состояние предприятия торговли.

Экономический вектор параметрической устойчивости и развития торгового предприятия как экономической функциональной системы (ЭФС) для нашей модели определяется:

– на основании расчета коэффициента эффективности хозяйственной деятельности ($K_{\text{Э}}$), который является средней арифметической величиной индексов (темпов роста) относительных показателей хозяйственной деятельности, расположенных под главной диагональю матрицы;

– на основе расчета интегрального показателя экономической эффективности хозяйственной деятельности ($J_{\text{ЭХД}}$) и темпов его изменения ($T_{J_{\text{ЭХД}}}$);

– на основании введенного нами коэффициента регулирования, представляющего собой отношение экономической прибыли предприятия после введения инноваций к таковой до их введения (K_r).

Финансовый вектор данной модели определяется платежеспособностью предприятия во времени и рассчитывается по абсолютным показателям (таблица 1). Финансовая устойчивость предприятия, как известно, бывает четырех типов [1–4, 7]:

– абсолютная устойчивость, когда все запасы и затраты ($ЗЗ$) полностью покрываются собственными оборотными средствами ($СОС$), т.е. $ЗЗ < СОС$. В экономике существует такое понятие, как «конкурентное преимущество фирмы». Это состояние как раз и обуславливает устойчивое финансовое положение фирмы, при котором потребности в денежных ресурсах в тактическом и стратегическом плане удовлетворяются полностью самой фирмой. Однако, такая ситуация в России встречается крайне редко и ее нельзя признать идеальной, поскольку возможности внешних источников финансирования ока-

зываются нереализованными;

– нормальная устойчивость финансового состояния предприятия соответствует условию $ЗЗ = СОС + ЗК$. Таким образом, предприятие для покрытия запасов и затрат привлекает заемный капитал (ЗК) и умело варьирует финансовые источники, обеспечивая свою платежеспособность;

Таблица 1. Параметры трехвекторной модели, определяющие тип устойчивости предприятия торговли

Экономический вектор (oy)	Финансовый вектор (ox)	Вектор конкурентоустойчивости (oz)	Оценка степени устойчивости (OR)
$K_э > 1$ $J_{эзд} > 1$ $TJ_{эзд} > 1$ $K_r > 1$	$ЗЗ < СОС$	Лидер отрасли (слабая конкуренция)	Абсолютная устойчивость
$K_э > 1$ $J_{эзд} > 1$ $TJ_{эзд} > 1$ $K_r > 1$	$ЗЗ = СОС + ЗК$	Одно из ведущих предприятий отрасли (слабая конкуренция)	Нормальная устойчивость
$K_э \geq 1$ $J_{эзд} \geq 1$ $TJ_{эзд} \geq 1$ $K_r < 1$	$ЗЗ = СОС + ЗК + ОИ$	Умеренная конкуренция	Неустойчивое состояние
$K_э < 1$ $J_{эзд} \leq 1$ $TJ_{эзд} \leq 1$ $K_r < 1$	$ЗЗ > СОС + ЗК$	Конкуренция – от умеренной до сильной	Кризисное положение
$K_э > 1$ $J_{эзд} \geq 1$ $TJ_{эзд} \geq 1$ $K_r \geq 1$	$ЗЗ \leq СОС + ЗК + ОИ$	Умеренная конкуренция	Метастабильное развитие

– неустойчивое финансовое положение предприятия возникает при нарушениях его платежеспособности. Однако остается возможность восстановления финансового равновесия за счет пополнения источников СОС и других источников, ослабляющих финансовую напряженность (ОИ): $ЗЗ = СОС + ЗК + ОИ$. Такими источниками (ОИ) могут быть различные фонды, финансовые резервы, кредиты и прочие заемные средства. Вполне понятно, что предприятие, которое наряду с собственными средствами использует заемный капитал, испытывает определенный предпринимательский и финансовый риск. Считается, что если на долю заемного капитала приходится 30–40%

всех активов предприятия, а собственные средства составляют соответственно 60–70%, то предприятию обеспечивается максимальная доходность;

– кризисное финансовое состояние наступает тогда, когда краткосрочные ценные бумаги и дебиторская задолженность не покрывают даже кредиторской задолженности и просроченных ссуд: $ЗЗ > СОС + ЗК$. В этой позиции предприятие находится на грани банкротства.

Финансовый вектор в нашей модели определяется в трехмерном пространстве следующим образом: $ЗЗ \leq СОС + ЗК + ОИ$. При этих условиях предприятие торговли способно длительное время поддерживать свой гомеостаз и обеспечивать свой гомеокинез, находясь в метастабильном состоянии. В обозначенном формате торговая бизнес-структура имеет гибкие возможности оперативного варьирования финансовыми ресурсами, сохраняя свою устойчивость (см. таблицу 1).

Давая рейтинговую оценку предприятию (вектор конкурентоспособности), необходимо отметить, что современная экономическая наука рассматривает устойчивое развитие предприятия не только как результат адаптации к изменяющимся условиям окружающей среды, но и как способность самого предприятия активно влиять на формирование своего бизнес-пространства и вырабатывать новые принципы взаимодействия с окружающей средой (гибкая ассортиментная политика; изменение хозяйственного профиля и организационно-управленческого звена, обеспечивающего эффективность принимаемых и реализуемых управленческих решений; диверсификация и реструктуризация предприятия; оптимизация структуры основных производственных фондов; создание различных корпоративных структур, объединений и т.д.), определяя место торговой бизнес-структуры среди предприятий-конкурентов. Интенсивность конкуренции в данном рыночном сегменте адекватно определяется по методу отраслевого анализа конкуренции М. Портера [6].

Результирующий вектор параметрической устойчивости развития (OR) соответствует в нашей модели метастабильному развитию предприятия, при котором данная ЭФС может находиться длительное время в термодинамически неустойчивом равновесии с окружающей средой, когда возможны и ожидаемы бифуркационные квантовые переходы системы на более высокий энергетический уровень

(безызлучательные квантовые переходы) или на уровень энергетически более низкий (излучательные квантовые переходы) [5]. Предприятие торговли, рассматриваемое нами как ЭФС, проходит свой онтогенез не в детерминистичной форме, а существует скорее как набор вероятностей в каждой фазе своего жизненного цикла, что и отражает предложенная трехвекторная модель. Метастабильное состояние предприятия торговли, по нашим данным, является энергетически более предпочтительным, по сравнению с другими (см. таблицу 1), поскольку характеризуется высоким экстенсивным темпом развития, относительно высокой степенью параметрической устойчивости и соответствует одному из минимумов термодинамического потенциала данной бизнес-структуры. Адаптированные к условиям рынка самоорганизация и самопрограммирование торговой бизнес-структуры осуществляются за счет набора стратегий поведения, укладывающихся в рамки предложенной модели. Оценочная функция оптимальной стратегии $F(t)$ будет в каждый момент времени определяться двумя слагаемыми. Первое из которых в виде функции F_1 характеризует параметрическое отклонение состояния системы x_n в момент t_1 , имевшей начальное состояние x_0 , от заданного конечного состояния x_1 в результате стратегии $n(t)$. Второе слагаемое является интегральной характеристикой изменения рассматриваемой системы от начального момента времени t_0 до конечного момента t_1 и выражает параметрическую устойчивость развития (надежность функционирования) торгового предприятия. В общем плане оценочная функция метастабильного поведения торговой системы имеет вид

$$F(t) = F_1 [x_n (t_1, x_0, t_0) t_1] + \int_{t_0}^{t_1} F_2 [x_n (t_1, x_0, t_0), n(t), t] dt.$$

В зависимости от критериев эффективности поведения минимизация одиночной функции может выражаться уменьшением или увеличением параметрических вариантов. Например, оптимальность стратегии предприятия торговли на данном бизнес-пространстве на основании трехвекторной модели, предложенной нами, будет определяться значениями описанных векторов, и в самом общем случае, при установившейся конкуренции – динамикой прибыли (финансовый вектор) и рентабельности (экономический вектор). В соответствии с этим, если прирост товарооборота будет ниже темпа прироста прибыли, то уровень рентабельности предприятия растет. Если

же темп прироста прибыли выше темпа прироста издержек обращения, то прирост прибыли обеспечивается за счет опережающего роста суммы валового дохода по сравнению с издержками обращения. В случае, когда темп прироста товарооборота превышает темп прироста прибыли, уровень рентабельности фирмы падает. И, наконец, когда темп прироста валового дохода равен темпу прироста издержек обращения, масса прибыли увеличивается [3].

Литература

1. Андрианов В.Д. Эволюция основных концепций регулирования экономики от теории меркантилизма до теории саморегуляции. М.: Экономика, 2008. 326 с.
2. Шишкин В.В. Параметрическая устойчивость, надежность и гомеостаз экономических функциональных систем // Международный технико-экономический журнал, 2010. № 3. С. 55–57.
3. Экономика, анализ и планирование на предприятии торговли / Под ред. А. Н. Соломатина. СПб.: Питер, 2009. 555 с.
4. Eungab K. Stochastic vender management replenishment with demand depend shipment // Eur. J. Operational Res., 2004. Vol. 152, № 3. P. 37–45.
5. Meyer P. Quantum probability for probabilists // Lecture notes in Mathem. Berlin, 1993. 310 p.
6. Porter M., Teisberg E. Redefining health care: creating value-based competition on results. Boston: Harvard Business School Press, 2006. 506 p.
7. Puu T. Nonlinear Economic Dynamics. Berlin: Verlag, 1997. 275 p.

Kostyunin S.¹, Palestini A.², Shevkoplyas E.¹

¹ Saint-Petersburg State University

² MEMOTEF, University of Rome "La Sapienza"

Differential game of resource extraction with random time horizon and different hazard functions

Consider a 2-firm differential game $\Gamma(x_0)$ with the following features:

- the game is played simultaneously;
- the i -th player starts extraction at the initial time $t_0 = 0$ and finishes at a random terminal time T_i ;
- the random variables denoting the terminal instants T_1 and T_2 are independent and both their probability distribution functions are known;
- $T = \min\{T_1, T_2\}$ is the terminal instant for the game;
- the strategic variable of the i -th agent is his rate of extraction $u_i \geq 0$;
- the state vector of the problem is $\mathbf{x} = (x_1, \dots, x_m) \in \mathbb{R}_+^m$, where x_i represents the stock of the i -th resource extracted by players.

The i -th player is endowed with a profit function $h_i(\mathbf{x}, u_1, u_2, t) \in C^2(\mathbb{R}_+^m \times \mathbb{R}_+^2 \times (0, +\infty))$, and is supposed to maximize the expected payoff flow:

$$\max_{u_i \geq 0} \mathbb{E} \left(\int_0^T h_i(\mathbf{x}, u_1, u_2, t) dt \right),$$

where $\mathbb{E}(\cdot)$ denotes the expected value of the integral with random terminal time T .

The dynamic constraints of the game are given by the following ODE system:

$$\begin{cases} \dot{\mathbf{x}}(t) = G(u_1(t), u_2(t), t), \\ \mathbf{x}(0) = \mathbf{x}_0, \end{cases}$$

where the coordinates $G_1(\cdot), \dots, G_m(\cdot) \in C^2(\mathbb{R}_+^2 \times \mathbb{R}_+)$ of the transition vector incorporate the shrinking effect of extraction on the resource stocks, that is:

$$\frac{\partial G_j}{\partial u_1} < 0, \quad \frac{\partial G_j}{\partial u_2} < 0, \quad j = 1, \dots, m.$$

Before delving into the problem, we need to point out some characteristics of the random variables at hand. At first, we are going to determine the probability distribution function and subsequently the density of the terminal instant of the game.

Proposition 1. *Given the probability distribution functions $F_1(t)$ and $F_2(t)$ of the the random terminal instants T_1 and T_2 , the probability distribution function $F(t)$ of $T = \min\{T_1, T_2\}$ reads as follows:*

$$F(t) = F_1(t) + F_2(t) - F_1(t)F_2(t).$$

Proof. Since $F_j(t) = \text{Prob}\{T_j < t\}$ for $j = 1, 2$, then by independence between T_1 and T_2 we have:

$$\begin{aligned} F(t) &= \text{Prob}\{T < t\} = 1 - \text{Prob}\{T \geq t\} = 1 - \text{Prob}\{\min\{T_1, T_2\} \geq t\} = \\ &= 1 - \text{Prob}\{T_1 \geq t\}\text{Prob}\{T_2 \geq t\} = \\ &= 1 - (1 - \text{Prob}\{T_1 < t\})(1 - \text{Prob}\{T_2 < t\}) = \\ &= 1 - (1 - F_1(t))(1 - F_2(t)) = F_1(t) + F_2(t) - F_1(t)F_2(t). \end{aligned}$$

Proposition 2. *If T_1 and T_2 respectively have densities $f_1(t)$ and $f_2(t)$, then*

$$f(t) = f_1(t)(1 - F_2(t)) + f_2(t)(1 - F_1(t))$$

is the density of T .

Proof. By applying Proposition 1 and the definition of density, we obtain:

$$\begin{aligned} f(t) &= F'(t) = f_1(t) + f_2(t) - (F_1(t)F_2(t))' = \\ &= f_1(t)(1 - F_2(t)) + f_2(t)(1 - F_1(t)). \end{aligned}$$

If we call

$$\lambda_i(t) = \frac{f_i(t)}{1 - F_i(t)}$$

the **hazard function** of the i -th player (also known as the force of mortality in actuarial science and survival analysis), it is straightforward to achieve the hazard function associated to the game as the sum of the hazard functions of the two players.

Proposition 3. $\lambda(t) = \lambda_1(t) + \lambda_2(t)$.

Proof. Employing Propositions 1 and 2 in the definition of hazard function yields:

$$\begin{aligned}\lambda(t) &= \frac{f_1(t)(1 - F_2(t)) + f_2(t)(1 - F_1(t))}{1 - [1 - (1 - F_1(t))(1 - F_2(t))]} = \\ &= \frac{f_1(t)}{1 - F_1(t)} + \frac{f_2(t)}{1 - F_2(t)} = \lambda_1(t) + \lambda_2(t).\end{aligned}$$

By integrating both sides in the definition of $\lambda(t)$ we achieve the following identity:

$$1 - F(t) = e^{-\int_0^t \lambda(s) ds} = e^{-\int_0^t \lambda_1(s) ds - \int_0^t \lambda_2(s) ds}.$$

Hence, the expected value of the payoff flow for the i -th firm on the horizon $[0, T]$ is:

$$\begin{aligned}\mathbb{E} \left(\int_0^T h_i(\mathbf{x}, u_1, u_2, t) dt \right) &= \\ &= \int_0^\infty \left(\int_0^t h_i(\mathbf{x}, u_1, u_2, \tau) d\tau \right) f(t) dt,\end{aligned}$$

and after rearranging the integrals by Fubini–Tonelli’s theorem, that amounts to

$$\begin{aligned}\int_0^\infty h_i(\mathbf{x}, u_1, u_2, t)(1 - F(t)) dt &= \\ &= \int_0^\infty h_i(\mathbf{x}, u_1, u_2, t) e^{-\int_0^t \lambda(s) ds} dt.\end{aligned}$$

Furthermore, if we consider any subgame starting in a subsequent instant $t > 0$, we have to take into account the occurrence that such game might not start at all, namely if $0 < T < t$. The related conditional probability can be expressed by dividing the payoff integral by the probability that $T > t$, so that the maximization problem for the i -th agent becomes:

$$\max_{u_i \geq 0} \frac{1}{1 - F(t)} \int_t^\infty h_i(\mathbf{x}, u_1, u_2, \tau)(1 - F(\tau)) d\tau. \quad (1)$$

From now on, consider the initial condition $x(t) = x > 0$ for the kinematic equation of (1) and call $W_i(\mathbf{x}, t)$ the i -th optimal value function

for this problem. After rearranging terms, the standard Hamilton–Jacobi–Bellman system reads as:

$$\frac{f(t)}{1 - F(t)} W_i(\mathbf{x}, t) = \frac{\partial W_i(\mathbf{x}, t)}{\partial t} + \max_{u_i \geq 0} \left(h_i(\mathbf{x}, u_1, u_2, t) + \sum_{j=1}^m \frac{\partial W_i(\mathbf{x}, t)}{\partial x_j} G_j(u_1, u_2, t) \right), \quad (2)$$

for $i = 1, 2$.

Example. Consider the following framework, borrowed from Rubio ([3], Example 2.1) and Dockner *et al.* ([1], Example 5.7) and modified with the above discount factor. This example originally describes the joint exploitation of a pesticide, but its peculiar structure makes it suitable for our aim. Note that, differently from Rubio [3], we confine our attention to the Nash equilibrium under simultaneous play, and we consider non-stationary feedback equilibrium structure, that is our optimal value function explicitly depends on the initial instant t .

We fix $m = 1$, i.e. a unique state variable $x(t)$, denoting the amount of the resource, whereas the i -th payoff function explicitly depends on the i -th rate of extraction and on x . This approach can be looked upon as a sort of endogenization of the variable T : firms cannot precisely know when the final instant comes, but as they get informed that they don't get any payoff any longer, the extraction stops and the first one who finishes extracting gets the reward. The payoff function for the i -th participant reads as:

$$h_i(\mathbf{x}, u_i(t)) = [u_i(t)x(t)]^\alpha,$$

where α is a positive real constant such that $0 < \alpha < \frac{1}{2}$. The reason for the choice of this interval relies on the positivity of the optimal feedback strategy and will be clearer at the end of the paper.

The transition function is linear and decreasing in the controls, so the dynamic constraint is:

$$\begin{cases} \dot{x} = -u_1 - u_2, \\ x(t) = x. \end{cases}$$

We are going to explicitly determine the optimal strategy in the feedback Nash structure, guessing the following ansatz for the solution of (2):

$$W_i(x, t) = A_i(t)x^{2\alpha} + B_i(t),$$

where $A_i(t)$ and $B_i(t)$ are unknown functions of t , such that the following limits are satisfied:

$$\lim_{t \rightarrow +\infty} A_i(t) = 0, \quad \lim_{t \rightarrow +\infty} B_i(t) = 0. \quad (3)$$

Application of the first order conditions in the r.h.s. of (2) yields:

$$\frac{\partial W_i(x, t)}{\partial x} = \alpha x^\alpha u_i^{\alpha-1} \iff 2\alpha A_i(t)x^{2\alpha-1} = \alpha x^\alpha u_i^{\alpha-1},$$

implying the following form for the feedback strategy:

$$u_i^*(x, t) = x (2A_i(t))^{\frac{1}{\alpha-1}}. \quad (4)$$

Hence, we need to exactly determine the expression for $A_i(t)$ by solving (2), whereas the expression of $B_i(t)$ is not necessary, as far as $u_i^*(\cdot)$ is concerned.

Assuming symmetry between players along the game, we can explicitly reckon a symmetric solution, so we posit: $A_1(t) = A_2(t) = A(t)$. Plugging $W_i(x, t)$ into (2), we get the unique ODE:

$$\dot{A}(t) = (2\alpha - 1) (2A(t))^{\frac{\alpha}{\alpha-1}} + \frac{f(t)}{1 - F(t)} A(t), \quad (5)$$

to be solved together with (3).

After some calculations, the solution of (5) amounts to:

$$A(t) = \frac{\left(C + 2^{\frac{\alpha}{\alpha-1}} \frac{2\alpha - 1}{1 - \alpha} \int_0^t (1 - F(s))^{\frac{1}{1-\alpha}} ds \right)^{1-\alpha}}{1 - F(t)}.$$

By (3), we achieve the value of C :

$$C = -2^{\frac{\alpha}{\alpha-1}} \frac{2\alpha - 1}{1 - \alpha} \int_0^\infty (1 - F(s))^{\frac{1}{1-\alpha}} ds,$$

therefore we can write down the optimal value function

$$W(x, t) = \frac{\left[2^{\frac{\alpha}{\alpha-1}} \frac{1-2\alpha}{1-\alpha} \left(\int_t^\infty (1-F(s))^{\frac{1}{1-\alpha}} ds \right) \right]^{1-\alpha} x^{2\alpha}}{1-F(t)},$$

and subsequently, by substitution in (4) one could acquire:

$$u^*(x, t) = \frac{1-\alpha}{1-2\alpha} \cdot \frac{x}{2 \left[\int_t^\infty (1-F(s))^{\frac{1}{1-\alpha}} ds \right] (1-F(t))^{\frac{1}{\alpha-1}}}.$$

Thus, the feedback Nash equilibrium of the game under symmetry assumption was obtained.

References

1. Dockner E., Jørgensen S., Van Long N., Sorger G. Differential games in economics and management science. Cambridge University Press, 2000. 394 p.
2. Petrosjan L.A., Shevkoplyas E.V. Cooperative solutions for games with random duration // Game Theory and Applications, 2003. Vol. 9. P. 125–139.
3. Rubio S. On coincidence of feedback Nash Equilibria and Stackelberg Equilibria in economic applications of differential games // Journal of Optimization Theory and Applications, 2006. Vol. 128, № 1. P. 203–221.
4. Sethi S.P., Thompson G.L. Optimal control theory: applications to management science and economics. Springer, 2000. 521 p.
5. Van der Ploeg F. Voracious transformation of a common natural resource into productive capital // International Economic Review, 2010. Vol. 51, № 2. P. 365–381.

Contents

1. Mathematical theory of control processes	9
<i>Dolgopolik M.V.</i> Codifferentialble functions in normed spaces	9
<i>Zhabko A.P., Zaranik U.P.</i> Domain of asymptotic stability of controlled systems with delay	15
<i>Kazantceva T.E.</i> Behavior of a certain dynamical system with two parameters	18
<i>Kuznetsova A.S.</i> Synchronization in coupled relay relaxation oscillators	24
<i>Lyubich S.Ya.</i> Mathematical model and algorithm of optimal planning of train schedule	29
<i>Medvedeva I.V.</i> Modification of algebraic method of differential-difference equations stability analysis	35
<i>Morozova D.N.</i> Synthesis of linear controls with linear delay	41
<i>Nguyen D.H.</i> On asymptotic stability of a class of nonlinear non-autonomous difference systems	47
<i>Ponomarev A.A.</i> On the parametric tuning of model predictive control	52
<i>Rekhlyasova L.V.</i> Calculation of the \mathcal{H}_2 norm for time delay systems	58
<i>Sukach M.P.</i> The example of two-level optimization problem solving	63
<i>Sumacheva V.A.</i> Computing of time-delay equation transfer function norm by means of the Lyapunov functions	68
<i>Uteshev N.A.</i> Solving identification problem via separating sphera construction	74
2. Mathematical methods in mechanics and physics	81
<i>Altsybeyev V.V.</i> Parametric beam dynamic optimization in hybrid systems	81
<i>Basinskiy K.Yu.</i> Distribution of capillary-gravity waves on free surface of low viscous fluids	86
<i>Belova O.A.</i> On a certain set of impact trajectories	92
<i>Bolshakov A.I.</i> Algorithm of a step size selection in dense output methods	98

<i>Bregman K.M.</i> Reduction of differential equations to polynomial system.....	103
<i>Vikulina Yu.I., Grekov M.A.</i> Stress-strain state of a half-plane in view of surface stress.....	109
<i>Vitvitskiy S.I., Nesterov A.V., Nesterov S.V.</i> Analysis of stable solutions of the Mathieu equation in modeling of periodically non-stationary control object.....	115
<i>Dron V.O.</i> On minimal splines of the third order on closed interval.....	121
<i>Ivanov A.N.</i> Symbolic computing in beam dynamic modeling....	127
<i>Ivanov V.A.</i> Two-component plane with curvilinear crack near interface boundary.....	133
<i>Klimakov A.A.</i> Electrostatic lens modeling.....	139
<i>Kodrik V.V., Kutuzov S.A.</i> Modeling of the Earth outer field via gravitating needles.....	145
<i>Kotov P.A.</i> On dynamic errors in dissipative dynamical systems.	151
<i>Kurushina S.E., Zhelnov Yu.V., Maximov V.V.</i> Phase transition «disorder - order - disorder» in stochastic reaction-diffusion systems.....	158
<i>Mikheev V.S., Mikheev S.E.</i> Interpolation on line table.....	165
<i>Podzyvalov E.A.</i> Globally optimal method of matching channel structure searching.....	171
<i>Prokopenko A.V.</i> Trajectories of stars in galaxy model allowing three integrals of motion.....	177
<i>Pronin A.O.</i> On cavitation instability of nonlinearly elastic solids.....	183
<i>Stepanova V.A.</i> Nonlinear problem for plane with straight interfacial crack.....	189
<i>Televny D.S.</i> Modeling of triode system with a field tip.....	195
<i>Ulanov E.A., Uteshev A. Yu.</i> Analytical solution for the generalized Fermat-Torricelli-Steiner problem.....	201
<i>Tselishev S.O., Eremin A.S.</i> On stability of structural method of ODE systems solving.....	207
<i>Shymanchuk D.V., Belova O.A.</i> Stabilization of orbital movement of a spacecraft in neighborhood of the collinear libration point...	213
<i>Shirokolobov A. Yu.</i> Mathematical optimization model of parameters of travelling-wave accelerator in case of smooth controls.....	220
<i>Yarish G.V., Pronina Yu.G.</i> Mechano-chemical corrosion of linear part of a pipeline.....	226

3. Mathematical models of medical and biological systems	233
<i>Abdrafiqov V.A.</i> Analysis of the morphology of red blood cells ...	233
<i>Kuznetsov S.R.</i> Spatial mathematical model of disbacteriosis development on human mucous membranes	239
<i>Mamochev V.A.</i> On application of the statistical method in classification problems	245
<i>Morozov P.D.</i> Modeling of basilar membrane oscillations based on Békésy's theory	249
<i>Morozov P.D., Miheev V.S.</i> Acoustic modulation of speech	255
<i>Cherneysanu E.K.</i> Strict <i>h</i> -separability and linear programming	259
4. Information and computer technologies	265
<i>Antipin I.A., Kozynchenko V.A.</i> On a modification of the Hamming neural network	265
<i>Borisov E.A.</i> Application of the optical Fourier transform in the problem of determining of orientation of the accident located text on image relatively to the lens aperture.....	271
<i>Bykov K.N.</i> Gauge invariance ratios for Hermite splines in a different order of nodes discarding.....	277
<i>Guznischeva A.P.</i> Program realization of non-linear object positional optimization.....	283
<i>Drokin I.S.</i> Computer analysis of the robust properties of superstable closed systems	289
<i>Durnovtseva S.A.</i> Modeling of artificial seismic vibrations of soil for design requirements	295
<i>Zhuravlev M.M.</i> Application of the minimum description length principle to the problem of mechanical destructions recognition on image.....	299
<i>Ivanova O.A., Peshekhonov K.A.</i> Implementation of elliptical functions computation method	304
<i>Lobov A.L.</i> Computer algorithms for identification of sea ship models.....	309
<i>Mikheyev P.A.</i> Individual speed of wireless LAN 802.11 station user.....	315
<i>Nwohiri A.M.</i> Ranking of the websites of Nigerian universities ...	321
<i>Pogozhev S.V., Sharlay A.S.</i> On displayed mathematical expressions extraction from document images	327

<i>Proskura E.S.</i> PKCS#12 file getting automation for MS Exchange mobile client	333
<i>Romanovsky L.M.</i> Local triangulation enlargement and gauge ratios	338
<i>Ryabov A.V., Osminin L.A.</i> Automation of the normative trains timetable development process	344
<i>Safonova A.V.</i> Machine learning approach using the LARS and Lasso to the Search Result Ranking problem	350
<i>Sevostyanov R.A.</i> Software for the motion control of the Futaba RBT-1 robot	356
<i>Smirnov M.N.</i> The algorithm for control synthesis that compensate bounded external disturbances act on a sea-going ship	362
<i>Fedorova M.A.</i> Synthesis and computer modeling of the non-static control system of motion control of sea-going ships	368
<i>Khakimyanova Yu.A.</i> BDD application to software development of trading systems	374
<i>Tsoy O.S.</i> Discretizing algorithm for ship routing in consideration of natural barriers	380
<i>Yakushkin O.O.</i> Artificial neural network for line segments detection	385
<i>Li M.H., Klyuev V.V., Wu S.H.</i> Sentence Alignment from Wikipedia as Comparable Corpora by using STF-IDTF	390
<i>Tikhomirov V.A.</i> Spectral image compression using different clustering methods	396
<i>Ueno R., Klyuev V.</i> Semantic Search Engine Query Expansion Using WordNet	402

5. Control problems in social and economical systems .. 409

<i>Andrienko V.A.</i> On a certain optimization problem in the interindustry balance model	409
<i>Afanasyeva E.V.</i> Nonlinear probabilistic chain-based modeling of the distribution of researchers across fields of science	416
<i>Barchenko M.K.</i> Non-Cramer model of insurance company ruin ..	422
<i>Baskov O.V.</i> Sequential algorithm for dual cone construction and its application to the multicriterion choice problem	427
<i>Bogdanova O.K.</i> Mathematical modeling of new product diffusion in social networks	432
<i>Brovkina L.A.</i> Cargo transportation traffic forecast model	436

<i>Bure V.M., Sergeeva A.A.</i> Optimal strategies of logistic company choice	440
<i>Gorbati M.V.</i> Dynamic model of licensing in clothing market	444
<i>Grigorova T.V.</i> Optimization of coalition structures in tree-like hierarchical games	450
<i>Grigoryeva K.V., Redinskikh N.D.</i> Absolute balance in a certain game	456
<i>Gromov I.A., Slobozhanin N.M.</i> On distribution of five-parameter resources	461
<i>Zakharov A.O.</i> Reduction of the Pareto set based on fuzzy closed information	467
<i>Zemtsova V.N.</i> Fuzzy sum – difference choice function	473
<i>Zenovich O.S., Malafeev O.A.</i> Optimization model of investor’s portfolio forming on the stock market	479
<i>Zubareva M.L., Gubar E.A.</i> Application of the Clarke & Wright method to construction of cash messenger routes in the ATM network	485
<i>Ivanchenko E.V., Kolbin V.V.</i> Peculiarities of pricing models of European and American options types in conditions of fuzzy data	491
<i>Iljina A.V., Petrosjan L.A.</i> The D.W.K. Yeung’s Condition for Dynamic Games with Nontransferable Payoffs	496
<i>Kantorovich A.V.</i> Game-theoretic model of interaction and cooperation in the supply chain	502
<i>Karpov M.I., Petrosjan L.A.</i> Costs minimization in the transport network	508
<i>Kumacheva S.Sh., Chugunnikova N.A.</i> Estimation and forecast of currency risk	514
<i>Mazalova A.V.</i> Braess’ paradox	519
<i>Melnikova E.A.</i> Game-theoretic model of the queue problem in service system	522
<i>Merzlyakova Yu.D., Gubar E.A.</i> Optimization of the ATM encashment process	528
<i>Nadezhdina A.M.</i> Queuing modeling of call-centers	533
<i>Orlov V.G.</i> Nash networks and the most efficient networks	539
<i>Petrov A.G., Slobozhanin N.M.</i> On potential function construction in the shortest-path problem	545
<i>Svishchikova M.V.</i> On a certain method of the Russian Federation GDP forecasting	550

<i>Seryakov I.A., Petrosjan L.A.</i> <i>N</i> -person transportation problem with non-overlapping paths	555
<i>Stepanova O.O.</i> Application of the Hamilton-Jacobi-Bellman equation and the method of lines in the investment processes ...	561
<i>Shishkin V.V.</i> Fractal model of the parametric stability evaluating of trade development	565
<i>Kostyunin S., Palestini A., Shevkoplyas E.</i> Differential game of resource extraction with random time horizon and different hazard functions	571
Contents	577
Author index	583

Именной указатель

Абдрафиков В.А.	233	Губар Е.А.	485, 528
Алцыбеев В.В.	81	Гузницева А.П.	283
Андриенко В.А.	409	Долгополик М.В.	9
Антипин И.А.	265	Дрокин И.С.	289
Афанасьева Е.В.	416	Дронь В.О.	121
Барченко М.К.	422	Дурновцева С.А.	295
Басинский К.Ю.	86	Еремин А.С.	207
Басков О.В.	427	Жабко А.П.	15
Белова О.А.	92, 213	Желнов Ю.В.	158
Богданова О.К.	432	Журавлев М.М.	299
Большаков А.И.	98	Зараник У.П.	15
Борисов Е.А.	271	Захаров А.О.	467
Бровкина Л.А.	436	Земцова В.Н.	473
Брэгман К.М.	103	Зенович О.С.	479
Буре В.М.	440	Зубарева М.Л.	485
Быков К.Н.	277	Иванов А.Н.	127
Викулина Ю.И.	109	Иванов В.А.	133
Витвицкий С.И.	115	Иванова О.А.	304
Горбати М.В.	444	Иванченко Е.В.	491
Греков М.А.	109	Ильина А.В.	496
Григорова Т.В.	450	Казанцева Т.Е.	18
Григорьева К.В.	456	Канторович А.В.	502
Громов И.А.	461	Карпов М.И.	508

Климаков А.А.	139	Надеждина А.М.	533
Кодрик В.В.	145	Нвохири А.М.	321
Козынченко В.А.	265	Нгуен Динь Хуен	47
Колбин В.В.	491	Нестеров А.В.	115
Котов П.А.	151	Нестеров С.В.	115
Кузнецов С.Р.	239	Орлов В.Г.	539
Кузнецова А.С.	24	Осьминин Л.А.	344
Кумачева С.Ш.	514	Петров А.Г.	545
Курушина С.Е.	158	Петросян Л.А.	496, 508, 555
Кутузов С.А.	145	Пешехонов К.А.	304
Лобов А.Л.	309	Погожев С.В.	327
Любич С.Я.	29	Подзывалов Е.А.	171
Мазалова А.В.	519	Пономарев А.А.	52
Максимов В.В.	158	Прокопенко А.В.	177
Малафеев О.А.	479	Пронин А.О.	183
Мамочев В.А.	245	Пронина Ю.Г.	226
Медведева И.В.	35	Проскура Е.С.	333
Мельникова Е.А.	522	Рединских Н.Д.	456
Мерзлякова Ю.Д.	528	Рехлясова Л.В.	58
Михеев В.С.	165, 255	Романовский Л.М.	338
Михеев П.А.	315	Рябов А.В.	344
Михеев С.Е.	165	Сафонова А.В.	350
Морозов П.Д.	249, 255	Свищикова М.В.	550
Морозова Д.Н.	41	Севостьянов Р.А.	356

Сергеева А.А.	440	Чернэуцану Е.К.	259
Серяков И.А.	555	Чугунникова Н.А.	514
Слобожанин Н.М.	461, 545	Шарлай А.С.	327
Смирнов М.Н.	362	Шиманчук Д.В.	213
Степанова В.А.	189	Широколов А.Ю.	220
Степанова О.О.	561	Шишкин В.В.	565
Сукач М.П.	63	Якушкин О.О.	385
Сумачева В.А.	68	Ярыш Г.В.	226
Телевный Д.С.	195	Klyuev V.V.	390, 402
Уланов Е.А.	201	Kostyunin S.	571
Утешев А.Ю.	201	Li M.H.	390
Утешев Н.А.	74	Palestini A.	571
Федорова М.А.	368	Shevkoplyas E.	571
Хакимьянова Ю.А.	374	Tikhomirov V.A.	396
Целищев С.О.	207	Ueno R.	402
Цой О.С.	380	Wu S.H.	390

Научное издание

ПРОЦЕССЫ УПРАВЛЕНИЯ И УСТОЙЧИВОСТЬ

Труды 42-й международной научной конференции
аспирантов и студентов

Печатается без издательского редактирования

Компьютерная верстка: *А. В. Згонников, М. Н. Смирнов,
Т. Е. Смирнова, М. А. Фёдорова*

Редактор текста на английском языке: *О. Н. Зиналиева*

Обложка художников: *С. Ю. Куццов, Г. А. Мартыненко*

Подписано в печать 02.06.11

Формат 60x84 1/16. Бумага офсетная. Печать офсетная.

Усл. печ. л. 34,2. Тираж 150 экз. Заказ № 148

Издательство СПбГУ. 199004, С.-Петербург, В.О. 6-я линия, 11/21

Тел. (812) 328-96-17, факс (812) 328-44-22

Типография Издательства СПбГУ
199061, С.-Петербург, Средний пр., д. 41