В. Д. НОГИН
И. О. ПРОТОДЬЯКОНОВ
И. И. ЕВЛАМПИЕВ

ОСНОВЫ ТЕОРИИ ОПТИМИЗАЦИИ

учебное пособие
для вузов

В. Д. НОГИН,
И. О. ПРОТОДЬЯКОНОВ,
И. И. ЕВЛАМПИЕВ

ОСНОВЫ ТЕОРИИ ОПТИМИЗАЦИИ

Под редакцией
И. О. Протодьяконова

Допущено Министерством высшего
и среднего специального образования СССР
в качестве учебного пособия
для студентов высших
технических учебных заведений

МОСКВА «ВЫСШАЯ ШКОЛА» 1986

ББК 22.11
Н72
УДК 51

Рецензенты: кафедра исследования операций факультета вычислительной математики и кибернетики Московского государственного университета им. М. В. Ломоносова и д-р физ.-мат. наук А. А. Петров.

Ногин В. Д., Протодьяконов И. О., Евлампиев И. И.
Н72 Основы теории оптимизации: Учеб. пособие для студентов втузов/под ред. И. О. Протодьяконова. — М.: Высш. шк., 1986. — 384 с., ил.

В пособии излагаются основные понятия и методы теории оптимизации. Рассматриваются задачи оптимизации в евклидовом и функциональных пространствах. Изучаются методы линейного, динамического, геометрического, нелинейного и многокритериального программирования, а также методы вариационного исчисления и оптимального управления системами.

Н $\frac{1502000000-543}{001(01)-86}$ КБ—28—19—86

ББК 22.11
517

Учебное издание

Владимир Дмитриевич Ногин
Игорь Орестович Протодьяконов
Игорь Иванович Евлампиев

ОСНОВЫ ТЕОРИИ ОПТИМИЗАЦИИ

Зав. редакцией Е. С. Гридасова. Редактор Ж. И. Яковлева. Мл. редакторы С. А. Доровских, Н. П. Майкова. Художественный редактор. В. И. Пономаренко. Технический редактор З. А. Муслимова. Корректор Г. И. Кострикова

ИБ № 4763

Изд. № ФМ—817. Сдано в набор 27.05.86. Подп. в печать 10.10.86. Т-18781. Формат $60\times90^1/_{16}$. Бум. кн.-журн. Гарнитура литературная. Печать высокая. Объем 24 усл. печ. л. 24 усл. кр.-отт. 23,74 уч.-изд. л. Тираж 40000 экз. Зак. № 339. Цена 1 р. 10 к.

Издательство «Высшая школа», 101430, Москва, ГСП-4, Неглинная ул., д. 29/14.

Московская типография № 8 Союзполиграфпрома при Государственном комитете СССР по делам издательств, полиграфии и книжной торговли. 101898, Москва, Центр, Хохловский пер., 7.

ПРЕДИСЛОВИЕ

В настоящее время теория оптимизации, успешному применению которой способствует современная вычислительная техника, вносит заметный вклад в ускорение научно-технического прогресса. Трудно назвать такую область инженерной деятельности, где бы не возникали задачи оптимизационного характера. Это, например, задачи определения наиболее эффективного режима работы различных технических систем, задачи организации производства, дающего наибольшую возможную прибыль при заданных ограниченных ресурсах, и др.

Учебное пособие включает основные разделы теории оптимизации (теории экстремальных задач), нашедшие применение при решении различных прикладных задач. В книге представлены *основы* теории оптимизации, т. е. такие понятия, теоремы и методы, знакомство с которыми не может дать исчерпывающего представления о всей теории, но которые позволят читателю войти в круг основных идей и методов решения прикладных задач и дадут возможность в дальнейшем ориентироваться в многообразии специальной литературы. Книга ориентирована на читателей, подготовленных в пределах программы по математике технического вуза.

Постановка каждой задачи оптимизации включает два объекта: множество допустимых решений и целевую функцию (функционал), которую следует минимизировать или максимизировать на указанном множестве. С этой общей точки зрения и рассматриваются различные классы экстремальных задач, составляющие предмет изучения линейного программирования, динамического программирования, нелинейного программирования, геометрического программирования, вариационного исчисления и теории оптимального управления.

Условно книгу можно разделить на две части. Первая часть (гл. 2—6) посвящены изучению и решению задач оптимизации, в которых целевая функция — это функция многих переменных, а допустимым множеством является подмножество евклидова пространства. В гл. 7 дается понятие о сравнительно новой ветви теории оптимизации — многокритериальной оптимизации, изучающей задачи с несколькими целевыми функциями. Во второй части (гл. 8—11) рассматриваются задачи оптимизации с допустимым множеством из функционального пространства, на котором минимизируется или максимизируется целевой функционал. При изложении этой части книги использован минимум дополнительных математических понятий, выходящих за рамки втузовской программы. Каждая глава, кроме первой, вспомогательной, заканчивается подробно разобранными примерами решения прикладных оптимизационных задач, что должно способствовать не только усвоению изложенного теоретического материала, но и развитию практических навыков решения конкретных прикладных задач.

Главы 1—11 (кроме § 9.7, 10.5, 11.3) написаны В. Д. Ногиным; § 9.7, 10.5 и 11.3 написаны И. О. Протодьяконовым и И. И. Евлампиевым.

Авторы выражают признательность С. А. Авдонину за полезное обсуждение материала книги. Кроме того, авторы благодарны В. В. Ананьеву, П. Н. Холодову и С. В. Чистякову, прочитавшим часть рукописи и высказавшим ряд ценных замечаний.

Авторы

Глава 1
ОСНОВНЫЕ МАТЕМАТИЧЕСКИЕ ПОНЯТИЯ

Цель данной главы — напомнить понятия и основные сведения из курса математического анализа, которые будут часто использоваться в дальнейшем. Более подробно с доказательствами формулируемых здесь утверждений можно ознакомиться в соответствующей учебной литературе (см., например, [16]).

В § 1.1 введены евклидово пространство и простейшие разновидности множеств в этом пространстве: прямая, отрезок, гиперплоскость. В § 1.2 сформулированы понятия открытого, замкнутого и компактного множеств, дано определение и приведен ряд примеров выпуклых множеств в евклидовом пространстве. Последний параграф посвящен функциям, определенным на множестве из евклидова пространства. Здесь напоминаются формулировки непрерывной, дифференцируемой, а также выпуклой функций. Приведены два обобщения понятия выпуклой функции, которые обычно не рассматривают в курсах математического анализа для втузов.

§ 1.1. ЕВКЛИДОВО ПРОСТРАНСТВО

1. Элементы теории множеств. Под *множеством* обычно понимают некоторый набор (совокупность) элементов произвольной природы. Примерами множеств могут служить совокупность страниц в данной книге, печатных знаков, формул. Множества вещественных, натуральных и целых чисел являются примерами числовых множеств.

Фразу «x является элементом множества X» («x принадлежит множеству X») записывают кратко в виде $x \in X$. Если x не принадлежит множеству X, то пишут $x \notin X$.

Множества X и Y называют *равными*, если они состоят из одних и тех же элементов.

Если каждый элемент множества X является элементом множества Y, то говорят, что X есть *подмножество* множества Y, и пишут $X \subset Y$. В частности, $X \subset X$.

Символы $\in$ и $\subset$ называют *знаками включения*, а соотношения вида $x \in X$ и $X \subset Y$ — *включениями*.

Равенство $X = Y$ имеет место тогда и только тогда, когда одновременно $X \subset Y$ и $Y \subset X$.

Запись $X = \{x_1, x_2, ...\}$ означает, что множество X состоит из элементов x_1, x_2 и, возможно, некоторых других, заданных тем или иным способом. Если множество X состоит из элементов x, обладающих определенным свойством $P(x)$, то пишут $X=\{x \mid P(x)\}$. Например, $(0, 1]=\{x \mid 0<x\leqslant 1\}$.

Для указания множества X, элементы которого принадлежат Y и, кроме того, обладают свойством $P(x)$, используют обозначение $X=\{x\in Y \mid P(x)\}$.

Множество, не содержащее ни одного элемента, называют *пустым множеством* и обозначают $\varnothing$.

Объединением двух множеств X, Y называют множество, обозначаемое $X\cup Y$, элементы которого принадлежат хотя бы одному из множеств X или Y. Операция объединения множеств обладает следующими свойствами:

1⁰) $X\cup Y=Y\cup X$,

2⁰) $(X\cup Y)\cup Z=X\cup(Y\cup Z)$,

3⁰) $X\cup X=X$, $X\cup\varnothing=X$,

где X, Y, Z — произвольные множества. Первые два свойства аналогичны свойствам операции сложения для чисел.

Пересечением множеств X, Y называют множество, обозначаемое $X\cap Y$, элементы которого принадлежат одновременно и множеству X, и множеству Y. Операция пересечения множеств обладает следующими свойствами:

1⁰) $X\cap Y=Y\cap X$,

2⁰) $(X\cap Y)\cap Z=X\cap(Y\cap Z)$,

3⁰) $X\cap X=X$, $X\cap\varnothing=\varnothing$.

Первые три свойства аналогичны соответствующим свойствам операции объединения.

Для введенных операций объединения и пересечения выполняются следующие свойства:

1⁰) $(X\cup Y)\cap Z=(X\cap Z)\cup(Y\cap Z)$,

2⁰) $(X\cap Y)\cup Z=(X\cup Z)\cap(Y\cup Z)$.

Разностью множеств X и Y называют множество, состоящее из элементов, которые принадлежат множеству X, но не принадлежат множеству Y (обозначение: $X\setminus Y$).

Рис. 1.1 иллюстрирует понятия объединения, пересечения и разности двух множеств.

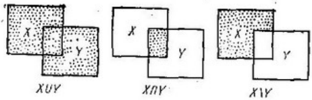

Рис. 1.1

Множество всех упорядоченных пар вида (x, y), где $x\in X$, $y\in Y$, называется *декартовым произведением множеств* X и Y и обозначается $X\times Y$. Так, например, $\{1, 2\}\times\{3, 4\}=\{(1,3), (1,4), (2,3), (2,4)\}$.

2. Евклидово пространство. Упорядоченный набор n ных чисел, записанных в виде столбца

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix},$$

называется *n-мерным вектором*. Числа $x_1, x_2, \ldots, x_n$ называют *компонентами* (или *координатами*) вектора x. Используя операцию транспонирования матриц, будем координаты вектора записывать и в виде строки:

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = (x_1, x_2, \ldots, x_n)^{\text{т}}.$$

Нередко для краткости, если это не приводит к неточности, знак транспонирования будем опускать.

Два n-мерных вектора x и y считают *равными*, если их соответствующие компоненты совпадают, т. е. если $x_i = y_i$, $i = 1, 2, \ldots, n$.

Суммой векторов x и y называют вектор, обозначаемый $x+y$, компоненты которого находят по формуле

$$x + y = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n). \tag{1.1}$$

Разность векторов x и y определяют следующим образом:

$$x - y = (x_1 - y_1, x_2 - y_2, \ldots, x_n - y_n). \tag{1.2}$$

Нулевым вектором (обозначение 0_n) называют n-мерный вектор, все компоненты которого равны нулю.

Произведением вектора x на вещественное число λ называют вектор, обозначаемый λx, компоненты которого находят по формуле

$$\lambda x = (\lambda x_1, \lambda x_2, \ldots, \lambda x_n). \tag{1.3}$$

Из приведенных определений непосредственно вытекает справедливость следующих равенств:

$$\begin{aligned} &x + y = y + x, && (x + y) + z = x + (y + z), \\ &x + 0_n = x, && 0x = 0_n, \\ &\lambda(x + y) = \lambda x + \lambda y, && (\lambda_1 + \lambda_2)x = \lambda_1 x + \lambda_2 x. \end{aligned}$$

Каждой паре векторов x, y сопоставим число, обозначаемое $\langle x, y \rangle$ и определяемое формулой

$$\langle x, y \rangle = \sum_{k=1}^{n} x_k y_k. \tag{1.4}$$

Это число называют *скалярным произведением* векторов x и y. Отметим свойства, которыми обладает скалярное произведение:

1°) $\langle x, y\rangle=\langle y, x\rangle$,

2°) $\langle x+y, z\rangle=\langle x, z\rangle+\langle y, z\rangle$,

3°) $\langle \lambda x, y\rangle=\lambda\langle x, y\rangle$,

4°) $\langle x, x\rangle\geqslant 0$ (равенство имеет место тогда и только тогда, когда x — нулевой вектор).

Справедливость приведенных свойств легко проверить, опираясь непосредственно на определение скалярного произведения.

Совокупность всех n-мерных векторов, для которых введены операции сложения, вычитания, умножение на вещественное число, а также скалярное произведение согласно формулам (1.1)—(1.4), называют *n-мерным вещественным евклидовым пространством* и обозначают $\mathbf{R}^n$. Для краткости это пространство будем называть просто *евклидовым пространством.*

Множество вещественных чисел далее будем обозначать буквой $\mathbf{R}$. Это множество является примером одномерного евклидова пространства. Пространство $\mathbf{R}^2$ будем геометрически интерпретировать как совокупность точек плоскости с фиксированной прямоугольной системой координат или же как совокупность векторов на этой плоскости, начало которых совпадает с началом прямоугольной декартовой системы координат. Пространство $\mathbf{R}^3$ имеет аналогичную интерпретацию в пространстве трех измерений.

Принимая во внимание указанную выше интерпретацию, элементы пространства $\mathbf{R}^n$ будем называть также *точками.*

Если $\langle x, y\rangle=0$, то векторы x и y называют *ортогональными.* В частности, нулевой вектор ортогонален любому вектору.

Нормой (или *длиной*) вектора $x\in\mathbf{R}^n$ называют число, обозначаемое $\|x\|$ и определяемое формулой

$$\|x\|=\sqrt{\langle x, x\rangle}=\sqrt{\sum_{k=1}^{n} x_k^2}.$$

Сформулируем основные свойства нормы вектора:

1°) $\|x\|\geqslant 0$, причем $\|x\|=0$ тогда и только тогда, когда $x=0_n$;

2°) $\|\lambda x\|=|\lambda|\,\|x\|$, где λ — число:

3°) $\|x+y\|\leqslant\|x\|+\|y\|$ (неравенство треугольника);

4°) $|\langle xy\rangle|\leqslant\|x\|\cdot\|y\|$ (неравенство Коши — Буняковского).

Расстояние между точками x и y евклидова пространства обозначают через $\rho(x, y)$ и определяют следующим образом:

$$\rho(x, y)=\|x-y\|=\sqrt{\sum_{k=1}^{n}(x_k-y_k)^2}.$$

Систему векторов $x^{(1)}, x^{(2)}, \ldots, x^{(m)}$ называют *линейно независимой системой*, если равенство $\lambda_1x^{(1)}+\lambda_2x^{(2)}+\ldots+\lambda_mx^{(m)}=0_n$ возможно лишь в случае $\lambda_1=\lambda_2=...=\lambda_m=0$. Систему векторов, не являющуюся линейно независимой, называют *линейно зависимой системой.*

В n-мерном пространстве существует линейно независимая система из n-векторов, а любая система из $n+1$ (и более) векторов является линейно зависимой. Любая линейно независимая система $\{e^{(k)}\}^n_{k=1}$ векторов n-мерного пространства образует *базис*; при этом каждый вектор пространства x единственным образом представляется в виде линейной комбинации базисных векторов: $x=\lambda_1e^{(1)}+\lambda_2e^{(2)}+...+\lambda_ne^{(n)}$ при некоторых $\lambda_1, \lambda_2, ..., \lambda_n$.

Для векторов $x\in\mathbf{R}^n$, $y\in\mathbf{R}^n$ условимся считать, что запись $x\geqq y$ означает, что $x_i\geqslant y_i$, $i=1, 2, ..., n$, а запись $x>y$ означает, что $x_i>y_i$, $i=1, 2, ..., n$.

3. Линейные множества. Непустое подмножество L пространства $\mathbf{R}^n$ называется *подпространством*, если в результате сложения любых двух векторов из L, а также умножения произвольного вектора из L на любое вещественное число получаются векторы, принадлежащие L. Если максимальное число линейно независимых векторов, которые можно найти в L, равно r, то говорят, что L — *r-мерное подпространство.* Само пространство $\mathbf{R}^n$ можно рассматривать как n-мерное подпрострачство.

Множество точек вида $\{z\in\mathbf{R}^n | z=x^*+x$ при некотором $x\in L\}$, где x^* — некоторый фиксированный вектор из $\mathbf{R}^n$, а L — r-мерное подпространство, называют *r-мерным линейным множеством.* При $x^*=0_n$ это множество совпадает с подпространством L. Геометрически r-мерное линейное множество представляет собой сдвиг подпространства на фиксированный вектор x^*. Одномерное линейное множество называют *прямой,* двумерное — *плоскостью,* а $(n-1)$-мерное — *гиперплоскостью.*

Любую прямую в $\mathbf{R}^n$ можно задать в виде

$$\{x\in\mathbf{R}^n | x=a+\lambda c \text{ при некотором } \lambda\in\mathbf{R}\}, \tag{1.5}$$

соответствущим образом подобрав векторы a и c из $\mathbf{R}^n$. Если в (1.5) число λ ограничено сверху или снизу, то получаем *луч.* Если λ ограничено и сверху, и снизу, то множество (1.5) задает *отрезок. Отрезок, соединяющий точки x', $y'\in\mathbf{R}^n$,*— это множество вида

$$\{x\in\mathbf{R}^n | x=\lambda x'+(1-\lambda)y' \text{ при некотором } 0\leqslant\lambda\leqslant 1\}.$$

Множество решений $x\in\mathbf{R}^n$ уравнения

$$\langle a, x\rangle=b, \tag{1.6}$$

где вектор $a\neq 0_n$ и число b фиксированы, представляет собой некоторую гиперплоскость в $\mathbf{R}^n$. Справедливо и обратное: любую гиперплоскость в $\mathbf{R}^n$ можно задать в виде множества решений уравнения (1.6), подобрав соответствующим образом вектор a и число b.

§ 1.2. МНОЖЕСТВА В ЕВКЛИДОВОМ ПРОСТРАНСТВЕ

1. Открытые, замкнутые, компактные множества. Множество вида

$$U_\varepsilon(x^{(0)})=\{x\in\mathbf{R}^n \mid \|x-x^{(0)}\|<\varepsilon\}$$

называют *открытым шаром радиуса* ε *с центром в точке* $x^{(0)}\in\mathbf{R}^n$ или ε-*окрестностью точки* $x^{(0)}$.

Пусть $X\subset\mathbf{R}^n$ — некоторое множество. Точку $x\in X$ называют *внутренней точкой* множества X, если найдется такое $\varepsilon>0$, что $U_\varepsilon(x)\subset X$, т. е. если точка x принадлежит множеству X вместе со своей некоторой окрестностью. В том случае, когда каждая точка множества X является внутренней, это множество называют *открытым множеством*. Например, $U_\varepsilon(x^{(0)})$ — открытое множество.

Точку пространства $\mathbf{R}^n$ (не обязательно принадлежащую множеству X) называют *граничной точкой* множества $X\subset\mathbf{R}^n$, если в ее окрестности любого радиуса $\varepsilon>0$ имеется хотя бы одна точка из X и хотя бы одна точка, не принадлежащая X. Совокупность граничных точек множества образует его *границу*.

Последовательность $\{x^{(k)}\}_{k=1}^{\infty}$ точек из пространства $\mathbf{R}^n$ называют *сходящейся к точке* $x^{(0)}\in\mathbf{R}^n$, если $\lim\limits_{k\to\infty}\|x^{(k)}-x^{(0)}\|=0$; при этом пишут $\lim\limits_{k\to\infty}x^{(k)}=x^{(0)}$.

Множество $X\subset\mathbf{R}^n$ называют *замкнутым*, если из соотношений

$$\lim_{k\to\infty}x^{(k)}=x^{(0)}\in\mathbf{R}^n,\ x^{(k)}\in X,\ k=1, 2,\ldots,$$

следует, что $x^{(0)}\in X$. Другими словами, множество X замкнуто, если предельная точка произвольной сходящейся последовательности точек из X принадлежит множеству X. Отрезки, прямые, гиперплоскости в $\mathbf{R}^n$ служат примерами замкнутых множеств. Само пространство $\mathbf{R}^n$ является одновременно открытым и замкнутым множеством. Известно, что множество является замкнутым тогда и только тогда, когда оно содержит все свои граничные точки.

Объединение, а также пересечение конечного числа открытых (замкнутых) множеств представляет собой открытое (замкнутое) множество.

Множество X называют *ограниченным*, если существует такое $\varepsilon>0$, что $X\subset U_\varepsilon(0_n)$.

Замкнутое и ограниченное множество называют *компактным множеством* или *компактом*. Примеры компактов: конечное множество точек в $\mathbf{R}^n$; отрезок, соединяющий пару точек в $\mathbf{R}^n$; *замкнутый шар*, т. е. множество вида

$$\{x\in\mathbf{R}^n \mid \|x-x^{(0)}\|\leqslant\varepsilon\}.$$

Гиперплоскость не является компактом, так как для нее не выполняется требование ограниченности.

Известно, что декартово произведение компактных множеств само является компактом.

2. Выпуклые множества. Множество X называют *выпуклым*, если для любой пары точек из X весь отрезок, соединяющий эти точки, также принадлежит X. Отрезок, прямая, гиперплоскость, шар — примеры выпуклых множеств.

Каждая гиперплоскость, заданная уравнением (1.6), порождает пару множеств

$$\{x \in \mathbf{R}^n | \langle a, x\rangle \geqslant b\}, \ \{x \in \mathbf{R}^n | \langle a, x\rangle \leqslant b\},$$

которые называются *замкнутыми полупространствами* (рис. 1.2). Если в этом определении использованы строгие знаки неравенств, то получаем *открытые полупространства*. Все указанные полупространства являются выпуклыми множествами. Проверим, например, выпукло ли полупространство $\{x \in \mathrm{R}^n | \langle a, x\rangle \geqslant b\}$. Для этого рассмотрим две произвольные точки $x^{(1)}$ и $x^{(2)}$ этого полупространства. Для этих точек выполнены неравенства $\langle a, x^{(1)}\rangle \geqslant b$, $\langle a, x^{(2)}\rangle \geqslant b$. Сложим эти два неравенства, предварительно умножив первое на произвольное число $\lambda \in [0, 1]$, а второе — на $1-\lambda$. В результате получим неравенство

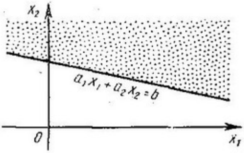

Рис. 1.2

$$\lambda \langle a, x^{(1)}\rangle + (1-\lambda)\langle a, x^{(2)}\rangle = \langle a, \lambda x^{(1)} + (1-\lambda) x^{(2)}\rangle \geqslant b.$$

Поскольку λ произвольно, весь отрезок, соединяющий выбранные точки, принадлежит данному полупространству. Следовательно, полупространство действительно является выпуклым множеством.

Можно проверить, что пересечение любого числа выпуклых множеств представляет собой выпуклое множество. Поэтому множество решений системы линейных неравенств и уравнений вида

$$\langle a^{(i)}, x\rangle \geqslant b_i, \quad i = 1, 2, \ldots, k_1,$$
$$\langle a^{(j)}, x\rangle \leqslant b_j, \quad j = k_1 + 1, \ldots, k_2,$$
$$\langle a^{(l)}, x\rangle = b_l, \quad l = k_2 + 1, \ldots, k_3,$$

является выпуклым множеством как пересечение замкнутых полупространств и гиперплоскостей.

§ 1.3. ФУНКЦИИ МНОГИХ ПЕРЕМЕННЫХ

1. Непрерывные, дифференцируемые функции. Пусть X и Y — два множества. Если указано правило, согласно которому каждому элементу множества X поставлен в соответствие определенный элемент множества Y, то говорят, что задана *функция* f, отображающая X в Y. Этот факт записывают в виде $f: X \to Y$ или $y = f(x)$, где $x \in X$, $y \in Y$. Множество X называют *областью задания* или *областью определения* функции f, а множество Y — *множеством значений*.

В первых семи главах настоящей книги в основном рассматривают функции многих переменных $y=f(x)=f(x_1, x_2, \ldots, x_n)$, т. е. функции с областью задания $X \subset \mathbf{R}^n$ и множеством значений $Y \subset \mathbf{R}$. Такие функции называют *числовыми функциями* в отличие от *векторных функций*, для которых $Y \subset \mathbf{R}^m$, $m > 1$. В дальнейшем числовые функции будем называть просто *функциями*.

Множество вида

$$\{(x, y) \in \mathbf{R}^{n+1} \mid y=f(x) \text{ при некотором } x \in X\}$$

называют *графиком функции* $y=f(x)$.

Функцию f называют *непрерывной в точке* $x^{(0)} \in X$, если для любого числа $\varepsilon > 0$ можно указать такое число $\delta_\varepsilon > 0$, что для всех $x \in X \cap U_{\delta_\varepsilon}(x^{(0)})$ выполняется неравенство $|f(x)-f(x^{(0)})| < \varepsilon$.

Заметим, что согласно данному определению функция f в изолированной точке всегда непрерывна. Напомним, что точка $x \in X$ называется *изолированной* точкой множества X, если существует такая ее окрестность, которая не содержит никаких других точек из X, кроме самой точки x.

Определение непрерывности функции можно сформулировать и на языке последовательностей. А именно: функция f непрерывна в точке $x^{(0)} \in X$ тогда и только тогда, когда для любой последовательности точек $\{x^{(k)}\}_{k=1}^{\infty}$, удовлетворяющих соотношениям $x^{(k)} \in X$, $\lim\limits_{k \to \infty} x^{(k)}=x^{(0)}$, имеет место равенство $\lim\limits_{k \to \infty} f(x^{(k)})=f(x^{(0)})$.

Функцию, непрерывную в каждой точке множества X, называют *непрерывной на множестве* X (или просто *непрерывной*, если $X=\mathbf{R}^n$).

В качестве примеров функций, непрерывных на $\mathbf{R}^n$, приведем *линейную функцию*

$$f_1(x)=\langle c, x\rangle+b=c_1x_1+c_2x_2+\ldots+c_nx_n+b$$

и *квадратичную функцию*

$$f_2(x)=\frac{1}{2}\langle Qx, x\rangle+\langle c, x\rangle+b,$$

где Q — числовая симметричная матрица размера $n \times n$, c — некоторый вектор из $\mathbf{R}^n$ и b — некоторое число, а Qx означает произведение матрицы на вектор по правилам перемножения матриц, принятым в линейной алгебре.

Пусть $X \subset \mathbf{R}^n$ — замкнутое множество и функция f непрерывна на нем. Тогда множество точек $x \in X$, удовлетворяющих неравенству $f(x) \leqslant 0$, является замкнутым множеством. В самом деле, пусть $x^{(k)} \in X$, $f(x^{(k)}) \leqslant 0$, $k=1, 2, \ldots$, $\lim\limits_{k \to \infty} x^{(k)}=x^{(0)} \in \mathbf{R}^n$. Множество X замкнуто, поэтому $x^{(0)} \in X$. А так как функция f непрерывна в точке $x^{(0)}$, то из $f(x^{(k)}) \leqslant 0$, $k=1, 2, \ldots$, следует неравенство $f(x^{(0)}) \leqslant 0$.

Пересечение замкнутых множеств само является замкнутым множеством. Поэтому множество вида $\{x \in X \mid g_j(x) \leqslant 0,\ j=1, 2, \ldots, k\}$ является замкнутым в том случае, если все функции $g_1(x)$, $g_2(x)$, ..., $g_k(x)$ непрерывны на замкнутом множестве X.

Пусть $x^{(0)}$ — внутренняя точка множества X. Функцию f называют *дифференцируемой в точке* $x^{(0)}$, если существует вектор $p \in \mathbf{R}^n$ такой, что для всех $h \in \mathbf{R}^n$, удовлетворяющих условию $(x^{(0)} + h) \in X$, имеет место формула

$$f(x^{(0)}+h) = f(x^{(0)}) + \langle p,\ h\rangle + \|h\|\,\alpha(x^{(0)},\ h), \tag{1.7}$$

где функция α обладает свойством

$$\lim_{\|h\|\to 0} \alpha(x^{(0)},\ h) = 0. \tag{1.8}$$

Если указанный вектор p существует, то его называют *градиентом функции* f *в точке* $x^{(0)}$ и обозначают $\nabla f(x^{(0)})$. Известно, что если функция f дифференцируема в точке $x^{(0)}$, то она в этой точке непрерывна и

$$\nabla f(x^{(0)}) = \left(\frac{\partial f(x^{(0)})}{\partial x_1},\quad \frac{\partial f(x^{(0)})}{\partial x_2}, \ldots, \frac{\partial f(x^{(0)})}{\partial x_n}\right)^{\mathrm{T}},$$

т. е. градиент представляет собой вектор частных производных первого порядка, вычисленных в точке $x^{(0)}$.

Функцию, дифференцируемую в каждой точке открытого множества X, называют *дифференцируемой на множестве* X (или просто *дифференцируемой*, если $X = \mathbf{R}^n$).

Если функция f имеет в некоторой точке x (или в каждой точке X) непрерывные частные производные первого порядка, то ее называют *непрерывно дифференцируемой* в этой точке (соответственно на множестве X). Известно [16], что функция, непрерывно дифференцируемая в точке, является также дифференцируемой в этой точке.

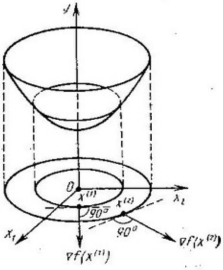
Рис. 1.3

Приведенные выше линейная и квадратичная функции непрерывно дифференцируемы, причем, как легко вычислить, $\nabla f_1(x) = c$, $\nabla f_2(x) = Qx + c$.

Градиент, как вектор пространства $\mathbf{R}^n$, обладает важным геометрическим свойством, которое проиллюстрируем на примере функции двух переменных $y = (x_1)^2 + (x_2)^2 + 1$. Графиком этой функции является параболоид вращения с вершиной в точке $(0,\ 0,\ 1)$ (рис. 1.3), а линиями уровня — окружности в плоскости $x_1 0 x_2$. Здесь градиент в точке $\left(x_1^{(0)},\ x_2^{(0)}\right)$ равен $\left(2x_1^{(0)},\ 2x_2^{(0)}\right)$ и перпендикулярен касательной к окружности, проходящей через точку $\left(x_1^{(0)},\ x_2^{(0)}\right)$. На рис. 1.3

градиенты изображены приложенными к точкам, в которых их вычисляют. Вообще градиент отличной от постоянной функции $y=f(x_1, x_2)$, вычисленный в точке $(x_1^{(0)}, x_2^{(0)})$, перпендикулярен касательной к линии уровня $\{x\in\mathbf{R}^2 \mid f(x_1, x_2)=f(x_1^{(0)}, x_2^{(0)})\}$, проходящей через точку $(x_1^{(0)}, x_2^{(0)})$. В случае функции трех переменных градиент ортогонален плоскости, проведенной через рассматриваемую точку и являющейся касательной к соответствующей поверхности уровня.

Пусть $x^{(0)}$ — внутренняя точка множества X. Функцию f называют *дважды дифференцируемой в точке* $x^{(0)}$, если существуют градиент $\nabla f(x^{(0)})$ и матрица чисел размера $n\times n$, обозначаемая $\nabla^2 f(x^{(0)})$, такие, что для всех векторов $h\in\mathbf{R}^n$, удовлетворяющих условию $(x^{(0)}+h)\in X$, имеет место формула

$$f(x^{(0)}+h)=f(x^{(0)})+\langle\nabla f(x^{(0)}), h\rangle+\frac{1}{2}\langle\nabla^2 f(x^{(0)})h, h\rangle+$$
$$+\|h\|^2\beta(x^{(0)}, h),$$

где функция β обладает следующим свойством:

$$\lim_{\|h\|\to 0}\beta(x^{(0)}, h)=0.$$

Здесь числовая матрица $\nabla^2 f(x^{(0)})$ называется *гессианом* (или *матрицей Гессе*). Если функция f имеет непрерывные частные производные второго порядка (*дважды непрерывно дифференцируема*) в точке $x^{(0)}$, то она дважды дифференцируема в $x^{(0)}$ и обладает матрицей Гессе вида

$$\nabla^2 f(x^{(0)})=\begin{pmatrix} \dfrac{\partial^2 f(x^{(0)})}{\partial x_1^2} & \dfrac{\partial^2 f(x^{(0)})}{\partial x_1\partial x_2} & \dots & \dfrac{\partial^2 f(x^{(0)})}{\partial x_1\partial x_n} \\ \dfrac{\partial^2 f(x^{(0)})}{\partial x_2\partial x_1} & \dfrac{\partial^2 f(x^{(0)})}{\partial x_2^2} & \dots & \dfrac{\partial^2 f(x^{(0)})}{\partial x_2\partial x_n} \\ \dots & \dots & \dots & \dots \\ \dfrac{\partial^2 f(x^{(0)})}{\partial x_n\partial x_1} & \dfrac{\partial^2 f(x^{(0)})}{\partial x_n\partial x_2} & \dots & \dfrac{\partial^2 f(x^{(0)})}{\partial x_n^2} \end{pmatrix},$$

причем эта матрица симметрична, т. е. $\dfrac{\partial^2 f(x^{(0)})}{\partial x_i\partial x_k}=\dfrac{\partial^2 f(x^{(0)})}{\partial x_k\partial x_i}$, $i\neq k$; $i, k=1, 2, \dots, n$.

Функцию, дважды дифференцируемую в каждой точке пространства $\mathbf{R}^n$, будем называть *дважды дифференцируемой*.

Квадратичная функция $f(x)=\frac{1}{2}\langle Qx, x\rangle+\langle c, x\rangle+b$ с симметричной матрицей Q дважды дифференцируема, причем, как легко найти, $\nabla^2 f(x)=Q$.

2. Выпуклые, псевдовыпуклые и квазивыпуклые функции. Выпуклые функции и их обобщения (псевдовыпуклые и квазивыпуклые функции) играют важную роль в теории оптимизации. В гл. 2 с помощью этих функций будут сформулированы достаточные условия оптимальности.

Числовую функцию f, определенную на выпуклом множестве X, $X \subset \mathbf{R}^n$, называют *выпуклой*, если для любых двух точек $x^{(1)}$, $x^{(2)} \in X$ и произвольного числа $\lambda \in [0, 1]$ выполняется неравенство

$$f(\lambda x^{(1)} + (1-\lambda) x^{(2)}) \leqslant \lambda f(x^{(1)}) + (1-\lambda) f(x^{(2)}).$$

Геометрически свойство выпуклости функции f означает, что линия ее графика, которая соответствует точкам отрезка, соединяющего точки $x^{(1)}$ и $x^{(2)}$, располагается не выше прямой, которая соединяет точки графика $(x^{(1)}, f(x^{(1)}))$ и $(x^{(2)}, f(x^{(2)}))$ (рис. 1.4).

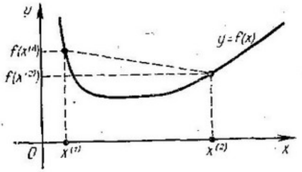

Рис. 1.4

Простейшими примерами выпуклых функций одной переменной служат парабола $y=x^2$ и экспонента $y=e^x$.

Если функция $-f$ выпуклая на множестве X, то саму функцию f называют *вогнутой*. Таким образом, функции $y=-x^2$ и $y=-e^x$ являются вогнутыми.

Справедливы [8] следующие утверждения.

Теорема 1.1. *Непрерывно дифференцируемая на выпуклом множестве X функция f выпукла на этом множестве тогда и только тогда, когда для любых $x^{(1)}$, $x^{(2)} \in X$ верно неравенство*

$$f(x^{(2)}) \geqslant f(x^{(1)}) + \langle \nabla f(x^{(1)}), x^{(2)} - x^{(1)} \rangle. \tag{1.9}$$

Примечание. Если множество X не является открытым, то в теореме 1.1 имеется в виду непрерывная дифференцируемость функции f на некотором открытом множестве, содержащем X. В этом же смысле далее в теореме 1.2 понимается дважды непрерывная дифференцируемость функции f на X.

Теорема 1.2. *Пусть функция f дважды непрерывно дифференцируема на выпуклом множестве X, содержащем хотя бы одну внутреннюю точку, и $\nabla^2 f(x)$ — ее гессиан. Тогда для выпуклости f на множестве X необходимо и достаточно, чтобы матрица $\nabla^2 f(x)$ была неотрицательно определена при всех $x \in X$, т. е. чтобы неравенство*

$$\langle \nabla^2 f(x) z, z \rangle \geqslant 0 \tag{1.10}$$

выполнялось для всех точек $x \in X$, $z \in \mathbf{R}^n$.

Аналогичные утверждения имеют место и для вогнутых функций. При этом в формулах (1.9) и (1.10) знак неравенства $\geqslant$ следует заменить на $\leqslant$.

Неотрицательная определенность матрицы M размера $n \times n$ по определению означает справедливость неравенства $\langle Mz, z \rangle \geqslant 0$ для всех векторов $z \in \mathbf{R}^n$. Матрица M называется *положительно-определенной*, если для всех векторов $z \in \mathbf{R}^n$, $z \neq 0_n$, верно неравенство $\langle Mz, z\rangle > 0$. Если здесь знаки неравенств $\geqslant$ и $>$ заменить соответственно на $\leqslant$ и $<$, то получим определения *неположительно-определенной* и *отрицательно-определенной* матрицы M.

Отметим некоторые важные свойства выпуклых функций, предполагая, что X — выпуклое множество.

1°. *Если числа $c_1, c_2, \ldots, c_m$ не отрицательны и функции $f_1, f_2, \ldots, f_m$ выпуклы на множестве X, то линейная комбинация, т. е. функция* $\sum\limits_{k=1}^{n} c_k f_k(x)$, *также выпукла на X.*

Убедимся в справедливости этого свойства для $m=2$. Рассмотрим две произвольные точки $x^{(1)}, x^{(2)} \in X$ и произвольное число $\lambda \in [0, 1]$. Функции f_1 и f_2 выпуклые, поэтому имеем

$$c_1 f_1(\lambda x^{(1)} + (1-\lambda) x^{(2)}) + c_2 f_2(\lambda x^{(1)} + (1-\lambda) x^{(2)}) \leqslant$$
$$\leqslant c_1(\lambda f_1(x^{(1)}) + (1-\lambda) f_1(x^{(2)})) + c_2(\lambda f_2(x^{(1)}) + (1-\lambda) f(x^{(2)})) =$$
$$= \lambda(c_1 f_1(x^{(1)}) + c_2 f_2(x^{(1)})) + (1-\lambda)(c_1 f_1(x^{(2)}) + c_2 f_2(x^{(2)})),$$

т. е. функция $c_1 f_1(x) + c_2 f_2(x)$ также выпукла.

2°. *Если функция f выпукла на множестве X, то для любого $\alpha \in \mathbf{R}$* множество вида

$$\{x \in X \mid f(x) \leqslant \alpha\} \tag{1.11}$$

является выпуклым.

В самом деле, рассмотрим произвольные две точки $x^{(1)}$ и $x^{(2)}$ множества (1.11), т. е. пусть выполнены неравенства $f(x^{(1)}) \leqslant \alpha$, $f(x^{(2)}) \leqslant \alpha$. Складывая почленно эти неравенства, предварительно умножив первое на $\lambda \in [0, 1]$, а второе — на $(1-\lambda)$, получим неравенство

$$\lambda f(x^{(1)}) + (1-\lambda) f(x^{(2)}) \leqslant \alpha.$$

Отсюда, используя выпуклость функции f, придем к неравенству

$$f(\lambda x^{(1)} + (1-\lambda) x^{(2)}) \leqslant \alpha.$$

Это означает, что точка $\lambda x^{(1)} + (1-\lambda) x^{(2)}$ принадлежит множеству (1.11) при любом $\lambda \in [0, 1]$, т. е. указанное множество действительно является выпуклым.

3°. *Если функции $f_1, f_2, \ldots, f_m$ выпуклы на множестве X, то функция вида*

$$f(x) = \max_{i=1,2,\ldots,m} f_i(x) = \max\{f_1(x), f_2(x), \ldots, f_m(x)\}$$

также является выпуклой на множестве X.

Действительно, в силу выпуклости функций $f_1, f_2, \ldots, f_m$ имеем

$$\max_{i=1,2,\ldots,m} f_i(\lambda x^{(1)} + (1-\lambda) x^{(2)}) \leqslant \max_{i=1,2,\ldots,m} [\lambda f_i(x^{(1)}) + (1-\lambda) f_i(x^{(2)})]$$

для произвольных $x^{(1)}, x^{(2)} \in X$, $\lambda \in [0, 1]$. Отсюда следует требуемое неравенство

$$\max_{i=1,2,\ldots,m} f_i(\lambda x^{(1)} + (1-\lambda) x^{(2)}) \leqslant \lambda \max_{i=1,2,\ldots,m} f_i(x^{(1)}) + \\ + (1-\lambda) \max_{i=1,2,\ldots,m} f_i(x^{(2)}).$$

4°. *Если функция f выпукла на множестве X, то для любых точек $x^{(1)}, x^{(2)}, \ldots, x^{(m)} \in X$ и произвольных неотрицательных чисел $\lambda_1, \lambda_2, \ldots, \lambda_m$ вида $\lambda_1+\lambda_2+\ldots+\lambda_m=1$ имеет место неравенство Иенсена*

$$f(\lambda_1 x^{(1)} + \lambda_2 x^{(2)} + \ldots + \lambda_m x^{(m)}) \leqslant \lambda_1 f(x^{(1)}) + \lambda_2 f(x^{(2)}) + \ldots + \\ + \lambda_m f(x^{(m)}).$$

При $m=2$ это неравенство совпадает с неравенством из определения выпуклой функции. Для доказательства неравенства Иенсена при произвольном натуральном $m>2$ используют метод математической индукции.

5°. *Предположим, что функция $\varphi(t)$ одной переменной t выпукла и не убывает на промежутке $[a, b]$ (этот промежуток может быть и бесконечным), а функция f выпукла на множестве $X \subset \mathbf{R}^n$, причем неравенство $a \leqslant f(x) \leqslant b$ верно при $x \in X$. Тогда сложная функция $g(x) = \varphi[f(x)]$ выпукла на множестве X.*

Рассмотрим две произвольные точки $x^{(1)}, x^{(2)} \in X$ и произвольное число $\lambda \in [0, 1]$. Имеем

$$g(\lambda x^{(1)} + (1-\lambda) x^{(2)}) = \varphi[f(\lambda x^{(1)} + (1-\lambda) x^{(2)})] \leqslant \\ \leqslant \varphi[\lambda f(x^{(1)}) + (1-\lambda) f(x^{(2)})] \leqslant \lambda \varphi[f(x^{(1)})] + \\ + (1-\lambda) \varphi[f(x^{(2)})] = \lambda g(x^{(1)}) + (1-\lambda) g(x^{(2)}).$$

Непосредственной проверкой можно установить, что линейная функция $y = \langle c, x \rangle + b$ одновременно и выпукла, и вогнута.

Ранее установлено, что гессиан квадратичной функции $y = \frac{1}{2}\langle Qx, x \rangle + \langle c, x \rangle + b$ с симметричной матрицей Q равен матрице Q. По теореме 1.2 квадратичная функция выпукла на $\mathbf{R}^n$ тогда и только тогда, когда Q — неотрицательно-определенная матрица.

Числовую функцию f называют *строго выпуклой* на выпуклом множестве X, если для любых двух точек $x^{(1)}, x^{(2)} \in X$, $x^{(1)} \neq x^{(2)}$, и произвольного числа λ, $0<\lambda<1$, справедливо строгое неравенство

$$f(\lambda x^{(1)} + (1-\lambda) x^{(2)}) < \lambda f(x^{(1)}) + (1-\lambda) f(x^{(2)}).$$

Для строго выпуклой функции имеет место утверждение, аналогичное теореме 1.1; следует лишь в (1.9) знак неравенства заменить знаком строгого неравенства и дополнительно считать, что $x^{(1)} \neq x^{(2)}$. Точно так же теорема 1.2 дает необходимые и достаточные условия строгой выпуклости функции в терминах матрицы вторых производных, если в (1.10) равенство исключить и дополнительно считать, что $z \neq 0_n$. На основании этого утверждения можно установить строгую выпуклость, например, функций e^x, $-\ln x$. В самом деле, для первой функции имеем $\langle \nabla^2 f(x)x, x\rangle = x^2 e^x > 0$ для всех $x \neq 0$, а для второй функции $\langle \nabla^2 f(x)x, x\rangle = 1 > 0$.

Линейная функция не является строго выпуклой. Квадратичная функция строго выпуклая, если матрица Q положительно определена.

Введем понятие строго вогнутой функции: функция $f(x)$ *строго вогнута* на выпуклом множестве X, $X \subset \mathbf{R}^n$, если функция $-f(x)$ строго выпукла.

Существует два важных обобщения понятия выпуклой функции. Одно из таких обобщений — псевдовыпуклая функция. Говорят, что дифференцируемая на выпуклом множестве X числовая функция f *псевдовыпукла* на X, если для всех точек $x^{(1)}$, $x^{(2)} \in X$, таких, что

$$\langle \nabla f(x^{(1)}), x^{(2)} - x^{(1)} \rangle \geqslant 0. \tag{1.12}$$

выполняется неравенство

$$f(x^{(2)}) \geqslant f(x^{(1)}). \tag{1.13}$$

Пусть функция f непрерывно дифференцируема на выпуклом множестве X. Тогда по теореме 1.1 имеет место неравенство (1.9) и поэтому, если верно неравенство (1.12), из (1.9) вытекает (1.13). Это означает, что *непрерывно дифференцируемая выпуклая функция всегда псевдовыпукла.*

Рассмотрим функцию одной переменной $f(x) = -e^x$. Здесь $\nabla f(x) = f'(x) = -e^x$. Пусть для точек x_1, $x_2 \in \mathbf{R}$ верно неравенство (1.12), т. е. $-e^{x_1} \cdot (x_2 - x_1) \geqslant 0$. Отсюда следует, что $x_1 \geqslant x_2$, а значит, верно неравенство (1.13), т. е. $-e^{x_2} \geqslant -e^{x_1}$. Таким образом, функция $-e^x$ псевдовыпукла на $\mathbf{R}$. Заметим, что эта функция не является выпуклой (она вогнута). Это означает, что класс псевдовыпуклых функций включает класс непрерывно дифференцируемых выпуклых функций и является «шире» последнего.

Функция $f(x)$ называется *псевдовогнутой* на выпуклом множестве X, если на этом множестве псевдовыпукла функция $-f(x)$.

Проверим, что *дробно-линейная функция,* т. е. функция вида

$$f(x) = \frac{\langle a, x \rangle + b}{\langle c, x \rangle + d},$$

где a, $c \in \mathbf{R}^n$; b, $d \in \mathbf{R}$, одновременно псевдовыпукла и псевдовогнута на произвольном выпуклом множестве X, на котором выполнено неравенство $\langle c, x\rangle + d \neq 0$. Имеем

$$\langle \nabla f(x^{(1)}), x^{(2)} - x^{(1)} \rangle =$$
$$= \frac{\langle a, x^{(2)} - x^{(1)} \rangle (\langle c, x^{(1)} \rangle + d) - \langle c, x^{(2)} - x^{(1)} \rangle (\langle a, x^{(1)} \rangle + b)}{(\langle c, x^{(1)} \rangle + d)^2} =$$
$$= \frac{(\langle a, x^{(2)} \rangle + b)(\langle c, x^{(1)} \rangle + d) - (\langle a, x^{(1)} \rangle + b)(\langle c, x^{(2)} \rangle + d)}{(\langle c, x^{(1)} \rangle + d)^2}.$$

Пусть

$$\langle \nabla f(x^{(1)}),\ x^{(2)}-x^{(1)}\rangle = \frac{\langle a,\ x^{(2)}\rangle + b}{\langle c,\ x^{(1)}\rangle + d} -$$

$$-\frac{(\langle a,\ x^{(1)}\rangle + b)(\langle c,\ x^{(2)}\rangle + d)}{(\langle c,\ x^{(1)}\rangle + d)^2} \geqslant 0.$$

Числа $\langle c,\ x^{(1)}\rangle+d$ и $\langle c,\ x^{(2)}\rangle+d$ имеют один и тот же знак, поскольку в противном случае нашлась бы такая точка $x^{(3)}$ на отрезке между точками $x^{(1)}$ и $x^{(2)}$, что $\langle c,\ x^{(3)}\rangle+d=0$ и $x^{(3)}\in X$. Умножая неравенство

$$\frac{\langle a,\ x^{(2)}\rangle + b}{\langle c,\ x^{(1)}\rangle + d} - \frac{(\langle a,\ x^{(1)}\rangle + b)(\langle c,\ x^{(2)}\rangle + d)}{(\langle c,\ x^{(1)}\rangle + d)^2} \geqslant 0$$

на положительное число $(\langle c,\ x^{(1)}\rangle+d)/(\langle c,\ x^{(2)}\rangle+d)$, получаем

$$\frac{\langle a,\ x^{(2)}\rangle + b}{\langle c,\ x^{(2)}\rangle + d} - \frac{\langle a,\ x^{(1)}\rangle + b}{\langle c,\ x^{(1)}\rangle + d} \geqslant 0$$

или $f(x^{(2)})\geqslant f(x^{(1)})$. Тем самым псевдовыпуклость дробно-линейной функции установлена. Аналогично можно проверить псевдовогнутость этой функции.

Другим важным обобщением выпуклой функции является квазивыпуклая функция. Числовая функция f, определенная на выпуклом множестве X, называется *квазивыпуклой* на множестве X, если для любых точек $x^{(1)},\ x^{(2)}\in X$ и любого числа $\lambda\in[0,\ 1]$ выполняется неравенство

$$f(\lambda x^{(1)}+(1-\lambda)\,x^{(2)}) \leqslant \max\{f(x^{(1)}),\ f(x^{(2)})\}.$$

Поскольку

$$\lambda f(x^{(1)})+(1-\lambda)\,f(x^{(2)}) \leqslant \max\{f(x^{(1)}),\ f(x^{(2)})\},\ \lambda\in[0,1],$$

выпуклая функция всегда квазивыпукла, т. е. класс квазивыпуклых функций содержит класс выпуклых функций. Более того, можно доказать, что *класс дифференцируемых квазивыпуклых функций содержит класс псевдовыпуклых функций*. На рис. 1.5, *а, б, в* изображены графики выпуклой, псевдовыпуклой и квазивыпуклой функций одной переменной.

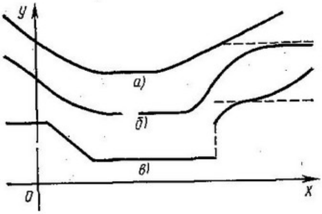

Рис. 1.5

Важное характеристическое свойство квазивыпуклых функций устанавливает следующая теорема.

Теорема 1.3. *Для того чтобы функция f, заданная на выпуклом множестве X, была на этом множестве квазивыпуклой, необходимо и достаточно, чтобы для любого числа $\alpha\in\mathbb{R}$ множество H_α вида*

$$H_\alpha=\{x\in X \mid f(x)\leqslant\alpha\}$$

было выпуклым.

□ Необходимость. Пусть функция f квазивыпукла на множестве X. Рассмотрим две произвольные точки $x^{(1)}, x^{(2)} \in H_\alpha$, где α — произвольное фиксированное число. Тогда выполнены неравенства $f(x^{(1)}) \leqslant \alpha$, $f(x^{(2)}) \leqslant \alpha$ и, поскольку функция f квазивыпукла, при любом $\lambda \in [0, 1]$ имеем

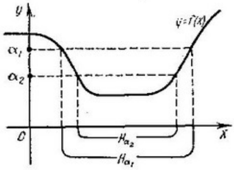

Рис. 1.6

$$f(\lambda x^{(1)} + (1-\lambda)x^{(2)}) \leqslant \max\{f(x^{(1)}), f(x^{(2)})\} \leqslant \alpha.$$

Это означает, что $\lambda x^{(1)} + (1-\lambda)x^{(2)} \in H_\alpha$. Выпуклость множества H_α установлена.

Достаточность. Теперь пусть для каждого $\alpha \in \mathrm{R}$ множество H_α выпукло. Рассмотрим две произвольные точки $x^{(1)}, x^{(2)} \in X$. Предположим для определенности, что имеет место неравенство

$$f(x^{(1)}) \geqslant f(x^{(2)}), \tag{1.14}$$

и введем число α равенством

$$\alpha = f(x^{(1)}). \tag{1.15}$$

В силу выпуклости множества H_α при данном α имеем

$$\lambda x^{(1)} + (1-\lambda)x^{(2)} \in H_\alpha \text{ для любого } \lambda \in [0, 1].$$

По определению множества H_α это означает, что выполнено неравенство

$$f(\lambda x^{(1)} + (1-\lambda)x^{(2)}) \leqslant \alpha.$$

Отсюда, используя формулы (1.14) и (1.15), получаем

$$f(\lambda x^{(1)} + (1-\lambda)x^{(2)}) \leqslant \alpha = f(x^{(1)}) = \max\{f(x^{(1)}), f(x^{(2)})\}.$$

Следовательно, функция f квазивыпукла. ■

На рис. 1.6 приведена геометрическая иллюстрация к доказанной теореме.

Глава 2

ТЕОРЕТИЧЕСКИЕ ОСНОВЫ ОПТИМИЗАЦИИ В ЕВКЛИДОВОМ ПРОСТРАНСТВЕ

В главе дана математическая постановка задачи оптимизации, приведены определения оптимальной точки и оптимумов (локального и глобального). Рассмотрены случаи, когда задача оптимизации разрешима, т. е. когда существует оптимальная точка и оптимум. Проводится классификация задач оптимизации и рассматриваются постановки задач линейного, нелинейного и геометрического программирования, а также дана математическая постановка задачи оптимального управления дискретной системой. Основное содержание составляют необходимые и достаточные условия оптимальности и теоремы двойственности выпуклого и линейного программирования.

§ 2.1. МАТЕМАТИЧЕСКАЯ ПОСТАНОВКА ЗАДАЧИ ОПТИМИЗАЦИИ

1. Допустимое множество и целевая функция. Постановка задачи оптимизации содержит *множество допустимых решений* X и числовую функцию f, определенную на множестве X, называемую *целевой функцией* (а также *критерием оптимальности* или *критерием качества*). Множество X также называют *допустимым множеством* или *множеством возможных решений.*

В первых семи главах настоящей книги понятие решения отождествляется с вектором (точкой) n-мерного евклидова пространства $\mathbf{R}^n$. В соответствии с этим допустимое множество X представляет собой некоторое подмножество пространства $\mathbf{R}^n$, т. е. $X \subset \mathbf{R}^n$, а целевая функция — это функция $f(x)$ n переменных $x_1, x_2, ..., x_n$. Не исключается случай равенства $X = \mathbf{R}^n$. Для указания элементов множества X наряду с термином «решения» далее используются термины «векторы» и «точки».

Нестрого говоря, задача оптимизации заключается в выборе среди элементов множества X такого решения, которое было бы с определенной точки зрения наиболее предпочтительным. Сравнение решений по предпочтительности осуществляется с помощью целевой функции. Выделяют два варианта сравнения произвольной пары решений $x^{(1)}, x^{(2)} \in X$ с помощью функции f. Можно считать, что решение $x^{(1)}$ предпочтительнее решения $x^{(2)}$, если выполнено неравенство $f(x^{(1)}) < f(x^{(2)})$. Тогда поиск наиболее предпочтительного решения среди всех элементов множества X заключается в нахождении решения, доставляющего наименьшее возможное значение целевой функции f на множестве X. В этом случае задача оптимизации представляет собой задачу минимизации. Если же считается, что решение $x^{(1)}$ предпочтительнее решения $x^{(2)}$, когда выполнено неравенство $f(x^{(1)}) > f(x^{(2)})$, то поиск наиболее предпочтительного решения представляет собой задачу максимизации функции f на множестве X.

Чтобы решить *задачу минимизации* функции f на множестве X, нужно найти такой вектор $x^{(0)} \in X$ (а также соответствующее значение $f(x^{(0)})$), чтобы неравенство $f(x^{(0)}) \leqslant f(x)$ было выполнено для всех $x \in X$. При этом решение $x^{(0)}$ называют *оптимальным* (точнее говоря, *минимальным*), а значение $f(x^{(0)})$ — *оптимумом* (*минимумом*). Тот факт, что решение $x^{(0)}$ оптимальное, т. е. доставляет наименьшее возможное значение функции f на множестве X, записывают в виде

$$f(x^{(0)}) = \min_{x \in X} f(x),$$

а для постановки задачи минимизации используют запись

$$f(x) \to \min_{x \in X}.$$

Задача максимизации заключается в отыскании такого вектора $x^{(0)} \in X$ (и соответствующего значения $f(x^{(0)})$), для которого неравенство $f(x^{(0)}) \geqslant f(x)$ имеет место при всех $x \in X$. Если $x^{(0)}$ — решение задачи максимизации, то используют запись

$$f(x^{(0)}) = \max_{x \in X} f(x).$$

При этом $x^{(0)}$ называют *максимальным (оптимальным) решением*, а значение $f(x^{(0)})$ — *максимумом (оптимумом)*.

Как указано выше, решить задачу оптимизации означает найти оптимальную точку $x^{(0)}$ и оптимальное значение $f(x^{(0)})$. Если найдена оптимальная точка, то определение оптимального значения обычно не составляет труда. Если же найдено только оптимальное значение $f(x^{(0)})$, то для отыскания оптимальной точки необходимо решить уравнение $f(x) = f(x^{(0)})$, что может составить сложную вычислительную задачу.

В настоящей книге в основном рассматриваются задачи минимизации. Все результаты, полученные для задач минимизации, можно легко переформулировать применительно к задачам максимизации.

2. Локальный и глобальный минимумы. В теории оптимизации удобно рассматривать два вида оптимумов: локальный и глобальный. Говорят, что точка $x^{(0)} \in X$ доставляет функции f на множестве X *локальный минимум*, если существует такая окрестность $U_\varepsilon(x^{(0)})$ $(\varepsilon > 0)$ точки $x^{(0)}$, что неравенство $f(x^{(0)}) \leqslant f(x)$ справедливо для всех $x \in X \cap U_\varepsilon(x^{(0)})$. *Глобальный минимум* функции f доставляет точка $x^{(0)} \in X$, для которой записанное выше неравенство выполняется при всех $x \in X$. Таким образом, глобальный минимум — это просто минимум в смысле определения, данного в предыдущем пункте. Прилагательное «глобальный» используется для того, чтобы подчеркнуть отличие этого минимума от локального минимума.

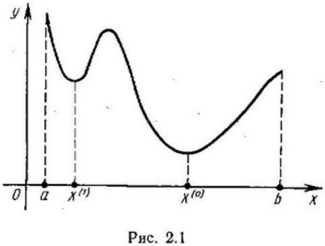

Рис. 2.1

В соответствии с приведенными определениями в первом случае точку $x^{(0)}$ называют *точкой локального минимума*, а во втором — *точкой глобального минимума*. Таким образом, выражения «$x^{(0)}$ — минимальная точка», «$x^{(0)}$ — решение задачи минимизации» и «$x^{(0)}$ — точка глобального минимума» означают одно и то же.

Аналогично определяются понятия *точки локального* и *глобального максимума*, а также *локальный* и *глобальный максимумы*.

Точка локального минимума не всегда является точкой глобального минимума (рис. 2.1), а значит, и локальный минимум не всегда совпадает с глобальным минимумом. В следующем ниже утверждении формулируются условия, накладываемые на множество X и функцию f, при выполнении которых указанные точки (и минимумы) совпадают.

Теорема 2.1. *Если множество $X \subset \mathbf{R}^n$ выпукло, а функция f выпукла или же псевдовыпукла на X, то всякая точка локального минимума является и точкой глобального минимума.*

Напомним, что определения выпуклой и псевдовыпуклой функций приведены в п. 2. § 1.3.

□ Пусть функция f выпукла и $x^{(0)} \in X$ — точка локального минимума, т. е. существует такое число $\varepsilon > 0$, что выполнено неравенство $f(x^{(0)}) \leqslant f(x)$ для всех $x \in X \cap U_\varepsilon(x^{(0)})$. Рассмотрим произвольную точку $x' \in X$. Множество X выпуклое, поэтому точка $x = \lambda x^{(0)} + (1-\lambda)x'$ принадлежит множеству X при любом $\lambda \in (0, 1)$. Но при значениях λ, близких к единице, точка x «попадает» в окрестность $U_\varepsilon(x^{(0)})$ и, следовательно, для нее выполняется записанное выше неравенство, т. е.

$$f(x^{(0)}) \leqslant f(\lambda x^{(0)} + (1-\lambda)x').$$

Отсюда, так как функция f выпуклая, имеем

$$f(x^{(0)}) \leqslant \lambda f(x^{(0)}) + (1-\lambda) f(x').$$

В результате несложных преобразований приходим к неравенству $f(x^{(0)}) \leqslant f(x')$. Так как x' — произвольная точка множества X, то $x^{(0)}$ — точка глобального минимума функции f на X.

Пусть теперь функция f псевдовыпуклая и $f(x^{(0)})$ — ее локальный минимум. Докажем, что $f(x^{(0)})$ является и глобальным минимумом. Для этого предположим противное: найдется точка $x^* \in X$, для которой выполнено неравенство $f(x^*) < f(x^{(0)})$. Функция f псевдовыпуклая, поэтому из последнего неравенства следует $\langle \nabla f(x^{(0)}), x^* - x^{(0)} \rangle < 0$.

Согласно определению дифференцируемости функции f в точке $x^{(0)}$ имеем

$$f(x^{(0)} + \lambda(x^* - x^{(0)})) = f(x^{(0)}) + \lambda \langle \nabla f(x^{(0)}), x^* - x^{(0)} \rangle + \\ + \lambda \|x^* - x^{(0)}\| \alpha(x^{(0)}, \lambda(x^* - x^{(0)})). \tag{2.1}$$

Поскольку участвующее в этом представлении скалярное произведение меньше нуля и $\alpha(x^{(0)}, \lambda(x^* - x^{(0)})) \to 0$ при $\lambda \to 0$, найдется такое положительное число λ_0, что для всех чисел $\lambda \in (0, \lambda_0)$ будет выполняться неравенство

$$\lambda [\langle \nabla f(x^{(0)}), x^* - x^{(0)} \rangle + \|x^* - x^{(0)}\| \alpha(x^{(0)}, \lambda(x^* - x^{(0)}))] < 0.$$

Следовательно, из равенства (2.1) получаем неравенство

$$f(x^{(0)}+\lambda(x^*-x^{(0)}))<f(x^{(0)}) \text{ для всех } \lambda\in(0,\lambda_0).$$

Здесь число λ можно выбрать настолько малым, чтобы выполнялось условие $x^{(0)}+\lambda(x^*-x^{(0)})\in X\cap U_\varepsilon(x^{(0)})$. Это противоречит тому, что $x^{(0)}$ — точка локального минимума. Следовательно, сделанное предположение неверно и на самом деле локальный минимум $f(x^{(0)})$ является глобальным. ■

Для квазивыпуклой функции f утверждение теоремы, вообще говоря, неверно. В этом позволяет убедиться простой пример функции одной переменной $f(x)=\operatorname{sign} x$, например на отрезке $[-1, 1]$. Здесь точка $x^{(0)}=0{,}5$ является точкой локального минимума, но не является точкой глобального минимума.

§ 2.2. РАЗРЕШИМОСТЬ ЗАДАЧИ ОПТИМИЗАЦИИ

1. Теорема Вейерштрасса. Сформулированная в предыдущем параграфе задача оптимизации имеет решение не при любых целевых функциях и допустимых множествах. Существуют задачи, в которых невозможно найти оптимальную точку и оптимальное значение. Например, не существует точек минимума функции одной переменной f на множестве X в случаях, изображенных на рис. 2.2. В первом случае точка минимума не существует, поскольку множество X незамкнутое. Во втором случае — вследствие неограниченно-

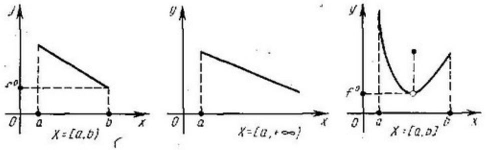

Рис. 2.2

сти X. Наконец, в третьем случае минимум не достигается из-за того, что функция f не является непрерывной. Эти примеры приводят к мысли, что в случае непрерывной целевой функции и замкнутого ограниченного (т. е. компактного) допустимого множества задача оптимизации должна иметь решение.

Действительно, имеет место следующее утверждение, которое называют теоремой Вейерштрасса.

Теорема 2.2. *Если множество $X\subset\mathbf{R}^n$ не пусто и компактно, а функция f непрерывна на нем, то множество точек глобального минимума (а также множество точек глобального максимума) не пусто и компактно.*

В частности, задача оптимизации всегда разрешима, если допустимое множество содержит конечное число элементов.

В теореме Вейерштрасса требование компактности множества X является довольно «жестким». На практике нередко встречается допустимое множество вида

$$X=\{x\in\mathbf{R}^n \mid x_i\geqslant 0, \quad i=1, 2,\ldots, n\},$$

которое не является ограниченным, а значит, и компактным. Приведем утверждение, аналогичное теореме Вейерштрасса, в котором не требуется ограниченность допустимого множества, однако предполагается, что целевая функция кроме непрерывности должна удовлетворять некоторым дополнительным условиям.

Теорема 2.3. *Пусть $X\subset\mathbf{R}^n$ — непустое замкнутое множество и функция f непрерывна на нем. Предположим, что выполняется по крайней мере одно из следующих условий:*

1) существует такая точка $x'\in X$, что множество вида $\{x\in X \mid f(x)\leqslant f(x')\}$ ограничено;

2) для любой последовательности $\{x^{(k)}\}_{k=1}^{\infty}$ точек множества X, обладающей тем свойством, что $\lim\limits_{k\to\infty}\|x^{(k)}\|=+\infty$ (если такая последовательность найдется), справедливо равенство $\lim\limits_{k\to\infty} f(x^{(k)})=+\infty$. Тогда множество точек глобального минимума не пусто и компактно.

Доказательства сформулированных теорем можно найти в [8]. Рис. 2.3 иллюстрирует теорему 2.3 при выполнении условия 1).

2. Обобщенная задача оптимизации. В теории оптимизации иногда удобно рассматривать более общую задачу оптимизации, в которой понятие решения определяется таким образом, что оно всегда существует. Для того чтобы сформулировать эту обобщенную задачу, понадобится определение точной нижней грани.

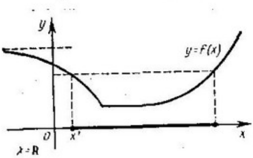

Рис. 2.3

Число (или символ $-\infty$) f^0 называют *точной нижней гранью* или *инфимумом* функции f на множестве X, если неравенство $f^0\leqslant f(x)$ * имеет место для всех $x\in X$ и, кроме того, для любого числа $f'>f^0$ найдется точка $x'\in X$ такая, что верно неравенство $f(x')<f'$. Тот факт, что f^0 — точная нижняя грань функции f на множестве X, записывают в виде

$$f^0=\inf_{x\in X} f(x). \tag{2.2}$$

Аналогично вводится понятие точной верхней грани. Число (или символ $+\infty$) f^* называют *точной верхней гранью* или *супремумом* функции f на множестве X, если неравенство $f^*\geqslant f(x)$ справедливо для всех $x\in X$ и для любого числа $f'<f^*$ найдется точка $x'\in X$ та-

* Считается, что $-\infty$ меньше любого числа.

кая, что верно $f(x') > f'$. Для точной верхней грани используют обозначение

$$f^* = \sup_{x \in X} f(x).$$

В курсе математического анализа доказывается, что произвольная числовая функция на любом непустом допустимом множестве имеет точную нижнюю и точную верхнюю грани. Так, например, точная нижняя грань в случаях 1) и 3), изображенных на рис. 2.2, обозначена через f^0. В случае 2) на этом рисунке $f^0 = -\infty$.

Из примеров видно, что не всегда можно указать точку, в которой точная грань достигается, т. е. точку $x^{(0)}$, для которой $f(x^{(0)}) = \inf_{x \in X} f(x)$. Поэтому в обобщенной задаче минимизации $f(x) \to \inf_{x \to X}$ под решением понимают не отдельную точку, как это имеет место в обычной задаче оптимизации, а последовательность точек $\{x^{(k)}\}_{k=1}^{\infty}$, $x^{(k)} \in X$, $k = 1, 2, ..$, такую, что

$$\lim_{k \to \infty} f(x^{(k)}) = f^0. \tag{2.3}$$

Эта последовательность всегда существует и называется *минимизирующей последовательностью.* Любая подпоследовательность минимизирующей последовательности сама является минимизирующей последовательностью. Поэтому, как правило, минимизирующих последовательностей «довольно много».

Таким образом, *обобщенная задача минимизации* целевой функции f на множестве X заключается в отыскании числа (или символа $-\infty$) f^0 и последовательности точек $\{x^{(k)}\}_{k=1}^{\infty}$, $x^{(k)} \in X$, $k=1$, 2, ..., таких, что выполняются равенства (2.2)—(2.3). Как отмечено выше, эта задача всегда имеет решение.

§ 2.3. КЛАССИФИКАЦИЯ ЗАДАЧ ОПТИМИЗАЦИИ

1. Задачи без ограничений и задачи с ограничениями. Если $X = \mathbf{R}^n$, то говорят о *задаче минимизации без ограничений.* Действительно, в этом случае нужно найти такую точку $x^{(0)}$, чтобы неравенство $f(x^{(0)}) \leqslant f(x)$ выполнялось для всех точек пространства $\mathbf{R}^n$ без ограничения. Часто задачу минимизации без ограничений называют также *задачей безусловной минимизации.* При этом для характеризации точки минимума и самого минимума добавляют прилагательное «безусловный».

Если $X \neq \mathbf{R}^n$, то имеет место *задача минимизации с ограничениями.* В этом случае также говорят о *задаче условной минимизации,* о *точках условного минимума* и об *условном минимуме.*

Соответствующая терминология вводится и для задач максимизации.

Если допустимое множество X задано в виде

$$X = \{x \in \mathbf{R}^n \mid g_j(x) \leqslant 0, \ j = 1, 2, ..., k; \ g_j(x) = 0,$$
$$j = k+1, ..., m\}, \tag{2.4}$$

где все числовые функции g_j определены на $\mathbf{R}^n$, то говорят о *задаче математического программирования.* Среди задач этого класса различают *задачи с ограничениями типа неравенств* — когда множество X имеет вид (2.4) и $m=k$; *задачи с ограничениями типа равенств* — когда в (2.4) неравенства отсутствуют, т. е. $k=0$, и *задачи со смешанными ограничениями* — когда в задании множества X встречаются как неравенства, так и равенства.

2. Задачи математического программирования. Если целевая функция f и функции $g_1, g_2, \ldots, g_m$ из (2.4) линейны, то имеем *общую задачу линейного программирования.* Часто рассматривают *стандартную задачу линейного программирования*, которая записывается следующим образом:

$$f(x)=\langle c, x\rangle \to \min$$

при условиях

$$\langle a^{(j)}, x\rangle \geqslant b_j, \; j=1, 2, \ldots, k,$$

$$x_i \geqslant 0, \; i=1, 2, \ldots, n,$$

где $c, a^{(1)}, a^{(2)}, \ldots, a^{(k)}$ — фиксированные векторы из $\mathbf{R}^n$, а $b_1, b_2, \ldots, b_k$ — фиксированные числа. Выделяют также *каноническую задачу линейного программирования:*

$$\langle c, x\rangle \to \min;$$

$$\langle a^{(j)}, x\rangle = b_j, \; j=1, 2, \ldots, k,$$

$$x_i \geqslant 0, \; i=1, 2, \ldots, n.$$

Линейные функции являются наиболее «простыми» функциями, поэтому задачи линейного программирования в определенном смысле проще остальных задач математического программирования и более детально исследованы.

Если среди функций $f, g_1, g_2, \ldots, g_m$ имеется хотя бы одна, не являющаяся линейной, то говорят о *задаче нелинейного программирования.*

В классе задач нелинейного программирования выделяют *задачу квадратичного программирования.* В этой задаче целевая функция является квадратичной

$$f(x)=\frac{1}{2}\langle Qx, x\rangle + \langle c, x\rangle,$$

где Q — числовая симметричная матрица размера $n \times n$, c — фиксированный вектор из $\mathbf{R}^n$, а $g_1, g_2, \ldots, g_m$ — линейные функции.

Задачей выпуклого программирования называют задачу минимизации выпуклой целевой функции f на множестве (2.4) при выпуклых функциях $g_1, g_2, \ldots, g_k$ и линейных функциях $g_{k+1}, \ldots, g_m$.

Особое место занимают задачи геометрического программирования и оптимального управления дискретной системой.

3. Задача геометрического программирования. В задаче геометрического программирования все функции f, g_1, g_2, ..., g_k представляют собой *позиномиальные функции (позиномы)* вида

$$f(x)=\sum_{i=1}^{m} c_i x_1^{a_{i1}} x_2^{a_{i2}} \dots x_n^{a_{in}}, \tag{2.5}$$

$$g_j(x)=\sum_{i=1}^{m_j} c_i^{(j)} x_1^{a_{i1}^{(j)}} x_2^{a_{i2}^{(j)}} \dots x_n^{a_{in}^{(j)}},\ j=1,\ 2,\dots,k, \tag{2.6}$$

где все числа $c_i, c_i^{(j)}$ — положительные. показатели степеней a_{il}, $a_{il}^{(j)}$, произвольные фиксированные числа и компоненты вектора x также положительные. Сама *задача геометрического программирования* записывается следующим образом:

$$f(x)\to \min;$$

$$g_j(x)\leqslant 1, \quad j=1,\ 2,\dots,k,$$

$$x_i>0, \qquad i=1,\ 2,\dots,n,$$

где целевая функция имеет вид (2.5), а функции-ограничения — вид (2.6). Поясним, почему ограничения записаны в виде неравенства $g_j(x)\leqslant 1$, а не как обычно: $g_j(x)\leqslant 0$. Поскольку все коэффициенты $c_i^{(j)}$ и компоненты вектора x положительны, для функции g_j вида (2.6) неравенство $g_j(x)\leqslant 0$ решений не имеет. Поэтому для позиномиальной функции $\tilde{g}_j$ ограничение естественно записать в форме неравенства $\tilde{g}_j(x)\leqslant b_j$, где $b_j>0$. Разделив обе части неравенства на b_j, приходим к неравенству $g_j(x)\leqslant 1$, где обозначено $g_j(x)=\tilde{g}_j(x)/b_j$.

Многие задачи технического проектирования приводят к задаче геометрического программирования, которая представляет собой задачу нелинейного программирования специального вида.

4. Задача оптимального управления. Предположим, что имеется некоторая система, в которой переход из одного состояния в другое происходит в дискретные моменты времени $t=1, 2, \dots, N$. В задаче оптимального управления имеется два типа переменных: *переменные управления* (r-мерные векторы)

$$u^{(t)}=(u_1^{(t)},\ u_2^{(t)},\dots,u_r^{(t)})^{\mathrm{T}}$$

и *переменные состояния* системы (n-мерные векторы)

$$x^{(t)}=(x_1^{(t)},\ x_2^{(t)},\dots,x_n^{(t)})^{\mathrm{T}}.$$

Векторы управления $u^{(t)}$ можно выбирать из соответствующих *множеств допустимых управлений:*

$$u^{(t)}\in U_t\subset \mathrm{R}^r,\ t=1,\ 2,\dots,N. \tag{2.7}$$

При заданном состоянии $x^{(t-1)}$ в момент времени $t-1$ изменение состояния от момента $t-1$ к моменту t задается *векторной функцией* *преобразования* вида

$$g^{(t)}(x^{(t-1)},\ u^{(t)})=\left(g_1^{(t)}(x^{(t-1)},\ u^{(t)}),\ g_2^{(t)}(x^{(t-1)},\ u^{(t)}),\ldots\right.$$
$$\left.\ldots,\ g_n^{(t)}(x^{(t-1)},\ u^{(t)})\right)^{\mathrm{T}},$$

зависящей от предыдущего состояния $x^{(t-1)}$ и управления $u^{(t)}$ в данный момент времени t. В векторных обозначениях закон изменения состояния системы с течением времени описывается равенствами

$$x^{(t)}=g^{(t)}(x^{(t-1)},\ u^{(t)}),\ t=1,\ 2,\ldots,\ N. \qquad (2.8)$$

Начальное состояние $x^{(0)}$ в момент времени $t=0$ считают заранее заданным.

Говорят, что соотношения (2.7) и (2.8) определяют *дискретную систему управления*. Процесс управления этой системой происходит следующим образом. При $t=0$ система находится в состоянии $x^{(0)}$. Выбор некоторого допустимого управления $u^{(1)}\in U_1$ переводит систему в состояние $x^{(1)}=g^{(1)}(x^{(0)},\ u^{(1)})$, соответствующее моменту времени $t=1$. Далее, если выбрано управление $u^{(2)}\in U_2$, то в момент времени $t=2$ состояние системы определяется равенством (2.8) и имеет вид $x^{(2)}=g^{(2)}(x^{(1)},\ u^{(2)})$ и т. д. В результате *последовательность управлений* $u^{(1)},\ u^{(2)},\ \ldots,\ u^{(N)}$, согласно (2.8), однозначно определит соответствующую последовательность состояний $x^{(0)},\ x^{(1)},\ \ldots,\ x^{(N)}$, называемую *траекторией системы*. Различным последовательностям допустимых управлений будут соответствовать, вообще говоря, различные траектории.

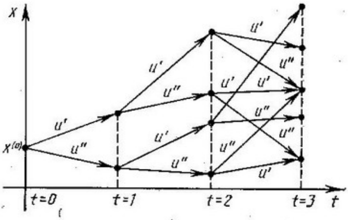

Рис. 2.4

Рис. 2.4 иллюстрирует переходы системы из одного состояния в другое для $N=3$ и двухэлементного множества допустимых управлений $U_t=\{u',\ u''\}$, $t=1,\ 2,\ 3$. Состояния обозначены точками, а переходам из одного состояния в другое соответствуют стрелки. В этом примере общее число последовательностей допустимых управлений и соответствующих траекторий равно $2^3=8$.

Качество управления дискретной системой оценивают *критерием оптимальности* суммарного вида

$$I=\sum_{t=1}^{N} f^t\ (x^{(t-1)},\ u^{(t)}). \qquad (2.9)$$

Задача оптимального управления заключается в отыскании такой последовательности допустимых управлений и соответствующей траектории, для которых критерий оптимальности (2.9) принимает наименьшее возможное значение. Другими словами, в задаче оптимального управления выбором допустимых векторов управлений требуется минимизировать критерий оптимальности I; при этом система должна удовлетворять начальному состоянию $x^{(0)}$ и изменяться во времени согласно равенствам (2.8).

В более сложных задачах оптимального управления кроме ограничений, накладываемых на выбор переменных управления (2.7), имеются также ограничения на переменные состояния (*фазовые ограничения*):

$$x^{(t)} \in X_t \subset \mathrm{R}^n, \ t=1, 2, \ldots, N. \tag{2.10}$$

Несмотря на внешнее несходство задачи нелинейного программирования и задачи оптимального управления, между ними имеется тесная связь, которая будет установлена в следующем параграфе.

Одним из распространенных методов решения задачи оптимального управления является метод динамического программирования. Изложению этого метода посвящена гл. 4.

§ 2.4. СВОДИМОСТЬ ЗАДАЧ ОДНОГО КЛАССА К ЗАДАЧАМ ДРУГОГО КЛАССА

1. Задачи минимизации и максимизации. Рассмотрим задачу максимизации

$$h(x) \to \max_{x \in X}. \tag{2.11}$$

Эта задача сводится к задаче минимизации

$$f(x) \to \min_{x \in X}, \tag{2.12}$$

если положить $f(x) = -h(x)$. Действительно, если $x^{(0)}$ — решение задачи минимизации, т. е.

$$f(x^{(0)}) \leqslant f(x) \text{ для всех } x \in X, \tag{2.13}$$

то, учитывая равенство $f(x) = -h(x)$, получаем отсюда неравенство

$$h(x^{(0)}) \geqslant h(x) \text{ для всех } x \in X.$$

Это означает, что каждое решение $x^{(0)}$ задачи минимизации (2.12) является также решением задачи максимизации (2.11). Аналогично можно проверить, что каждое решение задачи максимизации (2.11) является решением задачи минимизации (2.12), если $f(x) = -h(x)$. При этом оптимальные значения обеих задач связаны равенством $f(x^{(0)}) = -h(x^{(0)})$.

2. Задачи оптимизации и математического программирования. Задача математического программирования представляет собой частный случай задачи оптимизации, когда допустимое множество X задано в виде системы неравенств (2.4). Но общую задачу минимизации (2.12), в свою очередь, можно записать в виде следующей задачи математического программирования: $f(x)\to\min$ при ограничении $g(x)\leqslant 0$, где функция g определяется, например, так:

$$g(x)=\begin{cases}0, & \text{если } x\in X,\\ 1, & \text{если } x\notin X.\end{cases}$$

Следует отметить «неконструктивный» характер задания функции g.

3. Монотонное преобразование целевой функции. Пусть F — строго возрастающая функция одной переменной, определенная на множестве значений функции f (в частности, F может быть определена на всем множестве $\mathbf{R}$).

Множество точек минимума задачи (2.12) *совпадает с множеством точек минимума задачи*

$$F[f(x)]\to\min_{x\in X}. \tag{2.14}$$

В самом деле, если выполнено неравенство (2.13), то в силу строгого возрастания функции F получаем неравенство

$$F[f(x^{(0)})]\leqslant F[f(x)] \text{ для всех } x\in X,$$

т. е. $x^{(0)}$ — решение задачи (2.14). Обратно: если точка $x^{(0)}$ удовлетворяет последнему неравенству, то выполняется неравенство (2.13), так как в противном случае из неравенства $f(x^{(0)})>f(x)$ следует неравенство $F[f(x^{(0)})]>F[f(x)]$.

В соответствии с этим целевую функцию $f(x)=e^{\sum_{i=1}^{n}x_i^2}$ можно заменить, например, функцией $F[f(x)]=\sum_{i=1}^{n}x_i^2$, так как функция $F(t)=\ln t$ строго возрастает на множестве $(0, +\infty)$. Вторая целевая функция проще первой, а множества их точек минимума на $X\subset\mathbf{R}^n$ совпадают.

На этом же основании вместо недифференцируемой функции $|f(x)|$, где f — дифференцируемая функция, можно использовать новую дифференцируемую функцию $f^2(x)$ (так как функция $F(t)=t^2$ строго возрастает на $[0, +\infty)$).

Вообще, если значения целевой функции f на множестве X положительны, то ее можно заменить любой из следующих функций:

$$f^\alpha(x)\ (\alpha>0),\quad -1/f(x),\quad \log_\beta f(x)\ (\beta>1).$$

4. Ограничения в виде равенств и неравенств. Ограничение-равенство $g(x)=0$ в задаче математического программирования всегда можно заменить двумя ограничениями-неравенствами:

$$g(x)=0\longleftrightarrow\begin{cases}g(x)\geqslant 0,\\ g(x)\leqslant 0.\end{cases}$$

С другой стороны, ограничение-неравенство можно заменить эквивалентным ему ограничением-равенством:

$$g(x) \leqslant 0 \longleftrightarrow |g(x)| + g(x) = 0.$$

Этого же можно добиться, вводя дополнительную переменную x_{n+1}:

$$g(x) \leqslant 0 \longleftrightarrow g(x) + x_{n+1}^2 = 0.$$

5. Задачи оптимального управления и математического программирования. В предыдущем параграфе была сформулирована задача оптимального управления, которая представляет собой задачу оптимизации специального вида. Здесь будет показано, что между задачей оптимального управления и задачей математического программирования существует тесная связь. Оказывается, любую задачу оптимального управления всегда можно представить в виде некоторой задачи математического программирования и, наоборот, каждая задача математического программирования представима в виде одношаговой задачи оптимального управления.

Запишем задачу оптимального управления в следующем виде: найти последовательности векторов

$$\overline{x}^{(0)},\ \overline{x}^{(1)}, \ldots, \overline{x}^{(N)};\ \overline{u}^{(1)},\ \overline{u}^{(2)}, \ldots, \overline{u}^{(N)},$$

которые подчиняются равенствам

$$x^{(t)} = g^{(t)}(x^{(t-1)},\ u^{(t)}),\ t = 1,\ 2, \ldots, N,$$

и неравенствам

$$h_t(u^{(t)}) \leqslant 0,\ t = 1,\ 2, \ldots, N, \tag{2.15}$$

$$p_t(x^{(t)}) \leqslant 0,\ t = 1,\ 2, \ldots, N, \tag{2.16}$$

и, кроме того, доставляют наименьшее возможное значение критерию оптимальности

$$I = \sum_{t=1}^{N} f^t(x^{(t-1)},\ u^{(t)}).$$

Приведенная постановка отличается от постановки задачи оптимального управления из § 2.3 только видом ограничений на переменные состояния и переменные управления, т. е. вместо включений (2.7) и (2.10) здесь имеют место соответственно неравенства (2.15) и (2.16). Такая замена одних ограничений на другие правомерна в силу рассуждений, приведенных в п. 2 данного параграфа.

Введем переменную

$$z = (x^{(1)},\ x^{(2)}, \ldots, x^{(N)},\ u^{(1)},\ u^{(2)}, \ldots, u^{(N)})^{\tau},$$

представляющую собой вектор размерности $nN + rN$, составленный из компонент векторов состояний и компонент векторов управлений, которые записаны в порядке возрастания номеров компонент.

Введем следующие функции переменной z:

$$g_{tj}(z)=x_j^{(t)}-g_j^{(t)}(x^{(t-1)}, u^{(t)}),\ t=1,\ 2,\dots,N,\ j=1,\ 2,\dots,n,$$

$$g_t^{(1)}(z)=h_t(u^{(t)}),\ t=1,\ 2,\dots,N,$$

$$g_t^{(2)}(z)=p_t(x^{(t)}),\ t=1,\ 2,\dots,N,$$

$$f(z)=\sum_{t=1}^{N} f^t(x^{(t-1)}, u^{(t)}),$$

где $g_j^{(t)}$ — j-я компонента векторной функции преобразования $g^{(t)}$ из формулы (2.8).

Теперь задачу оптимального управления можно представить в виде следующей задачи математического программирования:

$$f(z)\to\min$$

при ограничениях

$$g_{tj}(z)=0,\ t=1,\ 2,\dots,N,\ j=1,\ 2,\dots,n,$$

$$g_t^{(1)}(z)\leqslant 0,\ t=1,\ 2,\dots,N,$$

$$g_t^{(2)}(z)\leqslant 0,\ t=1,\ 2,\dots,N.$$

С другой стороны, задачу математического программирования

$$f(x)\to\min;\ g_j(x)\leqslant 0,\ j=1,\ 2,\dots,\ k,$$

можно записать как одношаговую задачу оптимального управления системы, в которой переход из состояния $x^{(0)}$ в состояние $x^{(1)}$ происходит согласно закону

$$x^{(t)}=g^{(t)}(x^{(t-1)}, u^{(t)}),\ t=1,$$

при фазовых ограничениях

$$g_j^{(1)}\leqslant 0,\ j=1,\ 2,\dots,k,$$

с критерием оптимальности

$$I=f^1(x^{(0)}, u^{(1)}),$$

где

$$x^{(0)}=0_k,\ u^{(1)}=x,\ f^1(x^{(0)}, u^{(1)})=f(x),$$

$$g_j^{(1)}(x^{(0)}, u^{(1)})=g_j(x),\ j=1,\ 2,\ \dots,k.$$

6. Замечания. Рассмотренные простые приемы позволяют задачи оптимизации одного типа сводить к эквивалентным задачам другого типа. Это указывает на их общую природу и позволяет методы, разработанные для какого-то одного класса задач оптимизации использовать для исследования и решения задач другого класса.

Однако не следует считать, что, научившись решать задачи какого-либо одного типа (например, с ограничениями-равенствами), можно успешно решать и задачи другого типа (например, с ограничениями-неравенствами), сводящиеся к ним. Дело в том, что при сведении одних задач к другим может возрасти их сложность (увеличиться число переменных или ограничений) и, кроме того, в новой задаче могут быть потеряны некоторые полезные свойства старой задачи (например, дифференцируемость функций-ограничений).

§ 2.5. НЕОБХОДИМЫЕ И ДОСТАТОЧНЫЕ УСЛОВИЯ ОПТИМАЛЬНОСТИ В СЛУЧАЕ ДИФФЕРЕНЦИРУЕМЫХ ФУНКЦИЙ

1. Антиградиент и его свойство. Любой ненулевой вектор $y \in \mathbf{R}^n$ задает определенное *направление* в пространстве $\mathbf{R}^n$. Так, например, если $x^{(0)}$ — некоторая точка из $\mathbf{R}^n$, то множество вида

$$\{x \in \mathbf{R}^n \mid x = x^{(0)} + \tau y \text{ при некотором } \tau \geqslant 0\}$$

представляет собой луч, исходящий из точки $x^{(0)}$ в направлении вектора y (рис. 2.5).

Если функция f дифференцируема в точке $x^{(0)} \in X \subset \mathbf{R}^n$, то существует градиент $\nabla f(x^{(0)})$, который является некоторым вектором пространства $\mathbf{R}^n$. Вектор $-\nabla f(x^{(0)})$ называется *антиградиентом;* его длина совпадает с длиной градиента, а направление противоположно направлению градиента. Отличный от нулевого вектора антиградиент задает определенное направление в пространстве $\mathbf{R}^n$. Это направление обладает важным свойством, смысл которого раскрывает следующая лемма.

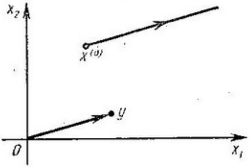

Рис. 2.5

Лемма 2.1. *Пусть функция f дифференцируема в точке $x^{(0)} \in \mathbf{R}^n$. Предположим, что для вектора $y \in \mathbf{R}^n$ выполняется неравенство $\langle -\nabla f(x^{(0)}), y \rangle > 0$. Тогда найдется такое положительное число σ, что для всех $\tau \in (0, \sigma)$ верно следующее неравенство:*

$$f(x^{(0)} + \tau y) < f(x^{(0)}). \tag{2.17}$$

□ Рассмотрим вектор

$$x = x^{(0)} + \tau y \text{ при } \tau > 0.$$

Так как функция f дифференцируема в точке $x^{(0)}$, то имеет место равенство

$$f(x^{(0)} + \tau y) = f(x^{(0)}) + \langle \nabla f(x^{(0)}), \tau y \rangle + \tau \| y \| \alpha(x^{(0)}, x^{(0)} + \tau y),$$

откуда следует

$$f(x^{(0)})-f(x^{(0)}+\tau y)=\tau[\langle-\nabla f(x^{(0)}),\ y\rangle-\|y\|\,\alpha(x^{(0)},\ x^{(0)}+\tau y)]. \qquad (2.18)$$

Поскольку $\alpha(x^{(0)},\ x^{(0)}+\tau y)\to 0$ при $\tau\to 0$ и по условию выполнено неравенство $\langle-\nabla f(x^{(0)}),\ y\rangle>0$, для некоторого достаточно малого числа σ правая часть равенства (2.18) положительна при каждом $\tau\in(0,\ \sigma]$. В этом случае положительной будет и левая часть равенства (2.18), что влечет неравенство (2.17). ■

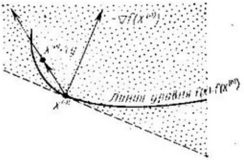

Рис. 2.6

Отличный от нулевого вектора антиградиент $-\nabla f(x^{(0)})$ определяет некоторое открытое полупространство $\{y\in\mathbf{R}^n \mid \langle-\nabla f(x^{(0)}),\ y\rangle>0\}$ (рис. 2.6). На этом рисунке все векторы изображены приложенными к точке $x^{(0)}$. Из леммы 2.1 следует, что любое достаточно малое перемещение из точки $x^{(0)}$ в направлении вектора y указанного полупространства ведет к уменьшению значения функции f по сравнению с ее значением в точке $x^{(0)}$. Таким образом, антиградиент указывает направление убывания целевой функции *. Это замечательное свойство антиградиента положено в основу многих вычислительных методов поиска оптимальных точек.

2. Необходимые и достаточные условия минимума в задачах без ограничений. Теорема 2.4. *Пусть функция f дифференцируема в точке $x^{(0)}\in\mathbf{R}^n$. Для того чтобы точка $x^{(0)}$ была точкой безусловного локального минимума функции f, необходимо, чтобы имело место равенство*

$$\nabla f(x^{(0)})=0_n. \qquad (2.19)$$

□ Действительно, если $x^{(0)}$ — точка безусловного локального минимума и $\nabla f(x^{(0)})\neq 0_n$, то для вектора $y=-\nabla f(x^{(0)})$ получаем

$$\langle-\nabla f(x^{(0)}),\ y\rangle=\langle-\nabla f(x^{(0)}),\ -\nabla f(x^{(0)})\rangle=$$
$$=\|\nabla f(x^{(0)})\|^2>0.$$

В этом случае согласно лемме 2.1, совершив достаточно малое перемещение из точки $x^{(0)}$ в направлении вектора $-\nabla f(x^{(0)}$, можно получить значение целевой функции меньшее, чем $f(x^{(0)})$. Это противоречит тому, что $x^{(0)}$ — точка локального минимума. ■

Теорема 2.4 представляет собой распространение известной из математического анализа теоремы Ферма на случай функции многих переменных. Поскольку точка глобального минимума является и точкой локального минимума, равенство (2.19) представляет

---

* Направление возрастания функции указывает ее градиент.

также необходимое условие того, чтобы точка $x^{(0)}$ была точкой глобального минимума в задаче без ограничений.

В соответствии с теоремой 2.4, если функция f дифференцируема, точками локального (значит, и глобального) минимума в задаче без ограничений могут быть только те точки, в которых градиент функции обращается в нулевой вектор. Таким образом, решение задачи безусловной минимизации функции f следует искать среди решений системы уравнений (2.19), которая в покоординатной форме имеет вид

$$\frac{\partial f(x)}{\partial x_i}=0,\ i=1,\ 2,\dots,n. \tag{2.19$'$}$$

Решения этой системы уравнений называют *стационарными точками* функции f. Такие точки являются «подозрительными» на оптимальные. Если известно, что число стационарных точек конечно, то остается найти их и отобрать точку минимума, сравнивая значения функции в этих точках. Следует, однако, заметить, что решение системы уравнений (2.19′) может составить сложную вычислительную задачу и поэтому практически реализовать указанный подход не всегда возможно.

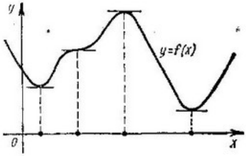

Рис. 2.7

Условие (2.19) является необходимым, но не достаточным условием оптимальности точки $x^{(0)}$. Это означает, что стационарная точка может оказаться точкой локального или глобального максимума, а в некоторых случаях не являться ни одной из них (рис. 2.7). В этом отношении полезным свойством обладают псевдовыпуклые (в частности, дифференцируемые выпуклые) функции: *любая стационарная точка псевдовыпуклой функции является точкой безусловного глобального минимума.*

В самом деле, если $x^{(0)}$ — стационарная точка функции f, то справедливо равенство $\nabla f(x^{(0)})=0_n$, а значит, равенство $\langle\nabla f(x^{(0)}), x-x^{(0)}\rangle=0$ верно для любого вектора $x\in\mathbf{R}^n$. Функция f псевдовыпуклая на множестве $\mathbf{R}^n$, отсюда следует, что неравенство $f(x^{(0)})\leqslant f(x)$ имеет место для любого $x\in\mathbf{R}^n$.

Таким образом, для псевдовыпуклых функций равенство (2.19) является как необходимым, так и достаточным условием того, чтобы точка $x^{(0)}$ была точкой безусловного глобального минимума. Для задач максимизации аналогичным свойством обладают псевдовогнутые функции.

Рассмотрим квадратичную функцию

$$f(x)=\frac{1}{2}\langle Qx,\ x\rangle+\langle a,\ x\rangle+b,$$

где $Q=(q_{ij})_{n\times n}$ — неотрицательно-определенная симметричная числовая матрица, $a\in\mathbf{R}^n$ и $b\in\mathbf{R}$. Как отмечено в предыдущей главе, такая квадратичная функция является выпуклой. Вычислим ее градиент:

$$\nabla f(x)=Qx+a$$

и приравняем его нулевому вектору:

$$Qx+a=0_n.$$

В развернутом виде это равенство можно записать так:

$$q_{11}x_1+q_{12}x_2+\ldots+q_{1n}x_n=-a_1,$$

$$\cdots\cdots\cdots\cdots\cdots$$

$$q_{n1}x_1+q_{n2}x_2+\ldots+q_{nn}x_n=-a_n.$$

В силу отмеченного выше, решения этой системы линейных уравнений и только они составляют множество точек безусловного глобального минимума данной квадратичной функции.

3. Условия оптимальности второго порядка. Предположим, что функция f дважды дифференцируема в точке $x^{(0)}\in\mathbf{R}$ и $\nabla^2 f(x^{(0)})$ — соответствующий гессиан (матрица вторых производных). Если $x^{(0)}$ — точка безусловного локального минимума функции f и неравенство $\langle\nabla^2 f(x^{(0)}),\ x,\ x\rangle<0$ выполнено при некотором $x\in\mathbf{R}^n$, то, используя определение дважды дифференцируемой функции и теорему 2.4, для точки вида $x^{(0)}+\tau x$ при $\tau>0$ можно записать равенство

$$f(x^{(0)}+\tau x)-f(x^{(0)})=\tau^2\left[\frac{1}{2}\langle\nabla^2 f(x^{(0)})x,\ x\rangle+\right.$$

$$\left.+\|x\|^2\beta(x^{(0)},\ x^{(0)}+\tau x)\right],$$

где правая часть отрицательная при всех достаточно малых положительных τ. Следовательно, имеет место неравенство $f(x^{(0)}+\tau x)<$ $<f(x^{(0)})$, что несовместимо с предположением о том, что $x^{(0)}$ — точка локального минимума. Таким образом, установлено следующее необходимое условие оптимальности: *если* $x^{(0)}$ *— точка безусловного локального минимума дважды дифференцируемой функции* f, *то имеет место равенство* (2.19) *и, кроме того, справедливо неравенство*

$$\langle\nabla^2 f(x^{(0)})x,\ x\rangle\geqslant 0 \text{ для всех } x\in\mathbf{R}^n.$$

Последнее неравенство означает неотрицательную определенность матрицы Гессе, вычисленной в точке $x^{(0)}$.

В формулировке полученных необходимых условий используется матрица вторых производных, поэтому их называют *условиями второго порядка* в отличие от необходимого *условия первого порядка* (2.19).

С помощью более сложных рассуждений [16] можно установить и достаточные условия оптимальности второго порядка: *если точка*

$x^{(0)}$ *стационарна и, кроме того, имеет место неравенство*

$$\langle \nabla^2 f(x^{(0)})\, x,\ x\rangle > 0 \text{ для любого } x \in \mathbf{R}^n,\ x \neq 0_n,$$

то $x^{(0)}$ — *точка безусловного локального минимума функции* f.

Неравенство, о котором здесь идет речь, означает положительную определенность матрицы Гессе. Напомним, что симметричная матрица является положительно-определенной тогда и только тогда, когда все ее угловые миноры положительны (критерий Сильвестра).

4. Необходимые условия минимума в задачах с ограничениями. Рассмотрим задачу условной минимизации вида

$$f(x) \to \min;\ g_j(x) \leqslant 0,\ j = 1,\ 2, \ldots, k, \tag{2.20}$$

где все функции f, g_1, g_2, ..., g_k предполагаются определенными на $\mathbf{R}^n$. Если точка минимума (локального или глобального) лежит на границе допустимого множества, определяемого записанной выше системой неравенств из (2.20), то градиент целевой функции не обязательно равен нулевому вектору. Это означает, что для задач оптимизации с ограничениями условия оптимальности, сформулированные в предыдущих пунктах, непригодны.

При решении задач условной оптимизации ограничения учитывают, вводя *функцию Лагранжа* переменных x и λ:

$$L(x,\ \lambda) = f(x) + \sum_{j=1}^{k} \lambda_j g_j(x).$$

Здесь λ_1, λ_2, ..., λ_k — некоторые числа, называемые *множителями Лагранжа*, которые составляют вектор, обозначаемый $\lambda \in \mathbf{R}^k$.

Условимся ограничение $g_j(x) \leqslant 0$ называть *активным в точке* $x^{(0)}$, если это неравенство при $x = x^{(0)}$ обращается в равенство, т. е. если $g_j(x^{(0)}) = 0$. В случае строгого неравенства $g_j(x^{(0)}) < 0$ данное ограничение называют *неактивным в точке* $x^{(0)}$.

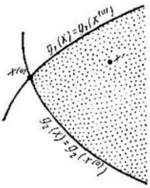

Рис. 2.8

На рис. 2.8 изображены точка $x^{(0)}$, в которой оба ограничения активны, и точка x', в которой оба ограничения неактивны.

Теорема 2.5. *Пусть все функции* f, g_1, g_2, ..., g_k *дифференцируемы в точке* $x^{(0)} \in \mathbf{R}^n$. *Если среди функций* g_1, g_2, ..., g_k *имеется хотя бы одна нелинейная, то дополнительно будем предполагать выполненным* у с л о в и е р е г у л я р н о с т и: *найдется такой вектор* $y \in \mathbf{R}^n$, *что неравенство* $\langle \nabla g_j(x^{(0)}),\ y\rangle < 0$ *справедливо для всех индексов* j *активных ограничений в точке* $x^{(0)}$. *Для того чтобы* $x^{(0)}$ *была точкой локального минимума в задаче* (2.20), *необходимо, чтобы существовали такие неотрицательные множители Лагранжа* λ_1, λ_2, ..., λ_k, *что*

$$g_j(x^{(0)}) \leqslant 0,\ j = 1,\ 2, \ldots, k, \tag{2.21}$$

$\nabla_x L(x, \lambda) = 0_n$ *или, что то же самое,*

$$\nabla f(x^{(0)}) + \sum_{j=1}^{k} \lambda_j \nabla g_j(x^{(0)}) = 0_n, \tag{2.22}$$

$$\lambda_j g_j(x^{(0)}) = 0, \ j = 1, 2, \ldots, k. \tag{2.23}$$

Здесь условие (2.21) очевидно: точка минимума $x^{(0)}$ должна удовлетворять исходным ограничениям, т. е. быть допустимой. Равенство (2.22) аналогично равенству (2.19) из теоремы 2.4, где вместо f использована функция Лагранжа. Согласно равенствам (2.23), те множители Лагранжа, которые соответствуют неактивным в точке $x^{(0)}$ ограничениям, должны обращаться в нуль. Среди перечисленных условий основным является равенство (2.22).

Теорема 2.5 утверждает, что необходимое условие равенства нулю градиента функции теоремы 2.4 можно распространить и на задачи с ограничениями, рассматривая вместо целевой функции функцию Лагранжа с соответствующим образом подобранными множителями $\lambda_1, \lambda_2, \ldots, \lambda_k$.

Если все ограничения в точке $x^{(0)}$ неактивны, то из (2.23) следует, что $\lambda_1 = \lambda_2 = \ldots = \lambda_k = 0$ и условие (2.22) превращается в условие (2.19). Это говорит о том, что в рассматриваемом случае $\lambda_1 = \lambda_2 = \ldots$ $\ldots = \lambda_k = 0$ ограничения фактически не учитываются и необходимым условием локального минимума в такой точке является равенство градиента целевой функции нулевому вектору.

Условие регулярности в теореме 2.5 геометрически означает, что все векторы $\nabla g_j(x^{(0)})$, отвечающие активным в точке $x^{(0)}$ ограничениям, расположены в одном и том же открытом полупространстве, порожденном некоторым вектором y. В случае когда все функции, участвующие в ограничениях, линейны, выполнение условия регулярности не требуется. Существуют примеры, показывающие, что если среди указанных функций имеются нелинейные, то без условия регулярности теорема 2.5, вообще говоря, не верна. На практике, как правило, условия регулярности выполняются. Часто более удобно использовать следующее условие: векторы $\nabla g_j(x^{(0)})$, отвечающие активным ограничениям в точке $x^{(0)}$, линейно независимы (можно доказать, что выполнение этого условия гарантирует выполнение условия регулярности из теоремы 2.5).

Нередко задача оптимизации содержит дополнительное условие неотрицательности переменных $x_1, x_2, \ldots, x_n$, т. е.

$$f(x) \to \min;$$

$$g_j(x) \leqslant 0, \ j = 1, 2, \ldots, k; \ -x \leq 0_n.$$

Этой задаче соответствует функция Лагранжа вида

$$L(x, \lambda, \mu) = f(x) + \sum_{j=1}^{k} \lambda_j g_j(x) - \sum_{i=1}^{n} \mu_i x_i.$$

Необходимые условия локального минимума в точке $x^{(0)}$ на основании теоремы 2.5 в этом случае можно записать следующим образом:

$$g_j(x^{(0)}) \leqslant 0,\ j=1,\ 2,\dots,k;\ x^{(0)} \geqq 0_n,$$

$$\nabla f(x^{(0)}) + \sum_{j=1}^{k} \lambda_j \nabla g_j(x^{(0)}) - \sum_{i=1}^{n} \mu_i e^{(i)} = 0_n,$$

$$\lambda_j g_j(x^{(0)}) = 0,\ j=1,\ 2,\dots,k;\ \mu_i x_i^{(0)} = 0,\ i=1,\ 2,\dots,n,$$

$$\lambda_j \geqslant 0,\ j=1,\ 2,\dots,k;\ \mu_i \geqslant 0,\ i=1,\ 2,\dots,n,$$

где $e^{(i)}$ — n-мерный вектор с нулевыми компонентами, среди которых только i-я отлична от нуля и равна единице. Условие регулярности в данной задаче соответственно превращается в следующее условие: существует такой вектор $y \in \mathbf{R}^n$, что $\langle \nabla g_j(x^{(0)}),\ y \rangle < 0$ и $\langle e^{(i)},\ y \rangle > 0$ для всех индексов j и i, при которых $g_j(x^{(0)}) = 0$ и $x_i^{(0)} = 0$.

Теперь на основании теоремы 2.5 запишем необходимые условия оптимальности применительно к *минимаксной задаче без ограничений*:

$$f(x) = \max_{i=1,2,\dots,m} f_i(x) \to \min_{x \in \mathbf{R}^n}, \tag{2.24}$$

где все функции $f_1,\ f_2,\ \dots,\ f_m$ считают дифференцируемыми на $\mathbf{R}^n$. Следует отметить, что теорема 2.4 непосредственно к задаче (2.24) неприменима, так как функция $\max\limits_{i=1,2,\dots,m} f_i(x)$, вообще говоря, не является дифференцируемой. Для того чтобы воспользоваться теоремой 2.5, сведем задачу (2.24) к эквивалентной задаче математического программирования, в которой все функции являются дифференцируемыми. Для этого введем дополнительную переменную x_{n+1} и рассмотрим следующую задачу с ограничениями:

$$h(x,\ x_{n+1}) \equiv x_{n+1} \to \min;$$

$$h_i(x,\ x_{n+1}) = f_i(x) - x_{n+1} \leqslant 0,\ i=1,\ 2,\dots,m. \tag{2.25}$$

Точка $x^{(0)}$ является точкой глобального минимума в задаче (2.24) *в том и только том случае, когда* $(x_1^{(0)},\dots,\ x_n^{(0)},\ x_{n+1}^{(0)})$ — *точка глобального минимума в задаче* (2.25); *при этом* $x_{n+1}^{(0)} = f(x^{(0)})$.

В самом деле, пусть $x^{(0)}$ — произвольная точка глобального минимума в задаче (2.24), т. е.

$$f(x^{(0)}) = \max_{i=1,2,\dots,m} f_i(x^{(0)}) \leqslant \max_{i=1,2,\dots,m} f_i(x) \text{ для всех } x \in \mathbf{R}^n. \tag{2.26}$$

Отсюда, полагая $x_{n+1}^{(0)} = f(x^{(0)})$, получаем неравенство

$$x_{n+1}^{(0)} \leqslant \max_{i=1,2,\dots,m} f_i(x) \text{ для всех } x \in \mathbf{R}^n. \tag{2.27}$$

Если взять произвольную допустимую точку задачи (2.25), то для нее верно неравенство $\max\limits_{i=1,2,\dots,m} f_i(x) \leqslant x_{n+1}$. Отсюда и из неравенства (2.27) следует, что

$$x_{n+1}^{(0)} \leqslant x_{n+1} \tag{2.28}$$

для любой допустимой точки задачи (2.25). Это означает, что $\left(x_1^{(0)}, \dots, x_n^{(0)}, x_{n+1}^{(0)}\right)$ — точка глобального минимума в задаче (2.25). Обратно: пусть $\left(x_1^{(0)}, \dots, x_n^{(0)}, x_{n+1}^{(0)}\right)$—произвольная точка глобального минимума в задаче (2.25). Так как эта точка является допустимой, то справедливо неравенство $f_i(x^{(0)}) - x_{n+1}^{(0)} \leqslant 0$, $i=1, 2, \dots, m$, а значит, верно неравенство $\max\limits_{i=1,2,\dots,m} f_i(x^{(0)}) \leqslant x_{n+1}^{(0)}$. Точка $\left(x_1^{(0)}, \dots, x_n^{(0)}, \max\limits_{i=1,2,\dots,m} f_i(x^{(0)})\right)$ также является допустимой, поэтому для нее выполняется неравенство (2.28), т. е. $x_{n+1}^{(0)} \leqslant \max\limits_{i=1,2,\dots,m} f_i(x^{(0)})$. Сравнивая два последних неравенства, приходим к равенству. $x_{n+1}^{(0)} = \max\limits_{i=1,2,\dots,m} f_i(x^{(0)})$. В соответствии с этим из неравенства (2.28) следует, что

$$\max_{i=1,2,\dots,m} f_i(x^{(0)}) \leqslant x_{n+1} \tag{2.29}$$

для всех допустимых точек задачи (2.25). Но точка $(x_1, x_2, \dots, x_n, \max\limits_{i=1,2,\dots,m} f_i(x))$ является допустимой для любого $x \in \mathbf{R}^n$. Для этой точки неравенство (2.29) принимает вид (2.26). Следовательно, $x^{(0)}$ — точка глобального минимума в задаче (2.24).

Прежде чем применить к задаче (2.25) теорему 2.5, заметим, что в любой точке ограничения этой задачи удовлетворяют условию регулярности; в качестве y можно взять, например, вектор $(0, 0, \dots, 0, 1) \in \mathbf{R}^{n+1}$.

Пусть $\left(x_1^{(0)}, \dots, x_n^{(0)}, x_{n+1}^{(0)}\right)$ — точка глобального минимума в задаче (2.25). Тогда, согласно теореме 2.5, существуют такие неотрицательные числа $\lambda_1, \lambda_2, \dots, \lambda_m$, что

$$(0, 0, \dots, 0, 1)^{\tau} + \sum_{i=1}^{m} \lambda_i \begin{pmatrix} \nabla f_i(x^{(0)}) \\ -1 \end{pmatrix} = 0_{n+1},$$

$$\lambda_i\left(f_i(x^{(0)}) - x_{n+1}^{(0)}\right) = \lambda_i (f_i(x^{(0)}) - f(x^{(0)})) = 0,\ i = 1, 2, \dots, m.$$

Заметим, что здесь рассмотрение ведется в пространстве $\mathbf{R}^{n+1}$. Первое из полученных равенств распадается на два равенства:

$$\sum_{i=1}^{m} \lambda_i \nabla f_i(x^{(0)}) = 0_n, \quad \sum_{i=1}^{m} \lambda_i = 1.$$

Окончательно приходим к следующему утверждению.

Следствие 2.1. *Для того чтобы точка* $x^{(0)}$ *была точкой безусловного глобального минимума в задаче* (2.24), *необходимо, что-*

бы нашлись такие числа $\lambda_1, \lambda_2, \ldots, \lambda_m$, *что*

$$\sum_{i=1}^{m} \lambda_i \nabla f_i(x^{(0)}) = 0_n,$$

$$\lambda_i (f_i(x^{(0)}) - f(x^{(0)})) = 0, \ i = 1, 2, \ldots, m, \tag{2.30}$$

$$\sum_{i=1}^{m} \lambda_i = 1, \ \lambda_i \geqslant 0, \ i = 1, 2, \ldots, m.$$

Нередко в задачах оптимизации среди ограничений встречаются ограничения-равенства. В принципе равенство $g_j(x) = 0$ можно заменить двумя неравенствами: $g_j(x) \leqslant 0$, $-g_j(x) \leqslant 0$. Однако если функция g_j нелинейная, то условие регулярности теоремы 2.5 требует, чтобы выполнялись два противоречащих друг другу неравенства: $\langle \nabla g_j(x^{(0)}), y \rangle > 0$, $\langle -\nabla g_j(x^{(0)}), y \rangle < 0$. Следовательно, непосредственно применять теорему 2.5 к задачам с нелинейными ограничениями-равенствами невозможно.

Рассмотрим задачу, в которой среди ограничений имеются как неравенства, так и равенства:

$$f(x) \to \min;$$

$$g_j(x) \leqslant 0, \ j = 1, 2, \ldots, k; \tag{2.31}$$

$$g_j(x) = 0, \ j = k+1, \ldots, m.$$

Для этой задачи справедливы следующие необходимые условия оптимальности.

Теорема 2.6. *Пусть функции* $f, g_1, g_2, \ldots, g_k$ *дифференцируемы в точке* $x^{(0)} \in \mathbf{R}^n$, *а функции* $g_{k+1}, \ldots, g_m$ *непрерывно дифференцируемы в этой точке. Будем предполагать выполненным расширенное условие регулярности: векторы* $\nabla g_j(x^{(0)})$, $j = k+1, \ldots, m$, *линейно независимы и существует такой вектор* $y \in \mathbf{R}^n$, *что* $\langle \nabla g_j(x^{(0)}), y \rangle = 0$, $j = k+1, \ldots, m$, а также $\langle \nabla g_j(x^{(0)}), y \rangle < 0$ *для всех индексов* $j \in \{1, 2, \ldots, k\}$ *активных ограничений в точке* $x^{(0)}$. *Для того чтобы точка* $x^{(0)}$ *была точкой локального минимума в задаче (2.31), необходимо, чтобы существовали такие множители Лагранжа* $\lambda_1, \lambda_2, \ldots, \lambda_m$, *что*

$$g_j(x^{(0)}) \leqslant 0, \ j = 1, 2, \ldots, k; \ g_j(x^{(0)}) = 0, \ j = k+1, \ldots, m;$$

$$\nabla f(x^{(0)}) + \sum_{j=1}^{m} \lambda_j \nabla g_j(x^{(0)}) = 0_n;$$

$$\lambda_j g_j(x^{(0)}) = 0, \ j = 1, 2, \ldots, k; \tag{2.32}$$

$$\lambda_j \geqslant 0, \ j = 1, 2, \ldots, k.$$

Заметим, что здесь множители Лагранжа, отвечающие ограничениям-равенствам, могут быть положительными, отрицательными или же равными нулю, тогда как множители, которые соответствуют ограничениям-неравенствам, отрицательными быть не должны.

В этой теореме не исключаются случаи, когда отсутствуют ограничения типа неравенств $g_j(x) \leqslant 0$ $(k=0)$ или типа равенств $g_j(x)=0$ $(m=k)$ или ограничения обоих типов $(m=0)$. При $k=0$ расширенное условие регулярности превращается в требование линейной независимости векторов $\nabla g_j(x^{(0)})$, $j=1, 2, \ldots, m$, поскольку равенства $\langle \nabla g_j(x^{(0)}), y \rangle = 0$, $j=1, 2, \ldots, m$, заведомо выполняются для $y=0_n$.

5. Достаточные условия минимума в задачах с ограничениями. Если все ограничения отсутствуют, то условия (2.32) превращаются в равенство $\nabla f(x^{(0)})=0_n$. Как отмечено в п. 2, точка $x^{(0)}$, удовлетворяющая этому равенству, не обязательно должна быть точкой безусловного глобального минимума функции f (она является точкой глобального минимума, если функция f псевдовыпуклая). Если функция f не является псевдовыпуклой, то условия (2.32) в частном случае $m=0$ не являются достаточными для оптимальности точки $x^{(0)}$, а значит, эти условия не будут достаточными и в общем случае $m \geqslant 0$. Более того, как показывает приведенный ниже пример, для того чтобы выполнение условий (2.32) гарантировало оптимальность точки $x^{(0)}$ в задаче (2.31), функции-ограничения $g_1, g_2, \ldots, g_m$ также должны удовлетворять определенным требованиям.

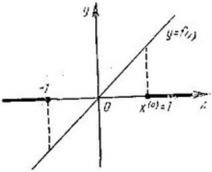

Рис. 2.9

Пусть $n=1$, $k=m=1$, $f(x)=x$, $g_1(x)=-x^2+1$. График данной линейной (а значит, псевдовыпуклой) функции $f(x)$ и соответствующее допустимое множество изображены на рис. 2.9. Рассмотрим точку $x^{(0)}=1$. Условия (2.32) при $\lambda_1=1/2$ для этой точки выполнены:

$$g_1(x^{(0)}) = -1+1 \leqslant 0,$$

$$\nabla f(x^{(0)}) + \lambda_1 \nabla g_1(x^{(0)}) = 1 + \frac{1}{2}(-2) = 0,$$

$$\lambda_1 g_1(x^{(0)}) = \frac{1}{2}(-1+1) = 0.$$

Однако она не является оптимальной. В этой задаче вообще не существует точки глобального минимума.

Достаточные условия глобального минимума применительно к общей задаче минимизации (2.31) сформулированы в следующем утверждении.

Теорема 2.7. *Предположим, что функция f псевдовыпукла, функции $g_1, g_2, \ldots, g_k$ квазивыпуклы на $\mathbf{R}^n$, а функции $g_{k+1}, \ldots, g_m$ линейны,* т. е.

$$g_j(x) = \langle a^{(j)}, x \rangle + b_j, \quad j=k+1, \ldots, m,$$

где $a^{(j)} \in \mathbf{R}^n$, $b_j \in \mathbf{R}$, $j=k+1, \ldots, m$. Если точка $x^{(0)}$ удовлетворяет со-

отношениям (2.32), *то она является точкой глобального минимума в задаче* (2.31).

□ Рассмотрим произвольную допустимую точку x, $x \neq x^{(0)}$, в задаче (2.31) и докажем неравенство $f(x^{(0)}) \leqslant f(x)$, где точка $x^{(0)}$ удовлетворяет (2.32).

Прежде всего заметим, что, поскольку соответствующие функции $g_1, g_2, \dots, g_m$ квазивыпуклы и линейны, допустимое множество в задаче (2.31) является выпуклым (см. п. 2 § 1.3). Значит, каждая точка отрезка, соединяющего точки x и $x^{(0)}$, должна удовлетворять ограничениям задачи (2.31).

Обозначим через J множество индексов активных ограничений-неравенств в точке $x^{(0)}$:

$$J = \{j \in \{1, 2, \dots, k\} \mid g_j(x^{(0)}) = 0\}.$$

На основании второго и третьего из соотношений (2.32) можно записать равенство

$$\nabla f(x^{(0)}) + \sum_{j \in J} \lambda_j \nabla g_j(x^{(0)}) + \sum_{j=k+1}^{m} \lambda_j a^{(j)} = 0_n.$$

Умножая это равенство скалярно на $x - x^{(0)}$, получаем

$$\langle \nabla f(x^{(0)}),\ x - x^{(0)} \rangle = -\sum_{j \in J} \lambda_j \langle \nabla g_j(x^{(0)}),\ x - x^{(0)} \rangle -$$

$$- \sum_{j=k+1}^{m} \lambda_j \langle a^{(j)},\ x - x^{(0)} \rangle. \tag{2.33}$$

Так как верны равенства $\langle a^{(j)}, x \rangle + b_j = 0$ и $\langle a^{(j)}, x^{(0)} \rangle + b_j = 0$, то выполняется и равенство $\langle a^{(j)}, x - x^{(0)} \rangle = \langle a^{(j)}, x \rangle - \langle a^{(j)}, x^{(0)} \rangle = 0$, $j = k+1, \dots, m$. Поэтому вторая сумма в правой части равенства (2.33) равна нулю. Далее, для каждого номера $j \in J$ имеет место неравенство

$$\langle \nabla g_j(x^{(0)}),\ x - x^{(0)} \rangle \leqslant 0. \tag{2.34}$$

В самом деле, если это не так, то для некоторого $j = j_0 \in J$ выполняется неравенство $\langle \nabla g_{j_0}(x^{(0)}),\ x - x^{(0)} \rangle > 0$. Отсюда согласно лемме 2.1 следует, что для некоторого числа $\tau \in (0, 1)$ окажется справедливым неравенство $g_{j_0}(x^{(0)} + \tau(x - x)^{(0)}) > g_{j_0}(x^{(0)})$. Но $x^{(0)} + \tau(x - x^{(0)}) = \tau x + (1 - \tau)\ x^{(0)}$ и $g_{j_0}(x^{(0)}) = 0$. Поэтому получаем неравенство

$$g_{j_0}(\tau x + (1 - \tau)\, x^{(0)}) > 0.$$

Каждая точка отрезка, соединяющего x и $x^{(0)}$, должна удовлетворять всем ограничениям задачи (2.31), тогда как последнее неравенство противоречит этому. Таким образом, неравенство (2.34) верно для любого номера $j \in J$. В соответствии с этим правая часть равенства (2.33) не отрицательна. Тогда неотрицательной является и левая часть: $\langle \nabla f(x^{(0)}),\ x-x^{(0)}\rangle \geqslant 0$. Функция f псевдовыпуклая, поэтому отсюда следует требуемое неравенство $f(x) \geqslant f(x^{(0)})$. ■

Доказанную теорему можно использовать и для установления достаточных условий оптимальности применительно к минимаксной задаче (2.24), учитывая при этом, что она эквивалентна задаче (2.25).

Следствие 2.2. *Предположим, что все функции* $f_i(x)-x_{n+1}$, $i=1, 2, \ldots, m$, *квазивыпуклы на* $\mathbf{R}^{n+1}$ *(это будет заведомо выполнено, когда все функции* f_i *выпуклы на* $\mathbf{R}^n$*). Если точка* $x^{(0)}$ *удовлетворяет соотношениям* (2.30), *то она является точкой глобального минимума в задаче* (2.24).

6. Замечания. Понятия и результаты, изложенные в этом параграфе, имеют принципиально важное значение как с теоретической, так и с практической точки зрения. Они относятся к широкому классу оптимизационных задач, в которых целевая функция и функции, входящие в ограничения, являются дифференцируемыми. Необходимые условия дают возможность «выбраковывать» среди множества допустимых решений те, которые заведомо не могут быть оптимальными. На основании необходимых условий также строятся различные численные методы решения задач оптимизации. С помощью достаточных условий проверяют, действительно ли найденное тем или иным методом решение является оптимальным.

§ 2.6. ВЫПУКЛОЕ ПРОГРАММИРОВАНИЕ. ДВОЙСТВЕННЫЕ ЗАДАЧИ

1. Условия оптимальности в терминах седловых точек функции Лагранжа. Если целевая функция и ограничения являются одновременно выпуклыми и дифференцируемыми, то соответствующие условия оптимальности в этом случае можно записать на основе теорем 2.5—2.7. В этом пункте будут сформулированы необходимые и достаточные условия оптимальности без предположения дифференцируемости указанных функций. Важным примером выпуклых, вообще говоря, недифференцируемых функций являются функции вида

$$h(x) = \max_{i=1,2,\ldots,m} h_i(x),$$

где все функции h_i предполагаются выпуклыми (в частности, линейными).

Рассмотрим следующую задачу выпуклого программирования:

$$f(x) \to \min;\ g_j(x) \leqslant 0,\ j=1,\ 2,\ldots,k;\ x \in D \subset \mathbf{R}^n, \qquad (2.35)$$

где функции f, g_1, g_2, ..., g_k предполагаются выпуклыми на некотором выпуклом множестве D. Эта задача отличается от рассмотренных ранее задач условием $x \in D$, которое добавлено из соображений удобства. В практических задачах, как правило, $D=\mathbf{R}^n$ или $D=\{x\in\mathbf{R}^n \mid x\geqq 0_n\}$.

Задаче (2.35) соответствует функция Лагранжа вида

$$L(x, \lambda)=f(x)+\sum_{j=1}^{k}\lambda_j g_j(x), \tag{2.36}$$

которую будем рассматривать при неотрицательных λ_j:

$$\lambda=(\lambda_1, \lambda_2, \ldots, \lambda_k)^{\mathrm{T}}\in\Lambda=\{\lambda\in\mathbf{R}^k \mid \lambda\geqq 0_k\}$$

и при $x\in D$.

Далее будет показано, что решения задачи (2.35) тесно связаны с седловыми точками функции Лагранжа.

Говорят, что пара векторов $(x^{(0)}, \lambda^{(0)})$, где $x^{(0)}\in D$, $\lambda^{(0)}\in\Lambda$, образуют *седловую точку* функции Лагранжа (2.36), если при всех $x\in D$ и $\lambda\in\Lambda$ выполняются неравенства

$$L(x^{(0)}, \lambda)\leqslant L(x^{(0)}, \lambda^{(0)})\leqslant \leqslant L(x, \lambda^{(0)}). \tag{2.37}$$

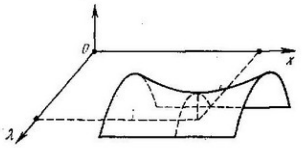

Рис. 2.10

График некоторой функции, имеющей седловую точку, где x и λ — скаляры, изображен на рис. 2.10. Седловая точка функции является результатом ее минимизации по x и максимизации по λ.

Используя конкретный вид функции Лагранжа (2.36), можно сформулировать и доказать следующее утверждение.

Лемма 2.2. *Пара векторов* $(x^{(0)}, \lambda^{(0)})$ *при* $x^{(0)}\in D$, $\lambda^{(0)}\in\Lambda$ *является седловой точкой функции Лагранжа* (2.36) *тогда и только тогда, когда выполнено*

$$L(x^{(0)}, \lambda^{(0)})=\min_{x\in D} L(x, \lambda^{(0)}), \tag{2.38}$$

$$g_j(x^{(0)})\leqslant 0,\ \lambda_j^{(0)} g_j(x^{(0)})=0,\ j=1, 2, \ldots, k. \tag{2.39}$$

□ Достаточность. Пусть выполнены соотношения (2.38), (2.39). Равенство (2.38) влечет справедливость правой части неравенства (2.37) при всех $x\in D$. Докажем выполнение левой части неравенства (2.37) для всех $\lambda\in\Lambda$. Для этого рассмотрим произвольные неотрицательные числа $\lambda_1, \lambda_2, \ldots, \lambda_k$. Очевидно, $\lambda_j g_j(x^{(0)})\leqslant 0$, $j=1, 2, \ldots, k$. Поэтому, используя равенства из (2.39), можно записать неравенства $\lambda_j g_j(x^{(0)})\leqslant \lambda_j^{(0)} g_j(x^{(0)})$, $j=1, 2, \ldots, k$. Следовательно,

$$f(x^{(0)})+\sum_{j=1}^{k}\lambda_j g_j(x^{(0)})\leqslant f(x^{(0)})+\sum_{j=1}^{k}\lambda_j^{(0)} g_j(x^{(0)}),$$

что означает справедливость левой части неравенства (2.37).

Необходимость. Пусть $(x^{(0)}, \lambda^{(0)})$ — седловая точка. Выполнение правой части неравенства (2.37) для всех $x \in D$ означает справедливость равенства (2.38). Проверим справедливость соотношений (2.39).

Если для некоторого номера j верно неравенство $g_j(x^{(0)}) > 0$, то, полагая числа λ_i, $i = 1, 2, \ldots, k$; $i \neq j$ нулевыми, а λ_j — достаточно большим, можно получить настолько большее значение $L(x^{(0)}, \lambda)$, что левая часть неравенства (2.37) не будет справедливой. Следовательно, $g_j(x^{(0)}) \leqslant 0$, $j = 1, 2, \ldots, k$.

Поскольку $\lambda_j^{(0)} \geqslant 0$ и $g_j(x^{(0)}) \leqslant 0$, $j = 1, 2, \ldots, k$, имеем неравенство

$$\sum_{j=1}^{k} \lambda_j^{(0)} g_j(x^{(0)}) \leqslant 0.$$

С другой стороны, левая часть неравенства (2.37) при $\lambda = 0_k$ принимает вид

$$0 \leqslant \sum_{j=1}^{k} \lambda_j^{(0)} g_j(x^{(0)}).$$

Поэтому имеет место равенство

$$\sum_{j=1}^{k} \lambda_j^{(0)} g_j(x^{(0)}) = 0.$$

Все слагаемые в этом неравенстве не положительны. Следовательно, сумма может обращаться в нуль лишь в том случае, когда каждое из слагаемых равно нулю. ■

Пусть $(x^{(0)}, \lambda^{(0)})$ — некоторая седловая точка функции Лагранжа. Используя равенства (2.39), равенство (2.38) можно записать в виде неравенства

$$f(x^{(0)}) \leqslant f(x) + \sum_{j=1}^{k} \lambda_j^{(0)} g_j(x) \text{ для всех } x \in D. \tag{2.40}$$

Для всех допустимых точек x в задаче (2.35), т. е. для всех x, удовлетворяющих включению $x \in D$ и неравенствам $g_j(x) \leqslant 0$, $j = 1, 2, \ldots, k$, справедливо неравенство $\sum_{j=1}^{k} \lambda_j^{(0)} g_j(x) \leqslant 0$. Поэтому из неравенства (2.40) при всех допустимых x получаем $f(x^{(0)}) \leqslant f(x)$. Это означает, что $x^{(0)}$ — точка глобального минимума в задаче (2.35). Тем самым установлены следующие достаточные условия оптимальности.

Теорема 2.8. *Если пара векторов* $(x^{(0)}, \lambda^{(0)})$, *где* $x^{(0)} \in D$, $\lambda^{(0)} \in \Lambda$, *является седловой точкой функции Лагранжа* (2.36), *то* $x^{(0)}$ — *точка глобального минимума в задаче* (2.35).

Необходимо отметить, что лемма 2.2 и теорема 2.8 получены без каких-либо предположений о свойствах функций $f, g_1, g_2, \ldots, g_k$ и структуре множества D.

В соответствии с теоремой 2.8, если удастся найти седловую точку $(x^{(0)}, \lambda^{(0)})$ функции Лагранжа (2.36), тем самым будет решена задача (2.35), в которой все функции $f, g_1, g_2, \ldots, g_k$ и множество D могут иметь различную природу (в частности, могут быть и невыпуклыми), причем $x^{(0)}$ будет точкой, доставляющей глобальный минимум. Теоретически такой подход безупречен. Однако его практическая реализация возможна не всегда. Дело в том, что если функции $f, g_1, g_2, \ldots, g_k$ и множество D не выпуклы, то функция Лагранжа часто не имеет ни одной седловой точки. Более того, даже если указанные функции и множество выпуклы и оптимальное решение задачи (2.35) существует, то в некоторых «вырожденных» случаях функция Лагранжа также может не обладать ни одной седловой точкой.

Для подтверждения этого положения рассмотрим следующую задачу, в которой $n=1$, $k=1$:

$-x \to \min$; $x^2 \leqslant 0$, $D=\{x \in \mathbf{R} \mid x \geqslant 0\}$.

Здесь функции $f(x)=-x$, $g_1(x)=x^2$ выпуклы. Множество D также выпукло. Существует единственная точка $x^{(0)}=0$, удовлетворяющая ограничениям задачи; она же является и точкой глобального минимума. Если седловая точка функции Лагранжа $L(x, \lambda)=$ $=-x+\lambda x^2$ при $x \geqslant 0$, $\lambda \geqslant 0$ существует, то она должна иметь вид $(0, \lambda^{(0)})$ и удовлетворять соотношениям (2.38) и (2.39), т. е. неравенство $0 \leqslant -x+\lambda^{(0)}x^2$ должно выполняться для всех $x \geqslant 0$. Следовательно, неравенство $\lambda^{(0)} \geqslant 1/x$ должно иметь место для любого (сколь угодно малого) положительного числа x. Легко видеть, что требуемое число $\lambda^{(0)}$ найти нельзя.

Условия, при выполнении которых хотя бы одна седловая точка функции Лагранжа всегда существует, сформулированы в следующем утверждении.

Теорема 2.9. *Предположим, что множество* $D \subset \mathbf{R}^n$ *выпукло, функции* $f, g_1, g_2, \ldots g_k$ *выпуклы на нем и имеет место условие Слейтера: найдется такая точка* $\bar{x} \in D$, *что* $g_j(\bar{x})<0$, $j=1, 2, \ldots, k$. Пусть $x^{(0)}$ — *точка глобального минимума в задаче* (2.35). *Тогда найдется такой вектор* $\lambda^{(0)} \in \Lambda$, *что пара* $(x^{(0)}, \lambda^{(0)})$ *образует седловую точку функции Лагранжа* (2.36).

Эта теорема (ее доказательство можно найти в [8]) представляет собой необходимое условие оптимальности: для того чтобы точка $x^{(0)}$ была точкой глобального минимума в задаче выпуклого программирования (2.35), необходимо, чтобы нашелся такой вектор $\lambda^{(0)} \in \Lambda$, что пара $(x^{(0)}, \lambda^{(0)})$ образует седловую точку функции Лагранжа.

Условие Слейтера исключает присутствие в задаче ограничений типа равенств. Это условие означает, что среди точек множества D найдется хотя бы одна, в которой все ограничения являются неактивными.

Теоремы 2.8 и 2.9 лежат в основе теории двойственности выпуклого программирования, элементы которой будут рассмотрены ниже. Кроме того, они находят применение в численных методах решения задач условной оптимизации. С их помощью исходную задачу (2.35) можно заменить задачей отыскания седловой точки функции Лагранжа.

2. Двойственность в выпуклом программировании. Вернемся к задаче (2.35). Зафиксируем точку $x \in D$ и введем функцию, которая может принимать и бесконечное значение:

$$L^*(x) = \sup_{\lambda \in \Lambda} \left[f(x) + \sum_{j=1}^{k} \lambda_j g_j(x) \right].$$

Имеем

$$L^*(x) = \begin{cases} f(x), & \text{если } g_j(x) \leqslant 0, \ j = 1, 2, \dots, k, \\ +\infty, & \text{если хотя бы для одного } j \text{ верно неравенство} \\ & \quad g_j(x) > 0. \end{cases}$$

Следовательно, исходная задача (2.35) равносильна задаче минимизации функции $L^*(x)$ по $x \in D$, т. е. задаче

$$\sup_{\lambda \in \Lambda} L_*(x, \lambda) \to \min_{x \in D}. \tag{2.41}$$

Эту равносильность можно записать и так:

$$\min_{x \in D} \sup_{\lambda \in \Lambda} L(x, \lambda) = \min_{\substack{x \in D \\ g_j(x) \leqslant 0,\ j=1,2,\dots,k}} f(x).$$

Зафиксируем теперь некоторый вектор $\lambda \in \Lambda$ и введем функцию, которая также может принимать бесконечное значение:

$$L_*(\lambda) = \inf_{x \in D} \left[f(x) + \sum_{j=1}^{k} \lambda_j g_j(x) \right].$$

Задача максимизации функции $L_*(\lambda)$ по $\lambda \in \Lambda$, т. е. задача

$$\inf_{x \in D} L_*(x, \lambda) \to \max_{\lambda \in \Lambda}, \tag{2.42}$$

называется *двойственной*, а задача (2.41) — *прямой*. Между этими двумя задачами существует тесная связь; установим ее.

По определению функций $L^*(x)$ и $L_*(\lambda)$, справедливы неравенства $L_*(\lambda) \leqslant L(x, \lambda) \leqslant L^*(x)$. Поэтому

$$L_*(\lambda) \leqslant L^*(x) \text{ для всех } \lambda \in \Lambda,\ x \in D. \tag{2.43}$$

Предполагая существование оптимальных решений прямой и двойственной задач, получаем отсюда неравенство

$$\max_{\lambda \in \Lambda} L_*(\lambda) \leqslant \min_{x \in D} L^*(x),$$

или, что то же самое, неравенство

$$\max_{\lambda \in \Lambda} \inf_{x \in D} L(x, \lambda) \leqslant \min_{x \in D} \sup_{\lambda \in \Lambda} L(x, \lambda). \tag{2.44}$$

Здесь в правой и левой частях неравенства находятся оптимальные значения соответственно прямой и двойственной задач. Неравенство (2.44) означает, что *оптимальное значение прямой задачи не может быть меньше, чем оптимальное значение двойственной задачи,* т. е. оптимальное значение двойственной задачи дает оценку снизу для оптимального значения исходной задачи (2.35). Отметим, что неравенство (2.44) было получено без использования каких-либо специальных свойств функций f, g_1, $g_2, \ldots, g_k$ и множества D.

Если оптимальные значения прямой и двойственной задач совпадают и двойственная задача в вычислительном отношении оказывается проще прямой, то можно вместо прямой решить более простую двойственную задачу и тем самым найти оптимальное значение исходной задачи (2.35). Возникает вопрос о том, при каких условиях значения обеих задач совпадают. Ответ на него дает следующая теорема.

Теорема 2.10. *Пусть функция Лагранжа* $L(x, \lambda)$ *вида* (2.36) *имеет седловую точку* $(x^{(0)}, \lambda^{(0)})$. *Тогда оптимальные решения прямой и двойственной задач существуют, причем оптимальные значения обеих задач совпадают и равны* $L(x^{(0)}, \lambda^{(0)})$, *т. е.*

$$\max_{\lambda\in\Lambda} \inf_{x\in D} L(x, \lambda) = \min_{x\in D} \sup_{\lambda\in\Lambda} L(x, \lambda) = L(x^{(0)}, \lambda^{(0)}). \tag{2.45}$$

□ Существование оптимального решения прямой задачи следует из существования седловой точки, теоремы 2.8 и равносильности прямой задачи и задачи (2.35).

Установим наличие оптимального решения двойственной задачи. Так как $(x^{(0)}, \lambda^{(0)})$ — седловая точка, то имеют место неравенства (2.37). Переходя к нижней грани по $x\in D$ в правой части неравенства (2.37), получаем $L(x^{(0)}, \lambda^{(0)}) \leqslant \inf\limits_{x\in D} L(x, \lambda^{(0)})$. Но по определению инфимума верно неравенство $L(x^{(0)}, \lambda^{(0)}) \geqslant \inf\limits_{x\in D} L(x, \lambda^{(0)})$. Поэтому

$$L(x^{(0)}, \lambda^{(0)}) = \inf_{x\in D} L(x, \lambda^{(0)}).$$

Далее, из левой части неравенства (2.37) получаем

$$L(x^{(0)}, \lambda^{(0)}) \geqslant L(x^{(0)}, \lambda) \geqslant \inf_{x\in D} L(x, \lambda) \text{ для всех } \lambda \in \Lambda.$$

Следовательно,

$$L(x^{(0)}, \lambda^{(0)}) = \inf_{x\in D} L(x, \lambda^{(0)}) = \max_{\lambda\in\Lambda} \inf_{x\in D} L(x, \lambda),$$

т. е. $\lambda^{(0)}$ — искомая точка максимума функции $L_*(\lambda)$ на множестве Λ.

Остается проверить равенства (2.45). Переходя в неравенствах (2.37) к нижней грани по $x\in D$ и к верхней грани по $\lambda\in\Lambda$, имеем

$$\sup_{\lambda\in\Lambda} L(x^{(0)}, \lambda) \leqslant \inf_{x\in D} L(x, \lambda^{(0)}).$$

Поэтому и

$$\min_{x\in D}\sup_{\lambda\in\Lambda} L(x, \lambda) \leqslant \max_{\lambda\in\Lambda}\inf_{x\in D} L(x, \lambda) = L(x^{(0)}, \lambda^{(0)}).$$

Сравнивая этот результат с неравенством (2.44), приходим к требуемым равенствам (2.45). ■

Таким образом, существование седловой точки функции Лагранжа гарантирует равенство оптимальных значений прямой и двойственной задач. В свою очередь, как показывает теорема 2.9, седловая точка заведомо существует, если функции f, g_1, g_2, ..., g_k выпуклы, множество D также выпукло, выполнено условие Слейтера и задача (2.35) имеет оптимальное решение.

3. Двойственность в линейном программировании. Напомним формулировку *стандартной задачи линейного программирования:*

$$\langle c, x\rangle \to \min;$$

$$\langle a^{(j)}, x\rangle \geqslant b_j,\ j=1, 2,\dots, k;\ x \geqq 0_n,$$

где c, $a^{(j)}\in\mathbf{R}^n$, $b_j\in\mathbf{R}$, $j=1, 2, \dots, k$. Если ввести вектор-столбец b с компонентами b_1, b_2, ..., b_k и матрицу A, строками которой служат компоненты соответствующих векторов $a^{(1)}$, $a^{(2)}$, ..., $a^{(k)}$, то стандартную задачу линейного программирования можно записать в следующем матричном виде:

$$\langle c, x\rangle \to \min;\ Ax \geqq b,\ x\in D_0, \tag{2.46}$$

где $D_0=\{x\in\mathbf{R}^n \mid x\geqq 0_n\}$. Этой задаче соответствует функция Лагранжа вида

$$L(x, \lambda)=\langle c, x\rangle+\langle\lambda, b-Ax\rangle=\langle\lambda, b\rangle+\langle c-A^{\mathrm{T}}\lambda, x\rangle$$

Отсюда в соответствии с определением функции $L_*(\lambda)$ имеем

$$L_*(\lambda)=\inf_{x\in D_0} L(x, \lambda)=\begin{cases}\langle\lambda, b\rangle, & \text{если } c-A^{\mathrm{T}}\lambda\geqq 0_n,\\ -\infty, & \text{если хотя бы одна компонента}\\ & \text{вектора } c-A^{\mathrm{T}}\lambda \text{ отрицательна.}\end{cases}$$

Точку $\lambda^{(0)}\in\Lambda$, в которой достигается максимум функции $L_*(\lambda)$ на множестве Λ, можно искать лишь среди тех $\lambda\in\Lambda$, для которых $c-A^{\mathrm{T}}\lambda\geqq 0_n$. Поэтому двойственную по отношению к (2.46) задачу можно записать в виде

$$\langle b, \lambda\rangle \to \max;\ A^{\mathrm{T}}\lambda \leqq c,\ \lambda\in\Lambda. \tag{2.47}$$

Это также задача линейного программирования, причем переменная x в ней отсутствует.

Задачу (2.46) называют *прямой задачей* линейного программирования, а задачу (2.47) — соответствующей *двойственной задачей* линейного программирования. Переменную x называют *прямой переменной*, а λ — *двойственной переменной.*

Прямая задача включает n переменных x_1, x_2, ..., x_n и k ограничений-неравенств $Ax\geqq b$, тогда как в двойственной задаче k переменных λ_1, λ_2, ..., λ_k и n ограничений-неравенств $A^{\mathrm{T}}\lambda\leqq c$.

Интересно, что если для задачи (2.47), в свою очередь, записать двойственную ей задачу, то придем к исходной задаче (2.46). Чтобы убедиться в этом, рассмотрим эквивалентную задачу

$$-\langle b, \lambda\rangle \to \min; \ -A^{\tau}\lambda \geqq -c, \ \lambda \geqq 0_k,$$

целевая функция которой отличается от целевой функции в (2.47) только знаком. Эта задача по форме такая же, что и задача (2.46). Поэтому двойственная по отношению к ней задача имеет следующий вид (при этом вектор двойственных переменных обозначаем через x):

$$-\langle c, x\rangle \to \max; \ (-A^{\tau})^{\tau} x \leqq -b, \ x \geqq 0_n.$$

Ограничения здесь легко преобразовать к виду $Ax \geqq b$, $x \in D_0$, а значит, последняя задача отличается от задачи (2.46) только знаком перед целевой функцией. Следовательно, двойственной по отношению к задаче (2.47) является именно задача (2.46).

Оказывается, для задачи линейного программирования теорема 2.9 верна и без условия Слейтера. Кроме того, для задачи линейного программирования справедливо более сильное утверждение, чем теорема 2.10, а именно справедлива следующая теорема, доказательство которой приведено, например, в [8].

Теорема 2.11. *Задачи* (2.46) и (2.47) *либо обе имеют оптимальные решения, либо его не имеют, причем в первом случае их оптимальные значения совпадают:*

$$\langle c, x^{(0)}\rangle = \langle b, \lambda^{(0)}\rangle,$$

где $x^{(0)}$ *— оптимальное решение задачи* (2.46), *а* $\lambda^{(0)}$ *— оптимальное решение задачи* (2.47).

Часто имеет место *каноническая задача линейного программирования* (в матричной записи):

$$\langle c, x\rangle \to \min; \ Ax = b, \ x \in D_0. \tag{2.48}$$

Запишем двойственную ей задачу. Векторное равенство $Ax=b$ всегда можно представить в форме неравенства

$$\begin{pmatrix} A & 0 \\ 0 & A \end{pmatrix} \begin{pmatrix} x \\ x \end{pmatrix} \geqq \begin{pmatrix} b \\ -b \end{pmatrix},$$

поэтому в данном случае функция Лагранжа имеет вид

$$L(x, \lambda', \lambda'') = \langle c, x\rangle + \langle \lambda', b - Ax\rangle - \langle \lambda'', b - Ax\rangle,$$

где $\lambda' \geqq 0_k$, $\lambda'' \geqq 0_k$. Вводя вектор $\lambda = \lambda' - \lambda''$, компонентами которого могут быть произвольные числа, получаем

$$L(x, \lambda) = \langle c, x\rangle + \langle \lambda, b - Ax\rangle = \langle \lambda, b\rangle + \langle c - A^{\tau}\lambda, x\rangle.$$

Поэтому

$$L_*(\lambda) = \inf_{x \in D_0} L(x, \lambda) = \begin{cases} \langle \lambda, b\rangle, & \text{если } c - A^{\tau}\lambda \geqq 0_n, \\ -\infty, & \text{если хотя бы одна компонента вектора } c - A^{\tau}\lambda \text{ отрицательна.} \end{cases}$$

В соответствии с этим двойственная по отношению к (2.48) задача записывается следующим образом:

$$\langle b, \lambda\rangle \to \min; \ A^{\tau}\lambda \leq c, \ \lambda \in \mathbf{R}^k. \qquad (2.49)$$

Эта задача отличается от задачи (2.47) лишь тем, что здесь переменные $\lambda_1, \lambda_2, \ldots, \lambda_k$ могут быть любыми числами, тогда как в задаче (2.47) они не отрицательны. Можно проверить, что двойственной для задачи (2.49) является исходная задача (2.48).

Для пары задач (2.48) и (2.49) теорема 2.11 также имеет место.

§ 2.7. ПРИКЛАДНЫЕ ЗАДАЧИ

1. Задача о распределении поисковых усилий. Пусть в заданной области производится поиск находящегося в ней объекта. Это может быть поиск месторождения полезных ископаемых, самолета или корабля или же поиск точки минимума функции n переменных. В соответствии с этим областью поиска являются части суши, моря или подмножество пространства $\mathbf{R}^n$. Вся область поиска разбита на m частей $X_1, X_2, \ldots, X_m$. Предполагается известной вероятность p_i обнаружения объекта в частичной области X_i, $i=1, 2, \ldots, m$. Во многих практических задачах значения вероятностей $p_1, p_2, \ldots, p_m$ могут быть получены на основе косвенных сведений о положении искомого объекта. Например, если речь идет о самолете, то для определения p_i может быть использована информация о его координатах во время последнего сеанса связи, его скорости, курсе, времени, прошедшем с момента связи, и т. п.

Поиск сопровождается затратами, которые обычно связаны с ресурсами времени и т. п. Считается, что поиск во всех областях $X_1, X_2, \ldots, X_m$ должен быть начат одновременно, так что поисковые усилия следует каким-то образом распределить по указанным областям. Эти поисковые усилия будем характеризовать соответствующими числами $q_1, q_2, \ldots, q_m$, причем $q_i \geqslant 0$, $\sum_{i=1}^{m} q_i = 1$.

Сформируем целевую функцию переменных $q_1, q_2, \ldots, q_m$. Если вероятность обнаружения объекта в некоторой частичной области есть p, то в результате k независимых попыток вероятность обнаружения составит

$$1-(1-p)^k = 1 - a\,e^{-bk},$$

где $a=1$ и $b=-\ln(1-p)$. Это дает основание вероятность обнаружения объекта в области X_i при условии, что он находится в данной области, выразить следующей функцией усилия q_i: $1-a_i\,e^{-b_i q_i}$, где a_i, b_i — фиксированные положительные числа. С помощью коэффициентов a_i и b_i можно в выражении для вероятности учитывать дополнительную информацию (например, более труднодоступным обследованию областям придавать меньшие значения b_i или же бо́льшие значения a_i). По формуле полной вероятности вероятность обнаружения объекта в исходной области равна

$$P(q_1, q_2, \ldots, q_m) = \sum_{i=1}^{m} p_i (1 - a_i e^{-b_i q_i}).$$

Поэтому математическая модель определения оптимального распределения поисковых усилий принимает вид следующей задачи нелинейного программирования:

$$P(q_1, q_2, \ldots, q_m) \to \max;$$

$$\sum_{i=1}^{m} q_i = 1, \ q_i \geqslant 0, \ i = 1, 2, \ldots, m.$$

Целевая функция представляет собой линейную комбинацию вогнутых функций одной переменной, а значит, сама является вогнутой функцией (см. п. 2 § 1.3). Таким образом, для решения задачи достаточно найти стационарную точку. Из ограничения-равенства выразим, например q_1 через остальные переменные и подставим полученное выражение в целевую функцию. Освободившись таким образом от ограничения-равенства и приравняв частные производные целевой функции нулю, получим

$$\frac{\partial P}{\partial q_2} = -p_1 a_1 b_1 e^{b_1 (q_2 + \ldots + q_m - 1)} + p_2 a_2 b_2 e^{-b_2 q_2} = 0,$$

$$\frac{\partial P}{\partial q_3} = -p_1 a_1 b_1 e^{b_1 (q_2 + \ldots + q_m - 1)} + p_3 a_3 b_3 e^{-b_3 q_3} = 0,$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\frac{\partial P}{\partial q_m} = -p_1 a_1 b_1 e^{b_1 (q_2 + \ldots + q_m - 1)} + p_m a_m b_m e^{-b_m q_m} = 0.$$

Эту систему равенств нетрудно преобразовать к следующей системе линейных относительно $q_2, q_3, \ldots, q_m$ уравнений:

$$\begin{cases} b_1 (q_2 + \ldots + q_m - 1) + b_2 q_2 = \ln \dfrac{p_2 a_2 b_2}{p_1 a_1 b_1}, \\ b_1 (q_2 + \ldots + q_m - 1) + b_3 q_3 = \ln \dfrac{p_3 a_3 b_3}{p_1 a_1 b_1}, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ b_1 (q_2 + \ldots + q_m - 1) + b_m q_m = \ln \dfrac{p_m a_m b_m}{p_1 a_1 b_1}. \end{cases}$$

Неотрицательное решение $q_2^*, q_3^*, \ldots, q_m^*$ этой системы вместе с $q_1^* = 1 - q_2^* - \ldots - q_m^* \geqslant 0$ образует искомое оптимальное распределение поисковых усилий.

2. Задача о выборе сечения проводников. Металлический проводник кругового сечения AB длиной l_0 разветвляется в точке B на n проводников BC_1, BC_2, ..., BC_n кругового сечения (изготовленных из того же материала, что и AB) с длинами l_1, l_2, ..., l_n. Требуется, чтобы сила тока в проводнике AB была равна I_0, а в проводниках BC_1, BC_2, ..., BC_n — соответственно I_1, I_2, ..., I_n. Какого сечения следует взять проводники, чтобы расход материала был минимально возможным, а разность потенциалов на каждом из участков ABC_j, $j=1, 2, ..., n$, не превышала заданного значения U?

Обозначим через x_i площадь поперечного сечения i-го проводника, $i=0, 1, ..., n$. Масса материала проводников с точностью до константы равна объему: $\sum_{i=0}^{n} l_i x_i$. На участке ABC_j, согласно закону Ома, разность потенциалов составляет

$$I_0(\rho l_0/x_0)+I_j(\rho l_j/x_j), \quad j=1, 2,..., n,$$

где ρ — удельное сопротивление материала. Следовательно, задача оптимизации имеет вид

$$\sum_{i=0}^{n} l_i x_i \to \min;$$

$$I_0(\rho l_0/x_0)+I_j(\rho l_j/x_j)-U \leqslant 0, \quad j=1, 2,..., n,$$

при дополнительном условии положительности всех переменных x_0, x_1, ..., x_n. Здесь целевая функция линейная, а все функции-ограничения выпуклые как линейные комбинации с положительными коэффициентами выпуклых функций $1/x_0$ и $1/x_j$ при $x_0, x_j>0$ (см. п. 2 § 1.3). Таким образом, имеем задачу выпуклого программирования. Составим функцию Лагранжа

$$\sum_{i=0}^{n} l_i x_i+\sum_{j=1}^{n} \lambda_j [I_0(\rho l_0/x_0)+I_j(\rho l_j/x_j)-U]$$

и запишем для нее соотношения типа (2.22)—(2.23):

$$l_0-I_0\rho l_0 \sum_{j=1}^{n} \lambda_j/x_0^2=0, \tag{2.50}$$

$$l_j-\lambda_j I_j l_j \rho/x_j^2=0, \quad j=1, 2,..., n, \tag{2.51}$$

$$\lambda_j[I_0(\rho l_0/x_0)+I_j(\rho l_j/x_j)-U]=0, \quad j=1, 2,..., n. \tag{2.52}$$

Из равенств (2.51) в силу положительности l_j следует, что $\lambda_j>0$, $j=1, 2, ..., n$. Таким образом, из равенств (2.52) получаем

$$x_j=I_j l_j/(U/\rho-I_0 l_0/x_0), \quad j=1, 2,..., n. \tag{2.53}$$

На основании (2.53) из (2.51) находим

$$\lambda_j = x_j^2/(I_j\rho) = I_j l_j^2/(U - I_0 l_0 \rho/x_0)^2, \quad j = 1, 2, \ldots, n.$$

Подставляя найденное выражение в (2.50), определяем

$$x_0 = \rho\left(\sqrt{I_0 \sum_{j=1}^{n} I_j l_j^2} + I_0 l_0\right) \Big/ U;$$

далее, используя формулу (2.53), находим

$$x_j = \rho I_j l_j \Big/ \left[U - I_0 l_0 U \Big/ \left(\sqrt{I_0 \sum_{j=1}^{n} I_j l_j^2} + I_0 l_0\right)\right], \quad j = 1, 2, \ldots, n.$$

Согласно теореме 2.7, найденное решение $x_0, x_1, \ldots, x_n$ является оптимальным.

3. Задача о пропускной способности канала связи. С помощью передающего устройства по каналу связи посылают в определенной последовательности сигналы, для обозначения которых будем использовать символы $a_1, a_2, \ldots, a_n$. Известно, что каждый символ a_i последовательности выбирают независимо от других с вероятностью $P(a_i)$, $i=1, 2, \ldots, n$. На другом конце канала связи имеется приемное устройство. Из-за помех в канале связи приемник принимает, вообще говоря, другие символы $b_1, b_2, \ldots, b_m$. Качество передачи символов по каналу связи характеризуется заданными величинами $P(b_j|a_i)$ — условной вероятностью приема символа b_j при условии, что был послан символ a_i, $j=1, 2, \ldots, m$; $i=1, 2, \ldots, n$. При передаче символа a_i какой-то символ из набора $b_1, b_2, \ldots, b_m$ обязательно поступит, поэтому должно выполняться равенство

$$\sum_{j=1}^{m} P(b_j \mid a_i) = 1, \quad i = 1, 2, \ldots, n.$$

Количество информации, которое несет символ a_i с вероятностью появления $P(a_i)$, определяется величиной $-\log_2 P(a_i)$. Тогда *среднее количество информации, посылаемое по каналу связи,* составит

$$H(a) = -\sum_{i=1}^{n} P(a_i) \log_2 P(a_i).$$

Аналогично, если $P(a_i|b_j)$ — условная вероятность передачи символа a_i при условии, что был принят символ b_j, то величину

$$H(a \mid b_j) = -\sum_{i=1}^{n} P(a_i \mid b_j) \log_2 P(a_i \mid b_j)$$

можно интерпретировать как среднее количество информации, которое теряется в канале связи из-за помех при условии регистрации символа b_j. Тогда *среднее количество всей потерянной в канале информации* равно

$$H(a \mid b) = -\sum_{j=1}^{m} P(b_j) \sum_{i=1}^{n} P(a_i \mid b_j) \log_2 P(a_i \mid b_j),$$

где $P(b_j) = \sum_{i=1}^{n} P(a_i) P(b_j \mid a_i)$ — вероятность приема символа b_j, $j=1, 2, \ldots, m$. Таким образом, разность $H(a) - H(a \mid b)$ представляет собой *количество информации, переданное через канал связи*, а максимальное значение этой разности

$$C = \max (H(a) - H(a \mid b)),$$

где максимизация ведется по всем возможным вероятностным распределениям $P(a_1), P(a_2), \ldots, P(a_n)$, называется *пропускной способностью* канала связи.

Запишем подробнее задачу максимизации, возникающую при вычислении пропускной способности канала. Для этого введем следующие более краткие обозначения: $p_i = P(a_i)$, $p_{ij} = P(b_j \mid a_i)$, $i=1, 2, \ldots, n$; $j=1, 2, \ldots, m$. В соответствии с этими обозначениями для всех i и j имеют место равенства

$$P(a_i \mid b_j) = p_i p_{ij} \Big/ \sum_{i=1}^{n} p_i p_{ij}, \quad P(b_j) = \sum_{i=1}^{n} p_i p_{ij}.$$

Задача нахождения пропускной способности принимает вид

$$f(p_1, p_2, \ldots, p_n) = -\sum_{i=1}^{n} p_i \log_2 p_i +$$

$$+ \sum_{j=1}^{m} \sum_{i=1}^{n} p_i p_{ij} \log_2 \left(p_i p_{ij} \Big/ \sum_{i=1}^{n} p_i p_{ij} \right) \to \max; \tag{2.54}$$

$$p_1 + p_2 + \ldots + p_n = 1, \quad p_1, p_2, \ldots, p_n \geqslant 0.$$

Ограничения в этой задаче линейные, причем $0 \leqslant p_i \leqslant 1$, $i = 1, 2, \ldots, n$. Таким образом, допустимое множество не пусто, выпукло, замкнуто и ограничено. Вместе с непрерывностью целевой функции это влечет разрешимость задачи максимизации (2.54) при произвольных фиксированных значениях p_{ij}.

Установим вогнутость целевой функции на допустимом множестве. Для этого убедимся в справедливости неравенства

$$\lambda f(p_1^{(1)}, \ldots, p_n^{(1)}) + (1-\lambda) f(p_1^{(2)}, \ldots, p_n^{(2)}) - f(\lambda_1^{(1)} +$$

$$+ (1-\lambda) p_1^{(2)}, \ldots, \lambda p_n^{(1)} + (1-\lambda) p_n^{(2)}) \leqslant 0 \tag{2.55}$$

для всех $\lambda \in [0, 1]$ и любых допустимых в задаче (2.54) наборов

$p_1^{(1)},\ldots,p_n^{(1)}$ и $p_1^{(2)},\ldots,p_n^{(2)}$. Используя равенство $\sum_{j=1}^{m} p_{ij}=1$, можно записать

$$f(p_1, p_2,\ldots, p_n)= -\sum_{i=1}^{n}\sum_{j=1}^{m} p_i p_{ij} \log_2 p_i + \sum_{i=1}^{n}\sum_{j=1}^{m} p_i p_{ij} \log_2 \Big(p_i \times$$

$$\times p_{ij} \Big/ \sum_{i=1}^{n} p_i p_{ij}\Big) = \sum_{i=1}^{n}\sum_{j=1}^{m} p_i p_{ij} \log_2 \Big(p_{ij} \Big/ \sum_{i=1}^{n} p_i p_{ij}\Big).$$

В соответствии с этим левую часть неравенства (2.55), которую обозначим через Δ, легко преобразовать к виду

$$\Delta = \lambda \sum_{i=1}^{n}\sum_{j=1}^{m} p_i^{(1)} p_{ij} \log_2 \Big[\sum_{i=1}^{n} (\lambda p_i^{(1)} + (1-\lambda) p_i^{(2)}) p_{ij} \Big/ \sum_{i=1}^{n} p_i^{(1)} p_{ij}\Big] +$$

$$+ (1-\lambda) \sum_{i=1}^{n}\sum_{j=1}^{m} p_i^{(2)} p_{ij} \log_2 \Big[\sum_{i=1}^{n} (\lambda p_i^{(1)} + (1-\lambda) p_i^{(2)}) p_{ij} \Big/ \sum_{i=1}^{n} p_i^{(2)} p_{ij}\Big].$$

Отсюда, используя неравенство $\log_2 x \leqslant (x-1) \log_2 e$ $(x>0)$, получаем неравенство

$$\Delta \leqslant \lambda \sum_{j=1}^{m}\sum_{i=1}^{n} [(\lambda p_i^{(1)} + (1-\lambda) p_i^{(2)}) p_{ij} - p_i^{(1)} p_{ij}] \log_2 e +$$

$$+ (1-\lambda) \sum_{j=1}^{m}\sum_{i=1}^{n} [(\lambda p_i^{(1)} + (1-\lambda) p_i^{(2)}) p_{ij} - p_i^{(2)} p_{ij}] \log_2 e.$$

С помощью несложных преобразований в правой части этого неравенства можно показать, что она равна нулю. Следовательно, неравенство (2.55) доказано, а значит, рассматриваемая задача (2.54) является задачей вогнутого программирования.

Глава 3

ЛИНЕЙНОЕ ПРОГРАММИРОВАНИЕ

Задачи линейного программирования составляют наиболее простой и хорошо изученный класс задач математического программирования.

Распространенным методом решения задачи линейного программирования является симплекс-метод, который излагается в § 3.1. На практике обычно используют видоизмененный вариант этого метода — модифицированный симплекс-метод, требующий меньшего объема вычислений. Рассмотрению модифицированного симплекс-метода посвящен § 3.2. Еще один вариант симплекс-метода — двойственный симплекс-метод — разобран в § 3.3.

§ 3.1. СИМПЛЕКС-МЕТОД

1. Геометрическая иллюстрация задач линейного программирования. *Общая задача линейного программирования* (ЛП) состоит в минимизации (максимизации) линейной целевой функции

$$\langle c,\ x\rangle = c_1x_1 + c_2x_2 + \ldots + c_nx_n, \tag{3.1}$$

когда допустимое множество задано конечным числом линейных ограничений вида

$$\begin{aligned} &a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n \geqslant b_i,\ i = 1,\ 2, \ldots, k_1, \\ &a_{j1}x_1 + a_{j2}x_2 + \ldots + a_{jn}x_n \leqslant b_j,\ j = k_1 + 1, \ldots, k_2, \\ &a_{l1}x_1 + a_{l2}x_2 + \ldots + a_{ln}x_n = b_l,\ l = k_2 + 1, \ldots, k_3. \end{aligned} \tag{3.2}$$

В частности, ограничениями могут быть только равенства (в этом случае $k_1 = k_2 = 0$) или только неравенства ($k_2 = k_3$). Часто среди ограничений встречаются условия неотрицательности переменных: $x_1 \geqslant 0,\ x_2 \geqslant 0, \ldots,\ x_n \geqslant 0$. Ограничения в задачах линейного программирования обязательно должны присутствовать, так как в противном случае функция (3.1) заведомо не будет достигать ни минимума, ни максимума.

Так, в задаче

$$\langle c,\ x\rangle = -x_1 + x_2 \to \min;$$

$$x_1 + 2x_2 \leqslant 2,$$

$$x_1 - x_2 \leqslant 1,$$

$$x_1,\ x_2 \geqslant 0$$

имеются две переменные x_1, x_2, два ограничения типа неравенств и условие неотрицательности переменных. Ограничения этой задачи определяют плоский четырехугольник (рис. 3.1). Линии уровня минимизируемой функции, т. е. множество точек, удовлетворяющих равенству

$$\langle c,\ x\rangle = -x_1 + x_2 = C$$

$$(C = \text{const}),$$

образуют семейство прямых, перпендикулярных вектор-градиенту $c = (-1,\ 1)$ (на рис. 3.1 изображены две линии уровня, соответствующие $C = 0$ и $C = -1$). Нас интересует минимальное значение целевой функции на четырехугольнике. Такое значение дает линия уровня, имеющая общие точки (одну точку) с данным четырехугольником и являющаяся крайней в направ-

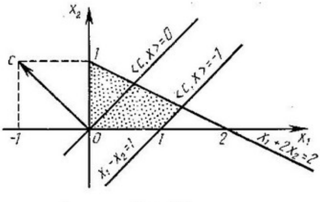

Рис. 3.1

лении, противоположном градиенту c. Это прямая $x_1 - x_2 = 1$. Следовательно, минимальное значение целевой функции равно -1 и достигается оно в тех точках прямой $x_1 - x_2 = 1$, которые принадлежат четырехугольнику. Заметим, что среди точек минимума имеются две вершины четырехугольника.

В рассмотренном примере линейные ограничения задают ограниченный плоский четырехугольник. Вообще же при $n=2$ допустимое множество в задаче линейного программирования может оказаться пустым (*а*) или неограниченным (*б*, *в*), может не иметь ни одной вершины (*в*) (рис. 3.2). Из геометрических соображений

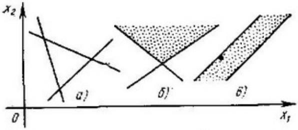

Рис. 3.2

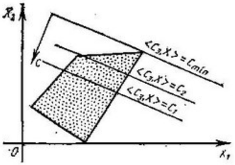

Рис. 3.3

очевидно, что в случае непустого плоского ограниченного многоугольника среди точек минимума линейной функции обязательно должна быть хотя бы одна вершина этого многоугольника (рис. 3.3).

Геометрическая интерпретация задач линейного программирования большей размерности ($n>2$) аналогична. Ограничения (3.2) определяют допустимое множество, являющееся пересечением конечного числа полупространств и гиперплоскостей, которое называют *многогранным множеством*. Оно выпукло и замкнуто (см. § 1.2). В частности, оно может оказаться пустым или неограниченным. *Многогранное множество имеет не более чем конечное число вершин.* Точка многогранного множества называется вершиной, если она не лежит внутри ни одного из отрезков, целиком принадлежащих данному множеству. Поверхностями уровня $\{x \in \mathbf{R}^n | \langle c, x\rangle = C\}$ в n-мерном случае являются гиперплоскости, ортогональные градиенту c. Справедливо [3] следующее утверждение: *если допустимое множество имеет вершины, то среди точек минимума (максимума) линейной функции всегда будет по крайней мере одна из вершин.* Здесь предполагается, что точки минимума (максимума) существуют.

На основе этого утверждения можно предположить следующий способ решения задач линейного программирования, множество ограничений которых имеет вершины: сначала определить все вершины допустимого многогранного множества (это осуществимо, так как количество вершин конечно), а затем среди найденных вершин выбрать ту, в которой целевая функция принимает наименьшее (наибольшее) значение. Таким образом, задача линейного программирования будет решена за конечный промежуток времени. Предложенный метод не вызывает принципиальных возражений, однако всерьез говорить о нем не приходится, так как количество вершин в реальных задачах оказывается слишком большим и перебор их даже с помощью самых быстродействующих ЭВМ может потребовать непомерно большого времени.

В излагаемом ниже симплекс-методе для нахождения оптимальной вершины также производится перебор вершин многогранного множества, но этот перебор осуществляется целенаправленно, что позволяет исключить из рассмотрения значительное количество вершин, заведомо не являющихся оптимальными.

2. Каноническая задача линейного программирования. Поскольку задача максимизации функции $\langle c, x\rangle$ эквивалентна задаче минимизации $-\langle c, x\rangle$, далее будем рассматривать задачи минимизации.

Простейшая версия симплекс-метода приспособлена для решения *канонической задачи* линейного программирования

$$\begin{aligned} &\langle c, x\rangle \to \min; \\ &Ax=b, \\ &x \geqq 0_n. \end{aligned} \tag{3.3}$$

Здесь $c=(c_1, c_2, ..., c_n)^{\mathrm{T}}$ — *вектор коэффициентов целевой функции*; $A=(a_{ij})$ — прямоугольная матрица размера $k\times n$, называемая *матрицей коэффициентов ограничений*; $b=(b_1, b_2, ..., b_k)^{\mathrm{T}}$ — *вектор ограничений* и $x=(x_1, x_2, ..., x_n)^{\mathrm{T}}$ — *вектор переменных*. В развернутом виде каноническая задача записывается следующим образом:

$$\begin{aligned} &c_1x_1+c_2x_2+\ldots+c_nx_n \to \min; \\ &a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n=b_1, \\ &a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n=b_2, \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ &a_{k1}x_1+a_{k2}x_2+\ldots+a_{kn}x_n=b_k, \\ &x_1\geqslant 0,\ x_2\geqslant 0, \ldots, x_n\geqslant 0. \end{aligned} \tag{3.3'}$$

Любую задачу линейного программирования можно привести к каноническому виду. Например, ограничение типа неравенства

$$a_{i1}x_1+a_{i2}x_2+\ldots+a_{in}x_n\leqslant b_i$$

можно представить в виде равенства, если ввести дополнительную неотрицательную переменную x_{n+1}:

$$a_{i1}x_1+a_{i2}x_2+\ldots+a_{in}x_n+x_{n+1}=b_i. \tag{3.4}$$

Если левая часть исходного ограничения-неравенства больше либо равна правой, то в соответствующем ограничении-равенстве (3.4) знак перед дополнительной переменной следует изменить на противоположный.

Если на переменную x_j не наложено условие неотрицательности, то ее можно заменить двумя неотрицательными переменными x_j' и x_j'', полагая $x_j = x_j' - x_j''$, x_j', $x_j'' \geqslant 0$. Такая замена правомерна, так как любое число можно представить в виде разности двух неотрицательных чисел.

Каноническую задачу можно еще сформулировать так: среди неотрицательных решений системы линейных уравнений

$$\begin{aligned}
&a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1,\\
&a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2,\\
&\cdots\cdots\cdots\cdots\cdots\cdots\\
&a_{k1}x_1 + a_{k2}x_2 + \ldots + a_{kn}x_n = b_k
\end{aligned} \tag{3.5}$$

найти такие, которые доставляют минимальное значение целевой функции $\langle c, x\rangle$. Если система (3.5) вообще не имеет решений, т. е. является несовместной, или же не имеет неотрицательных решений, то и каноническая задача (3.3) также не имеет решения. В том случае, когда система (3.5) имеет единственное неотрицательное решение, это решение, очевидно, является и оптимальным. Поэтому собственно о задаче оптимизации (3.3) имеет смысл говорить лишь тогда, когда система (3.5) имеет более чем одно неотрицательное решение.

Предполагая, что из системы (3.5) исключены «лишние» уравнения, которые выражаются линейной комбинацией остальных, можно считать, что ранг матрицы системы (3.5) равен числу уравнений, т. е. $\operatorname{rang} A = k$.

3. Идея симплекс-метода. Распространенным методом решения систем линейных уравнений является *метод последовательного исключения неизвестных (метод Гаусса)*. Напомним, что в его основе лежат следующие *элементарные преобразования:*

а) умножение какого-либо уравнения на число, отличное от нуля;

б) сложение двух уравнений и последующая замена одного из этих уравнений получившейся суммой;

в) перемена местами двух уравнений.

Элементарные преобразования не изменяют множества решений данной системы.

Поясним идею симплекс-метода на следующем простом примере: минимизировать функцию $y = x_1 + x_2 + x_3$ при следующих условиях:

$$\begin{aligned}
x_1 - x_2 + 2x_3 &= 1,\\
-x_1 + 2x_2 - x_3 &= 2,\\
x_1 \geqslant 0,\ x_2 \geqslant 0,\ &x_3 \geqslant 0.
\end{aligned}$$

Поставим в соответствие этой задаче следующую систему линейных уравнений:

$$
\begin{aligned}
y - x_1 - x_2 - x_3 &= 0,\\
x_1 - x_2 + 2x_3 &= 1,\\
-x_1 + 2x_2 - x_3 &= 2.
\end{aligned}
\tag{3.6}
$$

С помощью элементарных преобразований можно прийти к системе с неизвестными y, x_1, x_2, x_3:

$$y \qquad + 3x_3 = 7, \tag{3.6'}$$

$$x_1 \quad + 3x_3 = 4, \tag{3.6''}$$

$$x_2 + x_3 = 3. \tag{3.6'''}$$

Уравнение (3.6‴) — это сумма второго и третьего уравнений из системы (3.6); уравнение (3.6″) является суммой второго из уравнений (3.6) и уравнения (3.6‴); уравнение (3.6′) — сумма первого из уравнений (3.6) и уравнений (3.6″), (3.6‴).

Значения переменных $y=7$, $x_1=4$, $x_2=3$, $x_3=0$ являются решением системы (3.6′)—(3.6‴) (а значит, и системы (3.6), причем $x_1, x_2, x_3 \geqslant 0$, т. е. найденное решение допустимо. Величина y — это значение исходной целевой функции. Так как $y=7-3x_3$ (см. (3.6′)), то при увеличении x_3 величина y уменьшается. Необходимо получить наименьшее возможное значение y, поэтому увеличим насколько возможно x_3. Величина x_3 должна удовлетворять условиям (3.6″) и (3.6‴); следовательно, увеличивать неограниченно ее нельзя, иначе величины x_1 и x_2 станут отрицательными. Здесь $x_1=4-3x_3$ и $x_2=3-x_3$. Итак, наибольшее значение x_3, при котором величины x_1 и x_2 не отрицательны, таково: $x_3=\min\{^4/_3,\ ^3/_1\}={}^4/_3$. Пусть $x_3={}^4/_3$. Тогда $x_1=0$ (см. равенство (3.6″)). С помощью элементарных преобразований систему (3.6′)—(3.6‴) можно привести к виду

$$
\begin{aligned}
y \qquad - x_1 &= 3\\
x_3 \quad + 1/3x_1 &= 4/3,\\
x_2 - 1/3x_1 &= 5/3.
\end{aligned}
\tag{3.7}
$$

Теперь запишем решение системы (3.7): $y=3$, $x_1=0$, $x_2={}^5/_3$, $x_3={}^4/_3$. Здесь $y=3+x_1$, поэтому с увеличением x_1 значение y увеличивается. Следовательно, значение целевой функции $y=3$ уменьшить нельзя, а значит, оно является минимальным. Ему соответствует оптимальное решение $x_1=0$, $x_2={}^5/_3$, $x_3={}^4/_3$.

4. Алгоритм симплекс-метода. Пусть требуется минимизировать функцию

$$y = \langle c, x\rangle$$

при условиях

$$Ax=b,$$

$$x \geq 0_n,$$

где система уравнений $Ax=b$ $(\operatorname{rang} A = k < n)$ в результате элементарных преобразований уже приведена к виду

$$\begin{array}{llll} x_1 & & & +a_{1,k+1}x_{k+1}+\ldots+a_{1n}x_n=b_1, \\ & x_2 & & +a_{2,k+1}x_{k+1}+\ldots+a_{2n}x_n=b_2, \\ \ldots & & & \ldots\ldots\ldots\ldots\ldots\ldots \\ & & x_k & +a_{k,k+1}x_{k+1}+\ldots+a_{kn}x_n=b_k, \end{array} \qquad (3.8)$$

причем $b_i \geqslant 0$, $i=1, 2, \ldots, k$. Случай, когда такое преобразование выполнить достаточно трудно, рассмотрен в следующем пункте. Прибавляя к уравнению $y-\langle c, x\rangle=0$ все k уравнений (3.8), предварительно умноженные соответственно на $c_1, c_2, \ldots, c_k$, получаем уравнение для y, в котором все переменные $x_1, x_2, \ldots, x_k$ отсутствуют (однако может появиться свободный член). Полагая $y=x_0$, $b_i=a_{i0}$, $i=1, 2, \ldots, k$, уравнение, содержащее x_0, вместе с системой (3.8) окончательно можно записать в следующем виде:

$$\begin{array}{lllll} x_0 & & & & +a_{0,k+1}x_{k+1}+\ldots+a_{0n}x_n=a_{00}, \\ & x_1 & & & +a_{1,k+1}x_{k+1}+\ldots+a_{1n}x_n=a_{10}, \\ & & x_2 & & +a_{2,k+1}x_{k+1}+\ldots+a_{2n}x_n=a_{20}, \\ & & & & \ldots\ldots\ldots\ldots\ldots\ldots \\ & & & x_k & -a_{k,k+1}x_{k+1}+\ldots+a_{ku}x_n=a_{k0}. \end{array} \qquad (3.9)$$

Здесь $a_{0j}=\sum\limits_{m=1}^{k} c_m a_{mj}-c_j$, $j=k+1, \ldots, n$; $a_{00}=\sum\limits_{m=1}^{k} c_m b_m$.

Систему (3.9) называют *диагональной формой* относительно переменных $x_0, x_1, \ldots, x_k$ (или говорят, что система приведена к *диагональному виду* относительно указанных переменных). Решение этой системы (без учета уравнения, содержащего x_0) вида

$$x_i=a_{i0},\ i=1, 2, \ldots, k;\ x_{k+1}=\ldots=x_n=0 \qquad (3.10)$$

называется *базисным*, а $x_1, x_2, \ldots, x_k$ — *базисными переменными*. Ранее мы предположили, что компоненты вектора ограничений b не отрицательны, поэтому базисное решение является допустимым. Если среди элементов $a_{10}, a_{20}, \ldots, a_{k0}$ имеются нулевые, то соответствующие компоненты базисного решения также равны нулю. Базисное решение, хотя бы одна базисная компонента которого равна нулю, называют *вырожденным базисным решением*.

Как отмечено выше, базисное решение — это определенное решение системы уравнений, приведенной к диагональному виду. В теории линейного программирования понятие базисного решения используют и применительно к системе уравнений $Ax=b$ с произвольной матрицей A, имеющей ранг k. Заметим, что матрицу A можно рассматривать как совокупность составляющих ее столбцов. Количество столбцов совпадает с количеством переменных n. Таким образом, между номерами столбцов матрицы A и номерами переменных имеется взаимно однозначное соответствие: i-й переменной соответствует i-й столбец и обратно. Приведем определение базисного решения: ненулевое решение $x^{(0)}=(x_1^{(0)}, x_2^{(0)},\dots, x_n^{(0)})^{\text{т}}$ системы $Ax=b$ называется *базисным,* если вектор-столбцы, соответствующие ненулевым компонентам вектора $x^{(0)}$, линейно независимы. Нулевой вектор $x^{(0)}=0_n$, если он является решением системы, также причисляют к числу базисных решений. Если компоненты базисного решения не отрицательны, то его называют *допустимым базисным решением.* Так, решение (3.10) системы (3.9) (без первого уравнения), согласно последнему определению, является базисным, поскольку ненулевыми могут быть только компоненты $x_1, x_2, \dots, x_k$, а им соответствуют линейно независимые столбцы

$$\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix}, \begin{pmatrix}0\\1\\\vdots\\0\end{pmatrix},\dots, \begin{pmatrix}0\\0\\\vdots\\1\end{pmatrix}.$$

Компоненты допустимого базисного решения — это координаты вершины многогранного множества, заданного условиями $Ax=b$, $x\geqq 0_n$ (доказательство этого факта, устанавливающего геометрический смысл допустимых базисных решений, приведено, например, в [3]).

Систему (3.9) обычно записывают в виде табл. 3.1.

Таблица 3.1

		x_1	x_2	$\dots$	x_k	x_{k+1}	$\dots$	x_n
x_0	a_{00}	0	0	$\dots$	0	$a_{0,\,k+1}$	$\dots$	a_{0n}
x_1	a_{10}	1	0	$\dots$	0	$a_{1,\,k+1}$	$\dots$	a_{1n}
x_2	a_{20}	0	1	$\dots$	0	$a_{2,\,k+1}$	$\dots$	a_{2n}
$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$
x_k	a_{k0}	0	0	$\dots$	1	$a_{k,\,k+1}$	$\dots$	a_{kn}

Каждая строка табл. 3.1 задает соответствующее уравнение из (3.9), свободный член которого записан в самом левом (нулевом) столбце. Слева от таблицы записаны текущие базисные переменные. Столбцы, соответствующие базисным переменным, составляют единичную подматрицу таблицы. Табл. 3.1 называют *симплекс-таблицей.*

Исходную задачу линейного программирования можно привести к диагональной относительно x_0 и каких-то других k переменных из набора $x_1, x_2, \dots, x_n$ форме. В этом случае изменятся и базисные переменные. А столбцы, отвечающие этим базисным переменным, располагаясь в соответствующих графах симплекс-таблицы, составляют подматрицу, отличающуюся от единичной лишь перестановкой столбцов.

В любом случае *текущее базисное решение* — это такое решение, в котором все небазисные переменные имеют нулевое значение, а базисные переменные соответственно равны элементам нулевого столбца. Например, системе (3.7) соответствует симплекс-таблица 3.2, в которой текущее базисное решение таково: $x_1=0$, $x_2={}^5/_3$, $x_3={}^4/_3$.

Таблица 3.2

		x_1	x_2	x_3
	3	—1	0	0
x_3	4/3	1/3	0	1
x_2	5/3	—1/3	1	0

Рассмотрим *алгоритм симплекс-метода.* Начинать следует с таблицы, которая соответствует диагональной форме относительно x_0 и некоторых k (базисных) переменных из полного набора $x_1, x_2, \dots, x_n$. Все элементы нулевого столбца (кроме, возможно, a_{00}) должны быть неотрицательными.

Шаг 1. Если среди элементов нулевой строки (не считая a_{00}) нет положительных, то вычисления следует закончить, так как текущее базисное решение оптимально и минимальное значение целевой функции равно a_{00}. Таблица, соответствующая этому положению, называется *оптимальной таблицей.* Если таблица не является оптимальной, то следует среди положительных элементов нулевой строки (исключая a_{00}) выбрать некоторый элемент. Обычно среди положительных элементов выбирают максимальный, однако это не обязательно. Пусть выбран элемент $a_{0s}>0$. Тогда столбец с номером s называют *ведущим столбцом.*

Если все элементы нулевой строки таблицы не положительны, то значение x_0 целевой функции, найденное из нулевого уравнения, является разностью между a_{00} и суммой небазисных переменных с соответствующими коэффициентами: $x_0=a_{00}-\Sigma a_{0j}x_j$ (так как коэффициенты при базисных переменных равны нулю). Таким образом, величина x_0 зависит только от небазисных переменных. Коэффициенты a_{0j} не положительны, а небазисные переменные не отрицательны, поэтому минимальное значение x_0 достигается только в том случае, когда значения всех небазисных переменных равны нулю. Следовательно, число $x_0=a_{00}$ действительно минимум, а текущее базисное решение является оптимальным.

Отметим, что выбор максимального элемента нулевой строки на шаге 1 обеспечивает максимальное убывание целевой функции, т. е. в конечном счете ведет к сокращению общего числа вычислений.

Шаг 2. Выделить среди положительных элементов a_{is} ведущего столбца элемент, для которого отношение a_{i0}/a_{is} является наименьшим. Пусть этот элемент a_{rs}. Его называют *ведущим элементом*, а строку с номером r — *ведущей строкой*.

При выполнении вычислений шага 2 предполагается, что в ведущем столбце s найдется по крайней мере один положительный элемент. Если это предположение не выполняется, т. е. все элементы ведущего столбца (кроме a_{0s}) не положительны, то исходная задача не имеет конечного решения (целевая функция на допустимом множестве не ограничена снизу). Чтобы убедиться в этом, предположим, что симплекс-таблица имеет вид табл. 3.1 и $s=n$. Решение

$$x_i^{(\lambda)}=a_{i0}-a_{in}\lambda,\ i=1,\ 2,\dots,k,$$

$$x_i^{(\lambda)}=0,\ i=k+1,\dots,n-1,$$

$$x_n^{(\lambda)}=\lambda,$$

так как $a_{i0}\geqslant 0$, $a_{in}\leqslant 0$, $i=1, 2, \dots, k$, при любых $\lambda\geqslant 0$ является допустимым, т. е. удовлетворяет соответствующей системе уравнений и имеет неотрицательные компоненты. Из нулевой строки табл. 3.1 имеем

$$x_0=a_{00}-\sum_{j=1}^{n} a_{0j}x_j^{(\lambda)}=a_{00}-a_{0n}\cdot\lambda.$$

Поскольку $a_{0n}>0$, отсюда следует, что, увеличивая λ, можно неограниченно уменьшать значение x_0 целевой функции.

При выполнении вычислений шага 2 может получиться так, что минимум отношения a_{i0}/a_{is} $(a_{is}>0)$ окажется одинаковым для нескольких i, т. е. сразу несколько строк таблицы могут быть ведущими. В этом случае обычно в качестве ведущей выбирают строку с наименьшим номером. Такое правило выбора ведущей строки из нескольких равноправных строк на практике обычно не вызывает недоразумений, однако теоретически может приводить к *зацикливанию*, т. е. бесконечному повторению некоторого неоптимального базисного решения. Задачи линейного программирования, в которых возможно зацикливание, называют *вырожденными*. В п. 6 будет указано, как можно предотвратить зацикливание.

Шаг 3. Используя элементарные преобразования, сделать так, чтобы ведущий элемент a_{rs} стал равным единице, а все остальные элементы ведущего столбца — нулевыми. Иными словами, преобразовать данную систему уравнений к диагональному относительно нового набора базисных переменных виду. Новый набор отличается от прежнего тем, что в нем вместо переменной x_r участвует переменная x_s. Поэтому базисную переменную x_r слева от таблицы следует заменить на x_s. После этого вернуться к шагу 1.

Последовательное выполнение вычислений шагов 1—3 составляет одну *итерацию симплекс-метода.*

Итерации выполняют до тех пор, пока среди элементов нулевой строки таблицы (не считая a_{00}) не окажется ни одного отрицательного. Тогда текущее базисное решение является оптимальным, а минимальное значение целевой функции равно левому верхнему элементу таблицы.

Если целевая функция не является ограниченной снизу на допустимом множестве, то на одном из шагов 2 будет получен ведущий столбец, в котором нет ни одного положительного элемента (не считая элемента нулевой строки).

Замечание. 1. Алгоритм симплекс-метода обладает тем важным свойством, что после выполнения некоторого конечного числа итераций либо будет найдено оптимальное решение, либо установлена неразрешимость данной задачи.

Действительно, в канонической задаче количество различных базисных решений (вершин допустимого многогранного множества) конечно — оно не превышает C_n^k. Поэтому, если исходная задача разрешима, то на основании утверждения из п. 1 оптимальной является по крайней мере одна вершина многогранника. В невырожденной задаче в результате перебора конечного числа вершин будет найдена оптимальная. Если задача вырождена (т. е. возможно зацикливание), то, используя правило для предотвращения зацикливания, приведенное в п. 6, в конце концов также придем в оптимальную вершину. В том случае, когда исходная задача не имеет оптимального решения, на одном из шагов 2 будет получен ведущий столбец, в котором нет ни одного положительного элемента, и тем самым будет установлена неограниченность целевой функции.

Замечание 2. На каждой итерации симплекс-метода сохраняется допустимость базисного решения, т. е. элементы нулевого столбца (не считая верхнего) все время остаются неотрицательными. Это свойство следует из правила выбора ведущей строки на шаге 2.

Кроме того, алгоритм симплекс-метода таков, что при переходе от одного допустимого базисного решения к другому значение целевой функции, равное левому верхнему угловому элементу таблицы, уменьшается. В самом деле, если в табл. 3.1 ведущим выбран элемент a_{rs}, то после выполнения шага 3 в левой верхней клетке таблицы будет находиться разность $a_{00}-(a_{r0}/a_{rs})\cdot a_{0s}$, которая в силу положительности коэффициентов a_{r0}, a_{rs}, a_{0s} меньше, чем a_{00}. Величина этой разности и является новым значением целевой функции.

Указанное положение может нарушаться лишь в случае вырожденного базисного решения, причем когда имеет место зацикливание. В этом случае значение целевой функции может не изменяться, пока продолжается зацикливание.

Проиллюстрируем вычисления алгоритма симплекс-метода на следующем примере:

$$x_0=x_1+x_2+x_3-x_4+x_5-x_6\to\min;$$

$$x_1 \quad +2x_4-x_5+x_6=1,$$

$$x_2 \quad +x_5+x_6=2,$$

$$x_3-x_4+x_5-x_6=1,$$

$$x_1,\ x_2,\dots,x_6\geqslant 0.$$

Здесь в качестве исходного допустимого базисного решения следует взять $x_1=1$, $x_2=2$, $x_3=1$, $x_4=x_5=x_6=0$. Начальная таблица после сложения нулевого уравнения со всеми остальными уравнениями принимает вид табл. 3.3.

Таблица 3.3

		x_1	x_2	x_3	x_4	x_5	x_6
x_0	4	0	0	0	2	0	2
x_1	1	1	0	0	2	—1	1*
x_2	2	0	1	0	0	1	1
x_3	1	0	0	1	—1	1	—1

В нулевой строке имеются положительные элементы, поэтому текущее решение не является оптимальным. Среди двух равных положительных элементов нулевой строки выбираем, например, $a_{06}=2$. Таким образом, ведущим является последний столбец. В этом столбце имеются два положительных элемента: $a_{16}=a_{26}=1$. Сравнивая отношения $a_{01}/a_{16}=1/1$ и $a_{20}/a_{26}=2/1$, приходим к выводу, что первая строка является ведущей, а ведущим элементом будет $a_{16}=1$ (в таблице он помечен звездочкой). В соответствии с этим переменную x_1 из базиса следует вывести, а вместо нее ввести переменную x_6. Для этого первую строку табл. 3.3 последовательно умножаем на —2, —1, 1 и получившиеся строки складываем соответственно с нулевой, второй и третьей строками. В результате все элементы ведущего столбца (кроме ведущего элемента, который равен 1) станут нулевыми и симплекс-таблица на второй итерации принимает вид табл. 3.4.

Таблица 3.4

		x_1	x_2	x_3	x_4	x_5	x_6
x_0	2	—2	0	0	—2	2	0
x_6	1	1	0	0	2	—1	1
x_2	1	—1	1	0	—2	2*	0
x_3	2	1	0	1	1	0	0

Здесь в нулевой строке имеется единственный положительный элемент $a_{05}=2$, т. е. ведущий столбец определяется однозначно.

Так же однозначно определяется и ведущий элемент $a_{25}=2>0$. Теперь переменную x_2 следует вывести из базиса и ввести переменную x_5. Для этого сначала все элементы ведущей строки разделим на 2 (для того чтобы ведущий элемент стал равным 1), а затем с помощью элементарных преобразований добиваемся того, чтобы все остальные элементы ведущего столбца стали нулевыми. Приходим к оптимальной табл. 3.5, в которой среди элементов нулевой строки (не считая $a_{00}=1$) нет положительных.

Таблица 3.5

		x_1	x_2	x_3	x_4	x_5	x_6
x_0	1	-1	-1	0	0	0	0
x_6	1,5	0,5	0,5	0	1	0	1
x_5	0,5	$-0,5$	0,5	0	-1	1	0
x_3	2	1	0	1	1	0	0

Следовательно, текущее базисное решение $x_1=x_2=x_4=0$, $x_3=2$, $x_5=0,5$, $x_6=1,5$ оптимально, а минимальное значение целевой функции равно 1.

5. Двухфазный симплекс-метод. Выше алгоритм симплекс-метода был сформулирован применительно к канонической задаче линейного программирования специального вида. Мы предполагали, что исходная задача уже приведена к диагональной относительно x_0 и некоторых k базисных переменных форме. При этом считалось, что свободные члены во всех уравнениях (исключая нулевое) не отрицательны. Установим, как поступать в тех случаях, когда эти предположения не выполнены. Для приведения произвольной канонической задачи к диагональному виду с неотрицательными свободными членами можно использовать изложенный выше алгоритм симплекс-метода, если предварительно ввести дополнительные переменные.

Приведение исходной канонической задачи к диагональному виду равносильно указанию некоторого допустимого базисного решения системы ограничений. В самом деле, если диагональная форма имеет, например, вид (3.9), где $a_{i0}\geqslant 0$, $i=1, 2, \ldots, k$, то из этого представления сразу находим допустимое базисное решение (3.10). Обратно: пусть задано некоторое допустимое базисное решение $x^{(0)}=\left(x_1^{(0)}, x_2^{(0)}, \ldots, x_n^{(0)}\right)^{\mathrm{T}}\geqq 0_n$ системы ограничений $Ax=b$. Не уменьшая общности, можно считать, что положительными являются первые l, $l\leqslant k$, компонент вектора $x^{(0)}$. Согласно общему определению базисного решения, первые l вектор-столбцов матрицы A линейно независимы. Если $l<k$, то в силу того, что $\operatorname{rang} A=k$ должны найтись $k-l$ вектор-столбцов матрицы A, которые вместе с первыми l векторами образуют линейно независимую систему. Поэтому будем считать, что первые k столбцов матрицы A линейно независимы. Обозначим матрицу размера $k\times k$, которую они составляют, через $A_{\text{б}}$, а остальную часть матрицы A — через $A_{\text{н}}$. Кроме того, пусть $x_{\text{б}}=(x_1, x_2, \ldots x_k)^{\mathrm{T}}$ и $x_{\text{н}}=(x_{k+1}, \ldots, x_n)^{\mathrm{T}}$. Тогда систему уравнений $Ax=b$ можно записать так:

$$[A_{\text{б}}A_{\text{н}}]\begin{bmatrix} x_{\text{б}} \\ x_{\text{н}} \end{bmatrix}=A_{\text{б}}x_{\text{б}}+A_{\text{н}}x_{\text{н}}=b.$$

Для матрицы $A_б$ существует обратная матрица $A_б^{-1}$, поэтому после умножения последнего равенства слева на эту матрицу получаем

$$x_б + A_б^{-1}A_н x_н = A_б^{-1}b.$$

Эта система уравнений равносильна исходной системе $Ax=b$, и поэтому $x^{(0)}$ — ее решение. Из этого условия находим $x_б^{(0)} = A_б^{-1}b$, где $x_б^{(0)} = \left(x_1^{(0)}, x_2^{(0)}, \ldots, x_k^{(0)}\right)^т$, используя $x_i^{(0)}=0$, $i=k+1, \ldots, n$, и в результате получаем систему

$$x_б + A_б^{-1}A_н x_н = x_б^{(0)},$$

которая имеет диагональный вид (относительно $x_1, x_2, \ldots, x_k$) и неотрицательные свободные члены.

Итак, пусть требуется решить общую каноническую задачу линейного программирования:

$$\begin{gathered} y=c_1x_1+c_2x_2+\ldots+c_nx_n \to \min; \\ \begin{aligned} &a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n=b_1, \\ &a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n=b_2, \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ &a_{k1}x_1+a_{k2}x_2+\ldots+a_{kn}x_n=b_k, \\ &x_1, x_2, \ldots, x_n \geqslant 0. \end{aligned} \end{gathered} \tag{3.11}$$

Можно считать, что здесь все свободные члены $b_1, b_2, \ldots, b_k$ не отрицательны (так как уравнение, содержащее отрицательный свободный член, всегда можно умножить на -1). Выполнение равенства $\operatorname{rang} A=k$ не обязательно.

Введем дополнительные неотрицательные переменные $x_{n+1}, x_{n+2}, \ldots, x_{n+k}$, которые называют *искусственными переменными*, и рассмотрим новую задачу:

$$z=x_{n+1}+x_{n+2}+\ldots+x_{n+k} \to \min;$$

$$y-c_1x_1-c_2x_2-\ldots-c_nx_n=0,$$

$$\begin{array}{llll} a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n+ & x_{n+1} & & =b_1, \\ a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n & \quad + \; x_{n+2} & & =b_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots & \ldots\ldots & & \ldots \\ a_{k1}x_1+a_{k2}x_2+\ldots+a_{kn}x_n & & + \; x_{n+k} & =b_k, \end{array}$$

$$x_1, x_2, \ldots x_n, \ldots, x_{n+k} \geqslant 0.$$

Складывая уравнение $z-x_{n+1}-x_{n+2}-...-x_{n+k}=0$ со всеми уравнениями, содержащими переменные $x_{n+1}, ..., x_{n+k}$, придем к системе

$$\begin{array}{l|llll|l}
d_1x_1+d_2x_2+\ldots+d_nx_n+ & z & & & & =z_0, \\
-c_1x_1-c_2x_2-\ldots-c_nx_n & & +y & & & =0, \\
a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n & & +x_{n+1} & & & =b_1, \\
a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n & & & +x_{n+2} & & =b_2, \\
\cdots\cdots\cdots\cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
a_{k1}x_1+a_{k2}x_2+\ldots+a_{kn}x_n & & & & +x_{n+k} & =b_k,
\end{array} \quad (3.12)$$

$$x_1, x_2, \ldots, x_n, x_{n+k} \geqslant 0,$$

где $$d_m=\sum_{i=1}^{k} a_{im},\ m=1, 2, \ldots, n; \quad z_0=\sum_{m=1}^{k} b_m.$$

Система (3.12) имеет диагональный вид относительно z, y, $x_{n+1}, ..., x_{n+k}$ (где переменная y может принимать и отрицательные значения).

К системе (3.12) уже можно применить симплекс-метод, изложенный в предыдущем пункте. Все искусственные переменные не отрицательны, поэтому минимальное значение функции $z=x_{n+1}+ +...+x_{n+k}$ есть $z_{\min}=0$. Следовательно, если исходная система ограничений $Ax=b$, $x\geqq 0_n$ совместна, то значение $z_{\min}=0$ будет достигнуто и при этом все искусственные переменные должны стать равными нулю. Вместо них в число базисных переменных войдут переменные из $x_1, x_2, ..., x_n$. Далее, строку с z можно вычеркнуть и перейти к минимизации y. Такой способ решения исходной канонической задачи называют *двухфазным симплекс-методом*. На первой фазе, минимизируя z, из числа базисных переменных выводятся все искусственные переменные, а на второй фазе минимизируется целевая функция y и определяется оптимальное решение исходной задачи.

Рассмотрим более подробно случаи, которые могут встретиться на первой фазе:

1) Если в результате решения задачи (3.12) окажется, что $z_{\min}>0$, то это свидетельствует о том, что система ограничений $Ax=b$, $x\geqq 0_n$ несовместна.

2) Если $z_{\min}=0$ и слева от симплексной таблицы нет ни одной искусственной переменной, то можно приступать ко второй фазе.

3) Если $z_{\min}=0$, но слева от таблицы имеются искусственные переменные, то, используя элементарные преобразования, эти переменные следует вывести из числа базисных, а вместо них ввести исходные переменные. В этом случае базисное решение вырожденное (имеет нулевые базисные компоненты). Если искусственную переменную x_{n+r} невозможно вывести из базиса из-за того, что $a_{rj}=0$

для всех $j=1, 2, \ldots, n$, то это означает, что r-е уравнение системы (3.11) «лишнее» и его нужно вычеркнуть из таблицы.

Следующий пример иллюстрирует первую фазу симплекс-метода:

$$y=x_1-x_2+1\to\min$$
$$2x_1+x_2+3x_3=1,$$
$$x_1-3x_2+x_3=-3,$$
$$x_1+11x_2+3x_3=11,$$
$$x_1,\ x_2,\ x_3\geqslant 0.$$

Умножим второе из уравнений на —1 и составим задачу для первой фазы, введя искусственные переменные x_4, x_5, x_6:

$$\begin{aligned} z-x_4-x_5-x_6&=0,\\ y-x_1+x_2&=1,\\ 2x_1+x_2+3x_3+x_4&=1,\\ -x_1+3x_2-x_3+x_5&=3,\\ x_1+11x_2+3x_3+x_6&=11.\end{aligned}$$

Исключая из первого уравнения искусственные переменные, получаем систему

$$\begin{aligned} z+2x_1+15x_2+5x_3&=15,\\ y-x_1+x_2&=1,\\ 2x_1+x_2+3x_3+x_4&=1,\\ -x_1+3x_2-x_3+x_5&=3,\\ x_1+11x_2+3x_3+x_6&=11,\end{aligned}$$

которой соответствует табл. 3.6.

Таблица 3.6

		x_1	↓ x_2	x_3	x_4	x_5	x_6
z	15	2	15	5	0	0	0
y	1	−1	1	0	0	0	0
← x_4	1	2	1*	3	1	0	0
x_5	3	−1	3	−1	0	1	0
x_6	11	1	11	3	0	0	1

В качестве ведущего выбираем второй столбец. Сравнивая отношения 1/1, 3/3, 11/11, делаем вывод, что ведущей может быть любая из трех последних строк. Выбираем в качестве ведущей строку с наименьшим номером. Таким образом, базисную переменную x_4

следует заменить на x_2. В результате соответствующих преобразований приходим к табл. 3.7.

Таблица 3.7

		↓ x_1	x_2	x_3	x_5	x_6
z	0	−28	0	−40	0	0
y	0	−3	0	−3	0	0
x_2	1	2	1	3	0	0
← x_5	0	−7*	0	−10	1	0
x_6	0	−21	0	−30	0	1

Четвертый столбец, соответствующий выведенной базисной переменной x_4, далее не понадобится, поэтому из таблицы он исключен. Из табл. 3.7 следует, что минимальное значение z достигнуто и равно нулю. Однако искусственные переменные x_5 и x_6 еще не выведены из базиса. Выведем переменную x_5. Элементы предпоследней строки, соответствующие переменным x_1 и x_3, отрицательны, поэтому предварительно умножим эту строку на −1 (что допустимо, так как в нулевом столбце этой строки стоит элемент, равный нулю). Введем вместо x_5 переменную x_1. Для этого элементы ведущей строки разделим на 7 и с помощью элементарных преобразований добьемся появления нулевых элементов в ведущем столбце. В результате получаем табл. 3.8.

Таблица 3.8

		x_1	x_2	x_3	x_6
z	0	0	0	0	0
y	0	0	0	9/7	0
x_2	1	0	1	1/7	0
x_1	0	1	0	10/7	0
x_6	0	0	0	0	1

Переменную x_6 из числа базисных вывести нельзя, так как все соответствующие элементы равны нулю. Значит, третье из уравнений было «лишним». Исключив верхнюю и нижнюю строки, а также последний столбец, можно перейти ко второй фазе.

6. Правило предотвращения зацикливания. В п. 4 отмечалось, что если на одном из шагов 2 симплекс-метода ведущая строка определяется неоднозначно, то при решении может возникнуть зацикливание, т. е. бесконечное повторение одного и того же базисного решения. Сформулируем правило выбора ведущей строки, используя которое можно предотвратить зацикливание алгоритма симплекс-метода.

Предположим, что в качестве ведущего выбран столбец s и

$$\min_{i \in \{i \mid a_{is} > 0\}} a_{i0}/a_{is} = a_{r_1 0}/a_{r_1 s} = a_{r_2 0}/a_{r_2 s} = \ldots = a_{r_l 0}/a_{r_l s}.$$

Если $l=1$, то ведущая строка определена однозначно. Это строка r_1. Если $l>1$, то вычислим отношения a_{i1}/a_{is} для $i=r_1, r_2, ..., r_l$ и аналогично находим строки, которым соответствует минимум этого отношения. Если таковой является единственная строка, то ее следует считать ведущей. В противном случае составляем для найденных строк отношение a_{i2}/a_{is} и снова определяем строки, которым соответствует минимум отношения, и т. д. В результате ведущая строка будет определена единственным образом.

§ 3.2. МОДИФИЦИРОВАННЫЙ СИМПЛЕКС-МЕТОД

1. Алгоритм модифицированного симплекс-метода. Алгоритм симплекс-метода, описанного в п. 4 предыдущего параграфа, таков, что на каждой его итерации полностью вычисляется новая таблица, после чего таблица предыдущей итерации становится ненужной и может быть «забыта». Более экономным с вычислительной точки зрения является модифицированный симплекс-метод, согласно которому переход от одной итерации к другой осуществляется на основе одной и той же начальной таблицы. Кроме того, на каждой итерации модифицированного симплекс-метода вычисляются не все элементы таблицы, а лишь нулевая строка, нулевой и ведущий столбцы, знание которых позволяет перейти от старого набора базисных переменных к новому. Рассмотрим один из упрощенных вариантов этого метода.

Предположим, что каноническая задача линейного программирования уже приведена к диагональному виду (3.9) и ей соответствует начальная табл. 3.9.

Таблица 3.9

		x_1	x_2	...	x_k	x_{k+1}	...	x_n
x_0	a_{00}	0	0		0	$a_{0,\,k+1}$		a_{0n}
x_1	a_{10}	1	0		0	$a_{1,\,k+1}$		a_{1n}
x_2	a_{20}	0	1		0	$a_{2,\,k+1}$		a_{2n}
. .	. .	.	.	...	.	. .	...	.
x_k	a_{k0}	0	0		1	$a_{k,\,k+1}$		a_{kn}

Прежде чем рассмотреть модифицированный симплекс-метод, отметим следующее. Допустим, что ведущий столбец имеет номер s, а ведущая строка — номер r (т. е. a_{rs} — ведущий элемент). Шаг 3 обычного симплекс-метода заключается в применении элементарных преобразований, позволяющих привести исходную систему к диагональной форме относительно нового набора переменных, в котором вместо x_r участвует переменная x_s. Непосредственной проверкой можно убедиться в том, что такой переход к новой диагональной форме (новой таблице) можно осуществить, умножая табл. 3.9 слева на матрицу

$$M_{rs}=\begin{array}{c}\begin{pmatrix}1 & 0 & 0\cdot\cdot\cdot 0 & -a_{0s}/a_{rs} & 0\cdot\cdot\cdot 0\\ 0 & 1 & 0\cdot\cdot\cdot 0 & -a_{1s}/a_{rs} & 0\cdot\cdot\cdot 0\\ \cdots & \cdots & \cdots & \cdots & \cdots\\ 0 & 0 & 0\cdot\cdot\cdot 1 & -a_{r-1,s}/a_{rs} & 0\cdot\cdot\cdot 0\\ 0 & 0 & 0\cdot\cdot\cdot 0 & 1/a_{rs} & 0\cdot\cdot\cdot 0\\ 0 & 0 & 0\cdot\cdot\cdot 0 & -a_{r+1,s}/a_{rs} & 1\cdot\cdot\cdot 0\\ \cdots & \cdots & \cdots & \cdots & \cdots\\ 0 & 0 & 0\cdot\cdot\cdot 0 & -a_{ks}/a_{rs} & 0\cdot\cdot\cdot 1\end{pmatrix}\end{array}\begin{array}{l}\text{строка } 0\\ \\ \\ \\ \text{строка } r\\ \\ \\ \\ \end{array}$$

которая отличается от единичной матрицы $(k+1)$-го порядка лишь столбцом r. (Имеется в виду умножение по обычным правилам умножения матриц.)

Начало *алгоритма модифицированного симплекс-метода* такое же, как и в обычном симплекс-методе. Прежде всего выявляют положительные элементы нулевой строки табл. 3.9 (не считая a_{00}) и фиксируют некоторый из них, например $a_{0s}>0$. Далее, в ведущем столбце s среди положительных элементов из условия минимума отношения a_{i0}/a_{is} $(i:a_{is}>0)$ определяют ведущий элемент a_{rs}. Вместо переменной x_r в число базисных должна войти переменная x_s. На этом завершается нулевая итерация алгоритма модифицированного симплекс-метода.

Первая итерация. Последовательно умножают нулевую строку матрицы M_{rs} на столбцы табл. 3.9 и тем самым вычисляют элементы $a_{0i}^{(1)}$, $i=1, 2, ..., n$, нулевой строки новой таблицы. Если среди этих элементов не окажется положительных, то остается вычислить новый нулевой столбец, который определит оптимальные значения базисных переменных. В противном случае выбирают положительный элемент $a_{0l}^{(1)}$, определяющий ведущий столбец. Последовательно умножая строки* матрицы M_{rs} начиная с первой на нулевой столбец и столбец l исходной табл. 3.9, вычисляют новые элементы $a_{j0}^{(1)}$, $j=1, 2, ..., k$, нулевого столбца и элементы $a_{jl}^{(1)}$, $j=1, 2, ..., k$ столбца l. Остальные элементы новой таблицы далее не понадобятся. Сравнивая отношения $a_{i0}^{(1)}/a_{il}^{(1)}$ $(i:a_{il}^{(1)}>0)$, находят ведущий элемент. Если ведущий элемент определяется неоднозначно, то поступать нужно так же, как в обычном симплекс-методе, т. е. выбрать в качестве ведущего элемент строки, номер которого удовлетворяет условию минимума отношения $a_{i0}^{(1)}/a_{il}^{(1)}$ и является наименьшим номером. Или же следует использовать прави-

* Точнее говоря, следует умножать строки матрицы M_{rs} без левых элементов, соответствующих нулевому столбцу.

ло предотвращения зацикливания. Заметим, что использование этого правила предполагает предварительное вычисление соответствующих столбцов таблицы, которое также выполняется с помощью матрицы M_{rs}. Пусть $a_{tl}^{(1)}$ — новый ведущий элемент. Первая итерация закончена.

Новая матрица M_{tl} будет отличаться от единичной лишь столбцом t:

$$\begin{pmatrix} -a_{0l}^{(1)}/a_{tl}^{(1)} \\ -a_{1l}^{(1)}/a_{tl}^{(1)} \\ \vdots \\ -a_{t-1,l}^{(1)}/a_{tl}^{(1)} \\ 1/a_{tl}^{(1)} \\ -a_{t+1,l}^{(1)}/a_{tl}^{(1)} \\ \vdots \\ -a_{kl}^{(1)}/a_{tl}^{(1)} \end{pmatrix} \begin{array}{l} \\ \\ \\ \\ \text{строка } t \\ \\ \\ \\ \end{array}$$

С помощью матрицы M_{tl} осуществляется замена базисной переменной x_t на переменную x_l. Для этого следует таблицу, полученную в результате первой итерации, умножить слева на M_{tl}. Указанная таблица сама является произведением M_{rs} и табл. 3.9, поэтому будет получен тот же самый результат, если начальную табл. 3.9 умножить слева на произведение матриц $M_{tl}M_{rs}$.

Таким образом, следует вычислить матрицу $M_{tl}M_{rs}$ и последовательно умножить нулевую строку этой матрицы на столбцы исходной табл. 3.9. В результате получают нулевую строку таблицы второй итерации. Далее, вновь определяют ведущий элемент $a_{pq}^{(2)}$ (при условии, что в нулевой строке есть положительный элемент), составляют матрицу M_{pq} и вычисляют произведение $M_{pq}(M_{tl} \cdot M_{rs})$, с помощью которого можно определить элементы таблицы третьей итерации, и т. д.

Окончательно либо будет найдено оптимальное решение, либо установлено, что исходная задача неразрешима.

2. Пример. Проиллюстрируем применение модифицированного симплекс-метода на примере задачи, решенной ранее обычным симплекс-методом. Начальная таблица этой задачи имеет вид табл. 3.10.

Таблица 3.10

		4	0	0	0	2	0	2
←	x_1	1	1	0	0	2	−1	1
	x_2	2	0	1	0	0	1	1
	x_3	1	0	0	1	−1	1	−1

В качестве ведущего выберем четвертый столбец. Ведущим элементом является $a_{14}=2$. Составим матрицу

$$M_{14}=\begin{pmatrix}1 & -1 & 0 & 0\\ 0 & 1/2 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 1/2 & 0 & 1\end{pmatrix}$$

и вычислим элементы новой нулевой строки:

$$(1\ -1\ 0\ 0)\cdot\begin{pmatrix}4 & 0 & 0 & 0 & 2 & 0 & 2\\ 1 & 1 & 0 & 0 & 2 & -1 & 1\\ 2 & 0 & 1 & 0 & 0 & 1 & 1\\ 1 & 0 & 0 & 1 & -1 & 1 & -1\end{pmatrix}=(3\ -1\ 0\,0\,0\,1\,1).$$

Возьмем теперь в качестве ведущего, например, пятый столбец и найдем элементы нового нулевого и нового пятого столбцов. Имеем

$$\begin{pmatrix}1/2 & 0 & 0\\ 0 & 1 & 0\\ 1/2 & 0 & 1\end{pmatrix}\cdot\begin{pmatrix}1\\ 2\\ 1\end{pmatrix}=\begin{pmatrix}1/2\\ 2\\ 3/2\end{pmatrix},$$

$$\begin{pmatrix}1/2 & 0 & 0\\ 0 & 1 & 0\\ 1/2 & 0 & 1\end{pmatrix}\cdot\begin{pmatrix}-1\\ 1\\ 1\end{pmatrix}=\begin{pmatrix}-1/2\\ 1\\ 1/2\end{pmatrix}.$$

Новая таблица принимает вид табл. 3.11.

Таблица 3.11

	3	−1	0	0	0	1	1
x_4	1/2					−1/2	
← x_2	2					1	
x_3	3/2					1/2	

Пустые места в этой таблице соответствуют тем элементам, знание которых в данном случае не обязательно. Сравнивая отношения $2:1$ и $^3/_2:{}^1/_2$, видим, что переменную x_2 следует вывести из базиса. Составим новую матрицу

$$M_{25}=\begin{pmatrix}1 & 0 & -1 & 0\\ 0 & 1 & 1/2 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & -1/2 & 1\end{pmatrix}$$

и вычислим произведение:

$$M_{25}\cdot M_{14}=\begin{pmatrix}1 & -1 & -1 & 0\\ 0 & 1/2 & 1/2 & 0\\ 0 & 0 & 1 & 0\\ 0 & 1/2 & -1/2 & 1\end{pmatrix}.$$

Находим элементы новой нулевой строки:

$$(1\ -1\ -1\ 0)\cdot\begin{pmatrix}4 & 0 & 0 & 0 & 2 & 0 & 2\\ 1 & 1 & 0 & 0 & 2 & -1 & 1\\ 2 & 0 & 1 & 0 & 0 & 1 & 1\\ 1 & 0 & 0 & 1 & -1 & 1 & -1\end{pmatrix}=(1\ -1\ -1\ 0\ 0\ 0\ 0),$$

В этой строке положительных элементов (не считая левого) нет, поэтому минимальное значение равно 1. Наконец, имеем

$$\begin{pmatrix}1/2 & 1/2 & 0\\ 0 & 1 & 0\\ 1/2 & -1/2 & 1\end{pmatrix}\cdot\begin{pmatrix}1\\ 2\\ 1\end{pmatrix}=\begin{pmatrix}3/2\\ 2\\ 1/2\end{pmatrix}.$$

Таким образом, оптимальное решение имеет вид $x_1=x_2=0$, $x_3={}^1/_2$, $x_4={}^3/_2$, $x_5=2$, $x_6=0$.

§ 3.3. ДВОЙСТВЕННЫЙ СИМПЛЕКС-МЕТОД

1. Идея двойственного симплекс-метода. Рассмотрим следующую задачу линейного программирования:

$$x_0=x_3+2x_4\to\min;$$

$$x_1+\ x_3+x_4=1,$$

$$x_2-2x_3-x_4=-1,$$

$$x_1,\ x_2,\ x_3,\ x_4\geqslant 0.$$

Записав ее в виде

$$\begin{aligned}x_0\qquad\quad - x_3-2x_4&=0,\\ x_1\quad + x_3+\ x_4&=1,\\ x_2-2x_3-\ x_4&=-1,\end{aligned}$$

приходим к диагональной относительно переменных x_0, x_1, x_2 форме, которой соответствует табл. 3.12.

В табл. 3.12 в нулевой строке нет положительных элементов, однако сделать вывод об оптимальности нельзя, поскольку текущее базисное решение не является допустимым, так как $x_2=-1<0$. Для того чтобы решить данную задачу с помощью симплекс-метода, ее нужно предварительно привести к диагональной форме с неотрицательными свободными членами (элементами нулевого столбца). Но мы поступим иначе. Попробуем сделать базисное решение допустимым. Для этого переменную x_2 следует вывести из числа базисных. Вместо нее можно ввести переменную x_3 или

Таблица 3.12

x_0	0	0	0	−1	−2
x_1	1	1	0	1	1
← x_2	−1	0	1	−2	−1

x_4. Постараемся избежать появления положительных элементов в нулевой строке. Для этого сравним отношения $a_{03}/a_{23}=-1/-2=$ $={}^1/_2$, $a_{04}/a_{24}=-2/-1=2$ и выберем в качестве ведущего третий столбец, для которого такое отношение имеет наименьшее значение. Теперь, как и в симплекс-методе, введем в число базисных вместо x_2 переменную x_3. В результате преобразований получим табл. 3.13.

Таблица 3.13

x_0	0,5	0	−0,5	0	−1,5
x_1	0,5	1	0,5	0	2,5
x_3	0,5	0	−0,5	1	0,5

В табл. 3.13 все элементы нулевого столбца не отрицательны и среди элементов нулевой строки (не считая $a_{00}=0,5$) нет положительных. Следовательно, текущее базисное решение $x_1=0,5$, $x_2=$ $=x_4=0$, $x_3=0,5$ является оптимальным.

Способ, с помощью которого была решена исходная задача, называется *двойственным симплекс-методом.*

2. Алгоритм двойственного симплекс-метода. Дадим строгое описание двойственного симплекс-метода.

Начинать вычисления следует с составления таблицы, которая соответствует диагональной форме относительно x_0 и некоторых k (базисных) переменных из общего списка переменных $x_1, x_2, \ldots, x_n$. В этой таблице все элементы нулевой строки (не считая a_{00}) должны быть не положительны: $a_{0i}\leqslant 0$, $i=1, 2, \ldots, n$.

Шаг 1. Если среди элементов нулевого столбца (не считая a_{00}) нет отрицательных, то текущее базисное решение является оптимальным. В противном случае среди отрицательных элементов нулевого столбца следует выбрать некоторый элемент $a_{r0}<0$. Таким образом, r-я строка таблицы — ведущая.

Шаг 2. Среди отрицательных элементов ведущей строки найти элемент, который реализует минимум отношения a_{0j}/a_{rj} $(j : a_{rj}<$ $<0)$. Пусть это элемент a_{rs}. Значит, ведущим является столбец s.

Если в ведущей строке не окажется ни одного отрицательного элемента (кроме a_{r0}), то исходная задача не имеет допустимых решений. В самом деле, в этом случае уравнение, соответствующее ведущей строке с номером r, имеет вид

$$a_{r1}x_1+a_{r2}x_2+\ldots+a_{rn}x_n=a_{r0},$$

где $a_{rj}\geqslant 0$, $j=1, 2, \ldots, n$; $a_{r0}<0$. Этому уравнению не могут удовлетворить никакие допустимые значения переменных $x_i\geqslant 0$, $i=1, 2, \ldots, n$.

Отметим, что если минимум отношения a_{0j}/a_{rj} соответствует сразу нескольким столбцам, то обычно в качестве ведущего выбирают столбец с наименьшим номером.

Шаг 3. Так же как на шаге 3 прямого симплекс-метода, описанного в § 3.1, вместо переменной x_r в число базисных ввести переменную x_s и вернуться к шагу 1.

Последовательное выполнение шагов 1—3 составляет одну итерацию двойственного симплекс-метода. После конечного числа итераций либо будет найдено оптимальное решение, либо установлено, что данная задача не имеет допустимых решений.

В идейном отношении алгоритм двойственного симплекс-метода имеет много общего с алгоритмом прямого симплекс-метода. Образно говоря, двойственный симплекс-метод — это прямой симплекс-метод, взятый со знаком минус, в котором роль строк играют столбцы, а роль столбцов — строки. В самом деле, в прямом симплекс-методе вычисления нужно начинать с составления таблицы, в которой элементы нулевого столбца не отрицательны; в двойственном — с таблицы, в которой элементы нулевой строки не отрицательны. В прямом методе сначала определяют ведущий столбец, а затем ведущую строку; в двойственном методе — наоборот: сначала ведущую строку, а затем ведущий столбец. Признак оптимальности текущей симплексной таблицы, который позволяет установить момент окончания вычислений, для обоих методов одинаков: неотрицательность элементов нулевого столбца и неположительность элементов нулевой строки.

Отметим, что существует и модифицированный вариант двойственного симплекс-метода.

§ 3.4. ПРИКЛАДНЫЕ ЗАДАЧИ

1. Транспортная задача. Пусть имеется m поставщиков некоторой однородной продукции. Считается, что i-й поставщик располагает a_i единицами продукции, $i=1, 2, \ldots, m$. Продукцию поставщиков используют n потребителей, потребности которых характеризуются числами $b_1, b_2, \ldots, b_n$ (т. е. j-му потребителю необходимо b_j единиц продукции). Все числа a_i, b_j положительны и считаются заданными. Кроме того, известна стоимость перевозки единицы продукции от i-го поставщика j-му потребителю: c_{ij}, $i=1, 2, \ldots, m$; $j=1, 2, \ldots, n$. Продукция однородна, поэтому любой поставщик может предложить ее любому потребителю (рис. 3.4). Требуется определить такое количество единиц продукции, перевозимой от каждого поставщика к каждому потребителю, чтобы транспортные расходы были минимальными и потребности всех потребителей были удовлетворены. Задача предполагается разрешимой в том смысле, что общий объем возможных поставок должен быть не меньше общего объема потребностей.

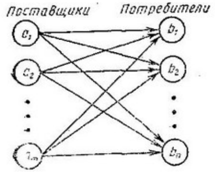

Рис. 3.4

Обозначим через x_{ij} число единиц продукции, перевозимой от i-го поставщика j-му потребителю. Тогда транспортные расходы этой пары поставщик-потребитель составят $c_{ij}x_{ij}$, а общая стоимость всех перевозок (от каждого поставщика каждому потребителю) равна $\sum\limits_{i=1}^{m}\sum\limits_{j=1}^{n} c_{ij}x_{ij}$. Тот факт, что потребности каждого потребителя должны быть удовлетворены, можно выразить в виде неравенств $\sum\limits_{i=1}^{m} x_{ij} \geqslant b_j$, $j=1, 2, \ldots, n$, а условие осуществимости выполнения поставок — в виде неравенств $\sum\limits_{j=1}^{n} x_{ij} \leqslant a_i$, $i=1, 2, \ldots, m$. Кроме того, все величины x_{ij} должны быть подчинены условию неотрицательности.

Таким образом, транспортная задача принимает вид следующей задачи линейного программирования:

$$
\begin{gathered}
\sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij}x_{ij} \to \min; \\
\sum_{j=1}^{n} x_{ij} \leqslant a_i, \quad i=1, 2, \ldots m, \\
\sum_{i=1}^{m} x_{ij} \geqslant b_j, \quad j=1, 2, \ldots, n, \\
x_{ij} \geqslant 0, \quad i=1, 2, \ldots, m; \; j=1, 2, \ldots, n.
\end{gathered} \tag{3.13}
$$

В этой задаче $m \cdot n$ переменных и $m+n$ ограничений типа неравенств (не считая условий неотрицательности переменных).

Если суммарные возможности поставок совпадают с требуемой суммарной потребностью $\left(\text{т. е. } \sum\limits_{i=1}^{m} a_i = \sum\limits_{j=1}^{n} b_j\right)$, то можно так организовать перевозки, чтобы вся имеющаяся у поставщиков продукция была вывезена и каждый потребитель получил необходимое количество продукции. В этом случае транспортная задача принимает вид канонической задачи линейного программирования

$$\sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij}x_{ij} \to \min;$$

$$\sum_{j=1}^{n} x_{ij}=a_i, \quad i=1, 2, \ldots, m, \tag{3.14}$$

$$\sum_{i=1}^{m} x_{ij}=b_j, \quad j=1, 2, \ldots, n,$$

$$x_{ij} \geqslant 0, \quad i=1, 2, \ldots, m; \quad j=1, 2, \ldots, n$$

и называется *закрытой транспортной задачей.*

Если $\sum_{i=1}^{m} a_i > \sum_{j=1}^{n} b_j$, то спрос потребителей будет удовлетворен, но у поставщиков останется «лишняя» продукция и ограничения транспортной задачи имеют вид

$$\sum_{j=1}^{n} x_{ij} \leqslant a_i, \quad i=1, 2, \ldots, m,$$

$$\sum_{i=1}^{m} x_{ij}=b_j, \quad j=1, 2, \ldots, n.$$

Если безразлично, у каких именно поставщиков и в каком количестве останется неотправленная продукция, то, вводя фиктивного $(n+1)$-го потребителя, потребность которого $\sum_{i=1}^{m} a_i - \sum_{j=1}^{n} b_j$, мы формально придем к закрытой транспортной задаче. А для того чтобы введение фиктивного потребителя не повлияло на суммарные расходы по перевозкам, следует принять $c_{i,\,n+1}=0$, $i=1, 2, \ldots, m$. Случай $\sum_{i=1}^{m} a_i < \sum_{j=1}^{n} b_j$ соответствует ситуации, когда потребности потребителей не могут быть удовлетворены полностью. Этот случай с самого начала был исключен из рассмотрения, хотя как с практической, так и с теоретической точек зрения он не лишен интереса.

Отметим, что *закрытая транспортная задача всегда имеет оптимальное решение* [3]. Это решение можно найти с помощью симплекс-метода. Однако обычно для решения транспортной задачи используют специальные методы, учитывающие ее специфику [3].

2. Задача об оптимальном распределении деталей по станкам. Пусть некоторый прибор состоит из k различных видов деталей, которые мы занумеруем числами 1, 2, ..., k. Имеется m типов различных станков, причем количество станков i-го типа равно a_i, $i=1, 2, ..., m$. Детали могут быть изготовлены на станках разного типа. Производительность станка i-го типа при изготовлении j-й детали составляет c_{ij}. После изготовления детали поступают на сборку. Требуется закрепить станки за деталями так, чтобы в единицу времени получать максимальное количество приборов.

Пусть x_{ij} — количество станков i-го типа, на которых можно изготовить j-ю деталь. Очевидно, количество станков i-го типа, изготовляющих детали 1, 2, ..., k видов, не должно превышать заданное число a_i:

$$\sum_{j=1}^{k} x_{ij} \leqslant a_i, \quad i=1, 2, \ldots, m. \tag{3.15}$$

Общее количество деталей j-го вида, изготовленное на станках за единицу времени, составляет $\sum_{i=1}^{m} c_{ij}x_{ij}$, $j=1, 2, ..., k$. В каждом приборе имеется ровно одна деталь с номером j ($j=1, 2, ..., k$), поэтому, для того чтобы не было изготовлено лишних и не было дефицита деталей, должны выполняться условия комплектности:

$$\sum_{i=1}^{m} c_{ij}x_{ij} = \sum_{i=1}^{m} c_{il}x_{il}, \quad j \neq l; \; j, \, l=1, 2, \ldots, k. \tag{3.16}$$

Общее количество комплектов деталей, необходимых для сборки прибора, равно общему количеству какой-либо одной детали, имеющей, например, номер 1. Поэтому решение задачи заключается в максимизации линейной функции $\sum_{i=1}^{m} c_{i1}x_{i1}$ при ограничениях (3.15)—(3.16) с дополнительным условием неотрицательности всех переменных x_{ij}.

Найденные оптимальные значения $x_{ij}^{(0)}$ этой задачи не обязательно целые числа. Например, $x_{11}^{(0)}=2{,}5$ означает, что на двух станках первого типа в течение единицы времени будут изготовлять деталь с номером 1, тогда как третий станок того же типа будет работать лишь половину указанного времени.

3. Задача о загрузке оборудования. Имеется m станков, на которых могут быть изготовлены n типов деталей. Производительность i-го станка при изготовлении детали j-го типа составляет c_{ij}. Величины плановых заданий A_j на изготовление j-й детали и ресурс времени B_i работы i-го станка приведены в табл. 3.14.

Требуется, учитывая ресурсы времени работы каждого станка распределить задания между станками таким образом, чтобы общее время работы всех станков было минимальным.

Пусть t_{ij} — время изготовления j-й детали i-м станком. Составим ограничения по ресурсу времени для каждого станка:

$$\sum_{j=1}^{n} t_{ij} \leqslant B_i, \quad i=1, 2, \ldots, m. \tag{3.17}$$

Условия выполнения плановых заданий имеют вид

$$\sum_{i=1}^{m} c_{ij} t_{ij} = A_j, \quad j=1, 2, \ldots, n. \tag{3.18}$$

Математическая модель сформулированной задачи состоит в минимизации линейной целевой функции (суммарное время) $\sum_{i=1}^{m} \sum_{j=1}^{n} t_{ij}$

Таблица 3.14

Станки	Тип деталей				Ресурс времени
	1	2	...	n	
1	c_{11}	c_{12}	...	c_{1n}	B_1
2	c_{21}	c_{22}	...	c_{2n}	B_2
.	..	..	...	...	.
.					.
.					.
m	c_{m1}	c_{m2}	...	c_{mn}	B_m
Требуемое количество деталей	A_1	A_2	...	A_n	

при ограничениях (3.17)—(3.18) и условии неотрицательности всех переменных t_{ij}.

4. Задача о производстве продукции при ограниченных запасах сырья. Из n видов сырья производится m различных типов продукции. Стоимость реализации изготовленной продукции i-го типа составляет α_i, $i=1, 2, \ldots, m$. Запас сырья j-го вида на планируемый период равен β_j, $j=1, 2, \ldots, n$. Потребность в сырье j-го вида для производства продукта i-го типа на данный период составляет p_{ij}. В ходе производства могут образовываться побочные продукты, представляющие собой сырье для других продуктов. Это приводит к тому, что среди коэффициентов p_{ij} могут быть и отрицательные. Исходные данные задачи приведены в табл. 3.15.

Требуется для каждого типа продукта $i=1, 2, \ldots, m$ определить такой объем производства x_i, чтобы обеспечить максимальную стоимость реализации изготовленной продукции при условии непревышения запасов имеющегося сырья.

Стоимость продукции равна

$$\sum_{i=1}^{m} \alpha_i x_i; \tag{3.19}$$

ограничения по запасам сырья имеют вид

$$\sum_{i=1}^{m} p_{ij} x_i \leqslant \beta_j, \quad j=1, 2, \ldots, n. \tag{3.20}$$

Таблица 3.15

Тип продукции	Вид сырья				Стоимость продукции
	1	2	....	n	
1	p_{11}	p_{12}	...	p_{1n}	α_1
2	p_{21}	p_{22}	...	p_{2n}	α_2
$\vdots$	...	.	...	...	$\vdots$
m	p_{m1}	p_{m2}	...	p_{mn}	α_m
Запас сырья	β_1	β_2	...	β_n	

Таким образом, задача заключается в максимизации линейной функции (3.19) при ограничениях (3.20) и дополнительном условии $x_i \geqslant 0$, $i=1, 2, \ldots, m$.

Глава 4
ДИНАМИЧЕСКОЕ ПРОГРАММИРОВАНИЕ

В настоящее время под динамическим программированием понимают методы решения оптимизационных задач, в основе которых лежит идея разбиения исходной задачи на последовательный ряд более простых задач. Основной областью приложения динамического программирования являются многошаговые процессы, т. е. процессы, протекающие во времени (дискретном или непрерывном).

В настоящей главе динамическое программирование рассматривается применительно к решению задачи оптимального управления дискретной системой. В § 4.1 напоминается формулировка этой задачи, а также приводятся различные ее разновидности. В § 4.2 на примере демонстрируется идея динамического программирования, а затем формулируются и доказываются условия оптимальности в форме рекуррентных соотношений динамического программирования. Эти условия являются теоретической основой численного метода решения задачи оптимального управления дискретной системой, который приводится в § 4.3. Обсуждаются достоинства и недостатки численного метода.

§ 4.1. ЗАДАЧА ОПТИМАЛЬНОГО УПРАВЛЕНИЯ И ЕЕ РАЗНОВИДНОСТИ

1. Постановка задачи оптимального управления. Напомним формулировку задачи оптимального управления дискретной системой.

Изменение состояний системы происходит в дискретные моменты времени t и описывается уравнениями

$$x^{(t)}=g^{(t)}(x^{(t-1)}, u^{(t)}), \quad t=1, 2, \ldots, N, \tag{4.1}$$

где $x^{(t)}$—n-мерный *вектор состояния (фазовый вектор)*, а $u^{(t)}$—r-мерный *вектор управления* в момент времени t, $t=1, 2, \ldots, N$. Векторная функция $g^{(t)}$ считается заданной на множестве $\mathbf{R}^n\times\mathbf{R}^r$, т. е. $g^{(t)}: \mathbf{R}^n\times\mathbf{R}^r\rightarrow\mathbf{R}^n$, $t=1, 2, \ldots, N$. Управления $u^{(t)}$ могут выбираться из соответствующих *множеств допустимых управлений*

$$u^{(t)}\in U_t\subset\mathbf{R}^r, \quad t=1, 2, \ldots, N. \tag{4.2}$$

Чтобы была возможность выбора, обычно предполагают, что каждое множество U_t содержит по крайней мере два элемента.

Соотношения (4.1)—(4.2) определяют *дискретную управляемую систему. Начальное состояние* $x^{(0)}$ считается заранее заданным. Управления $u^{(1)}, u^{(2)}, \ldots, u^{(N)}$, удовлетворяющие (4.2), образуют *последовательность допустимых управлений*, которую будем обозначать через $\mathfrak{U}$. Последовательности допустимых управлений при фиксированном начальном состоянии, согласно (4.1), соответствует определенная последовательность состояний $x^{(0)}, x^{(1)}, \ldots, x^{(N)}$, называемая *траекторией системы.* Траекторию будем обозначать через $\mathfrak{X}$.

Управление системой осуществляется следующим образом. При $t=0$ система находится в состоянии $x^{(0)}$. В результате выбора управления $u^{(1)}\in U_1$, согласно (4.1), однозначно определится состояние $x^{(1)}$ в момент времени $t=1$. Дальнейший выбор управления $u^{(2)}\in U_2$ в силу (4.1) приведет систему в некоторое состояние $x^{(2)}$ и т. д. Окончательно последовательность допустимых управлений $\mathfrak{U}=\{u^{(1)}, u^{(2)}, \ldots, u^{(N)}\}$ определит траекторию системы $\mathfrak{X}=\{x^{(0)}, x^{(1)}, \ldots, x^{(N)}\}$, соответствующую данной последовательности управлений. Выбору некоторой другой последовательности допустимых управлений $\mathfrak{U}'$ будет отвечать, вообще говоря, новая траектория системы $\mathfrak{X}'$.

В качестве *критерия оптимальности* I будем рассматривать критерий вида

$$I=\sum_{t=1}^{N} f^t(x^{(t-1)}, u^{(t)}), \tag{4.3}$$

где числовые функции f^t, $t=1, 2, \ldots, N$, считаются заданными на множестве $\mathbf{R}^n\times\mathbf{R}^r$.

Задача оптимального управления состоит в том, чтобы для заданного начального состояния $x^{(0)}$ среди всех последовательностей допустимых управлений $\mathfrak{U}$ найти такую последовательность $\overline{\mathfrak{U}}$ и соответствующую ей траекторию, которые доставляют наименьшее возможное значение критерию оптимальности (4.3). При этом $\overline{\mathfrak{U}}$ будем называть *оптимальной последовательностью управлений,* а отвечающую этой последовательности управлений траекторию — *оптимальной траекторией.*

2. Условие разрешимости задачи оптимального управления. Согласно (4.1), при $t=1$ верно равенство $x^{(1)}=g^{(1)}(x^{(0)}, u^{(1)})$. Поэтому на основании (4.1) при $t=2$ можно записать

$$x^{(2)}=g^{(2)}(x^{(1)}, u^{(2)})=g^{(2)}(g^{(1)}(x^{(0)}, u^{(1)}), u^{(2)}).$$

Последовательно используя равенства (4.1), все остальные состояния $x^{(3)}, \ldots, x^{(N-1)}$ можно выразить через $u^{(1)}, u^{(2)}, \ldots, u^{(N-1)}$. Подставляя найденные выражения для $x^{(1)}, x^{(2)}, \ldots, x^{(N-1)}$ в правую часть равенства (4.3), получаем критерий оптимальности I, зависящий только от управлений $u^{(1)}, u^{(2)}, \ldots, u^{(N)}$ (напомним, что $x^{(0)}$ фиксировано). При условии непрерывности всех функций $g_1^{(t)}, g_2^{(t)}, \ldots, g_n^{(t)}$* и f^t, $t=1, 2, \ldots, N$, на $\mathbf{R}^n\times\mathbf{R}^r$ полученный критерий оптимальности представляет собой непрерывную функцию $r\cdot N$ переменных $u_1^{(1)}, \ldots, u_r^{(1)}, u_r^{(2)}, \ldots, u_r^{(2)}, \ldots, u_1^{(N)}, \ldots, u_r^{(N)}$ на множестве $U_1\times U_2\times \ldots \times U_N$. При этом если указанное декартово произведение — непустой компакт (это заведомо выполнено, если каждое из множеств $U_1, U_2, \ldots, U_N$ является непустым компактом), то согласно теореме Вейерштрасса (см. § 2.2) существует набор управлений $\overline{u}^{(1)}, \overline{u}^{(2)}, \ldots, \overline{u}^{(N)}$, доставляющий наименьшее возможное значение критерию оптимальности I. Таким образом, имеет место следующее утверждение.

Теорема 4.1. *Пусть все функции* $g_1^{(t)}, g_2^{(t)}, \ldots, g_n^{(t)}$, f^t, $t=1, 2, \ldots, N$, *непрерывны на* $\mathbf{R}^n\times\mathbf{R}^r$ *и непустые множества допустимых управлений* $U_1, U_2, \ldots, U_N$ *компактны. Тогда задача оптимального управления разрешима.*

В частности, задача оптимального управления разрешима всегда, если каждое из множеств $U_1, U_2, \ldots, U_N$ содержит конечное число элементов.

3. Разновидности задачи оптимального управления. Нередко в задаче оптимального управления вместо критерия оптимальности I (4.3) рассматривают критерий оптимальности, зависящий только от конечного состояния:

$$I=\varphi(x^{(N)}). \tag{4.4}$$

Возникающую при этом задачу называют *задачей оптимизации конечного состояния* или *задачей терминального управления.* За-

* Эти функции представляют собой компоненты вектор-функции $g^{(t)}$.

дачу терминального управления легко свести к задаче оптимального управления с критерием оптимальности (4.3), если положить

$$f^t\,(x^{(t-1)},\ u^{(t)})\equiv 0,\ t=1,2,\dots,N-1;\ f^N(x^{(N-1)},\ u^{(N)})=$$
$$=\varphi\,(g^{(N)}\,(x^{(N-1)},\ u^{(N)})).$$

С другой стороны, задача оптимального управления с критерием оптимальности (4.3) может быть приведена к некоторой задаче терминального управления. Для этого следует увеличить размерность вектора состояния на единицу, введя дополнительную переменную согласно следующему правилу:

$$x_{n+1}^{(0)}=0;\ x_{n+1}^{(t)}=\sum_{i=1}^{t} f^i\,(x^{(i-1)},\ u^{(i)}),\ t=1,2,\dots,N.$$

Тогда к векторному равенству (4.1) добавится еще одно равенство

$$x_{n+1}^{(t)}=x_{n+1}^{(t-1)}+f^t\,(x^{(t-1)},\ u^{(t)}),\quad t=1,2,\dots,N,$$

а суммарный критерий оптимальности принимает вид функции конечного состояния

$$I=x_{n+1}^{(N)}.$$

Таким образом, задачи оптимального управления с критериями оптимальности вида (4.3) и (4.4) эквивалентны.

Иногда в качестве критерия оптимальности рассматривают

$$I=\varphi\,(x^{(N)})+\sum_{t=1}^{N} f^t\,(x^{(t-1)},\ u^{(t)}). \tag{4.5}$$

Задача оптимального управления с критерием (4.5) также эквивалентна задаче с критерием (4.3), поскольку $\varphi\,(x^{(N)})=\varphi\,(g^{(N)}\,(x^{(N-1)},\ u^{(N)}))$.

Существуют задачи с дополнительными ограничениями на переменные состояния:

$$x^{(t)}\in X_t,\quad t=1,2,\dots,N, \tag{4.6}$$

где $X_t\neq\mathbf{R}^n$ хотя бы для одного $t\in\{1,2,\dots,N\}$. Это *задачи оптимального управления с фазовыми ограничениями,* поскольку вектор состояния $x^{(t)}$ называют и фазовым вектором. В частном случае, когда из всех условий (4.6) имеется только одно условие $x^{(N)}\in X_N$, причем множество X_N состоит из единственного элемента, соответствующую задачу называют *задачей оптимального управления с фиксированным конечным состоянием.*

Задачи оптимального управления с фазовыми ограничениями в условиях выполнения предположений теоремы (4.1) и компактности непустых множеств $X_1, X_2, \dots, X_N$ не обязательно имеют решение. Это может быть, например в случае, когда не найдется ни одной траектории системы $x^{(0)}, x^{(1)}, \dots, x^{(N)}$, для которой были бы выполнены все условия (4.6).

В *задаче управления со свободным начальным состоянием* состояние $x^{(0)}$ заранее не фиксируют, а выбирают из заданного множества: $x^{(0)} \in X_0 \subset \mathbb{R}^n$. По существу, $x^{(0)}$ является дополнительным вектором управления и задачу со свободным начальным состоянием легко привести к обычной задаче с фиксированным состоянием.

Встречаются также *задачи с нефиксированной продолжительностью,* где время управления N заранее не задано. Важным примером задач подобного типа является *задача об оптимальном быстродействии.* В этой задаче требуется найти последовательность управлений, переводящую систему из начального состояния в заданное конечное состояние за минимальное возможное число шагов N.

Кроме дискретных управляемых систем вида (4.1)—(4.2), в которых состояние $x^{(t)}$ однозначно определяется предыдущим состоянием $x^{(t-1)}$ и управлением $u^{(t)}$, встречаются системы, имеющие з а п а з д ы в а н и е. В системах с запаздыванием состояние $x^{(t)}$ в момент времени t зависит не только от предыдущего состояния $x^{(t-1)}$, но и от «более ранних» состояний в моменты времени $t-2$, $t-3$ и т. д.

В общем случае *система с запаздыванием на один шаг* описывается следующим образом. Известны состояния $x^{(-1)}$, $x^{(0)}$, а дальнейшее изменение состояний описывают равенства

$$x^{(t)} = g^{(t)}(x^{(t-1)}, x^{(t-2)}, u^{(t)}), \quad t = 1, 2, \ldots, N. \tag{4.7}$$

Критерий оптимальности для систем с запаздыванием на один шаг имеет вид

$$I = \sum_{t=1}^{N} f^t(x^{(t-1)}, x^{(t-2)}, u^{(t)}), \tag{4.8}$$

а выбор управлений стеснен ограничениями (4.2). *Задача оптимального управления системой с запаздыванием* заключается в том, чтобы для заданных состояний $x^{(-1)}$, $x^{(0)}$ среди всех последовательностей допустимых управлений найти такую последовательность и соответствующую ей согласно (4.7) траекторию, которые доставляют минимальное значение критерию оптимальности (4.8). Данную задачу можно привести к задаче оптимального управления, сформулированной в п. 1. Для этого введем дополнительный вектор состояния $y^{(t)} = (x_{n+1}^{(t)}, x_{n+2}^{(t)}, \ldots, x_{2n}^{(t)})$ по правилу

$$y^{(t)} = x^{(t-1)}, \quad t = 0, 1, 2, \ldots, N. \tag{4.9}$$

В новых обозначениях равенства (4.7) можно записать так:

$$x^{(t)} = g^{(t)}(x^{(t-1)}, y^{(t-1)}, u^{(t)}), \quad t = 1, 2, \ldots, N. \tag{4.10}$$

В результате приходим к следующей задаче, в которой $2n$ переменных состояния $x_1, x_2, \ldots, x_n, x_{n+1}, \ldots, x_{2n}$ и r переменных управления $u_1, u_2, \ldots, u^r$. Задана дискретная система управления (4.10), (4.9), (4.2), начальное состояние $\left(x_1^{(0)}, \ldots, x_n^{(0)}, y_1^{(0)}, \ldots, y_n^{(0)}\right)$ и критерий оптимальности

$$I=\sum_{t=1}^{N} f^t\left(x^{(t-1)}, y^{(t-1)}, u^{(t)}\right).$$

Требуется найти оптимальную последовательность управлений $\bar{u}^{(1)}, \bar{u}^{(2)}, \ldots, \bar{u}^{(N)}$ и соответствующую оптимальную траекторию. Эта задача того же типа, что и задача оптимального управления, рассмотренная в п. 1.

Если дискретная система имеет *запаздывание на два шага*, то функции в правых частях равенств (4.7) и (4.8) зависят еще и от $x^{(t-3)}$. В этом случае, для того чтобы соответствующую задачу управления привести к задаче из п. 1, необходимо исходное число переменных состояния $x_1, x_2, \ldots, x_n$ утроить.

Запаздывание может иметь место не только по переменным состояния, но и по переменным управления. Так, запаздывание на один шаг как по состояниям, так и по управлениям будет выражаться в том, что функции в правых частях равенств (4.7) и (4.8) будут зависеть еще и от $u^{(t-1)}$. В этом случае для сведéния получающейся задачи управления к задаче из п. 1 потребуется ввести $n+r$ дополнительных переменных состояния $y^{(t)}=\left(x_{n+1}^{(t)}, x_{n+2}^{(t)}, \ldots, x_{2n}^{(t)}\right)$, $z^{(t)}=\left(x_{2n+1}^{(t)}, \ldots, x_{2n+r}^{(t)}\right)$ согласно (4.9) и $z^{(t)}=u^{(t)}$, $t=1, 2, \ldots, N$.

§ 4.2. МЕТОД ДИНАМИЧЕСКОГО ПРОГРАММИРОВАНИЯ

1. Идея динамического программирования. Для того чтобы минимизировать функцию аддитивного вида $h(u^{(1)}, u^{(2)}, \ldots, u^{(N)})=h_1(u^{(1)})+h_2(u^{(2)})+\ldots+h_N(u^{(N)})$ по $u^{(i)} \in U_i$, $i=1, 2, \ldots, N$, достаточно найти минимальное значение каждого слагаемого $h_i(u^{(i)})$ на соответствующем множестве U_i. Тем самым решение исходной задачи сведется к решению N более простых задач. Это возможно благодаря аддитивной структуре минимизируемой функции и выбору переменной $u^{(i)}$ в пределах соответствующего множества U_i независимо от остальных переменных.

Критерий оптимальности (4.3) в задаче оптимального управления дискретной системой (4.1), (4.2) также обладает аддитивной структурой и управление $u^{(t)}$ выбирают независимо от остальных управлений. Однако минимизация критерия (4.3) не сводится к минимизации каждого слагаемого $f^t(x^{(t-1)}, u^{(t)})$ на множестве U_t, поскольку это слагаемое зависит еще и от $x^{(t-1)}$. В свою очередь, состояние $x^{(t-1)}$, согласно равенству (4.1), зависит от $x^{(t-2)}$ и $u^{(t-1)}$ и т. д. Таким образом, состояние $x^{(t-1)}$ (а значит, и $f^t(x^{(t-1)}, u^{(t)})$) определяется последовательностями предыдущих состояний $x^{(1)}, x^{(2)}, \ldots, x^{(t-2)}$ и управлений $u^{(1)}, u^{(2)}, \ldots, u^{(t-1)}$. Итак, для того чтобы указать оптимальное управление u в момент времени t, необходимо знать последовательности предыдущих состояний и управлений. Тем не менее исходную задачу минимизации критерия оптимальности (4.3) также можно разбить на ряд более простых подзадач, которые решают в определенной последовательности.

Пример. Пусть $n=1$, $r=1$, $N=4$ и задача оптимального управления имеет вид

$$x^{(t)}=(-1)^t x^{(t-1)}u^{(t)}, \quad t=1, 2, 3, 4; \tag{4.11}$$

$$U_t=U=\{-1;\ 1\}, \quad t=1, 2, 3, 4; \ x^{(0)}=1; \tag{4.12}$$

$$I=\sum_{t=1}^{4} x^{(t-1)}2^{u^{(t)}} \to \min. \tag{4.13}$$

На рис. 4.1 изображены все возможные переходы из одного состояния в другое под действием допустимых управлений —1 и 1. Из 16

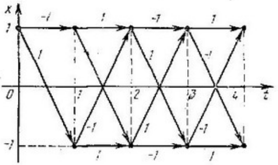

Рис. 4.1

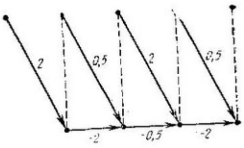

Рис. 4.2

возможных последовательностей допустимых управлений требуется выбрать такую последовательность и соответствующую ей траекторию, которым отвечает наименьшее возможное значение критерия оптимальности (4.13).

Обозначим искомую оптимальную траекторию через 1, $\bar{x}^{(1)}$, $\bar{x}^{(2)}$, $\bar{x}^{(3)}$, $\bar{x}^{(4)}$. Рассмотрим последнее слагаемое в сумме (4.13), т. е. функцию $f^4(x^{(3)}, u^{(4)}) = x^{(3)}2^{u^{(4)}}$. При $t=3$ возможны два состояния: $x^{(3)}=1$ и $x^{(3)}=-1$. Предположим, что оптимальным является первое состояние: $\bar{x}^{(3)}=x^{(3)}=1$. Для этого состояния имеем

$$f^4(x^{(3)}, -1)=0{,}5<2=f^4(x^{(3)}, 1),$$

а значит,

$$\sum_{t=1}^{3} \bar{x}^{(t-1)}2^{\bar{u}^{(t)}}+0{,}5<\sum_{t=1}^{3} \bar{x}^{(t-1)}2^{\bar{u}^{(t)}}+2.$$

Следовательно, если $\bar{x}^{(3)}=1$, то управление $u^{(4)}=1$ заведомо не будет оптимальным; таковым является управление $u^{(4)}=-1$. Поэтому стрелку, которая соединяет точки, соответствующие состояниям $x^{(3)}=1$ и $x^{(4)}=1$, и соответствует управлению $u^{(4)}=1$, можно исключить из дальнейшего рассмотрения (рис. 4.2). На рис. 4.2 также указано значение $f^4(x^{(3)}, -1)=0{,}5$.

Теперь предположим, что $\bar{x}^{(3)}=x^{(3)}=-1$. В этом случае

$$f^4(x^{(3)},\ -1)=-0{,}5>-2=f^4(x^{(3)},\ 1),$$

а значит, управление $u^{(4)}=-1$ (при условии $\bar{x}^{(3)}=-1$) не может быть оптимальным. Этот факт также отражен на рис. 4.2.

Таким образом, для каждого возможного состояния системы при $t=3$ найдено «условно-оптимальное» управление $u^{(4)}$, которое является действительно оптимальным при условии, что оптимальным окажется соответствующее состояние.

При $t=2$ также возможны два состояния: $x^{(2)}=1$ и $x^{(2)}=-1$. Предположим вначале, что $\bar{x}^{(2)}=x^{(2)}=1$. В этом случае выбор управления $u^{(3)}=1$ приводит к состоянию $x^{(3)}=-1$, а в результате действия управления $u^{(3)}=-1$ система перейдет в состояние $x^{(3)}=1$. Поэтому $f^3(x^{(2)},\ -1)+f^4(1,\ -1)=0{,}5+0{,}5>2-2=f^3(x^{(2)},\ 1)+f^4(-1,\ 1)$, т. е. если $\bar{x}^{(2)}=1$, то оптимальным является управление $u^{(3)}=1$. Аналогично находим, что при $\bar{x}^{(2)}=-1$ оптимальным является управление $u^{(3)}=-1$.

Аналогично для $t=1$ получаем, что при $\bar{x}^{(1)}=1$ и $\bar{x}^{(1)}=-1$ соответственно $u^{(2)}=-1$ и $u^{(2)}=1$.

Если $t=0$, то состояние задано однозначно: $x^{(0)}=1$. Поскольку $f^1(x^{(0)},\ -1)+0{,}5-0{,}5-2=-1{,}5>-2{,}5=f^1(x^{(0)},\ 1)-2-0{,}5-2$, оптимальным является управление $u^{(1)}=\bar{u}^{(1)}=1$.

Теперь легко воссоздать всю последовательность оптимальных управлений $\bar{u}^{(1)}=1$, $\bar{u}^{(2)}=1$, $\bar{u}^{(3)}=-1$, $\bar{u}^{(4)}=1$ и соответствующую оптимальную траекторию $x^{(0)}=1$, $\bar{x}^{(1)}=-1$, $\bar{x}^{(2)}=-1$, $\bar{x}^{(3)}=1$, $\bar{x}^{(4)}=-1$ (см. рис. 4.1 и 4.2).

Если задачу (4.11)—(4.13) рассматривать только для $N=3$, то оптимальной последовательностью управлений является $\bar{u}^{(1)}=1$, $\bar{u}^{(2)}=1$, $\bar{u}^{(3)}=1$, где значение $\bar{u}^{(3)}=1$ не совпадает с полученным выше значением $\bar{u}^{(3)}$ для $N=4$.

2. Условия оптимальности. Рассмотрим дискретную управляемую систему из п. 1 предыдущего параграфа:

$$x^{(t)}=g^{(t)}(x^{(t-1)},\ u^{(t)}),\ t=1,2,\dots,N,\quad u^{(t)}\in U_t,\ t=1,2,\dots,N, \tag{4.14}$$

с критерием оптимальности

$$I=\sum_{t=1}^{N} f^t(x^{(t-1)},\ u^{(t)}). \tag{4.15}$$

Предположим, что множества $U_1, U_2, \dots, U_N$ непустые и компактные, а все функции $g_1^{(t)}, g_2^{(t)}, \dots, g_n^{(t)}$, f^t, $t=1, 2, \dots, N$, из (4.14) и (4.15) непрерывные на $\mathbf{R}^n\times\mathbf{R}^r$. При этих предположениях, согласно теореме 4.1, задача оптимального управления разрешима.

Определим рекуррентным способом функции $\omega_N(x)$, $\omega_{N-1}(x),\dots$, $\omega_0(x)$, заданные на $\mathbf{R}^n$:

$$\omega_N(x)\equiv 0,$$

$$\omega_{t-1}(x)=\min_{u\in U_t}[f^t(x,\ u)+\omega_t(g^{(t)}(x,\ u))],\quad t=1, 2,\dots, N. \tag{4.16}$$

Функция $\omega_N(x)$ тождественно равна нулю для всех $x \in \mathbf{R}^n$. Поэтому из равенства (4.16) при $t=N$ имеем

$$\omega_{N-1}(x)=\min_{u \in U_N} f^N(x, u).$$

Далее, используя функцию $\omega_{N-1}(x)$, из формулы (4.16) при $t=N-1$ определяем следующую функцию:

$$\omega_{N-2}(x)=\min_{u \in U_{N-1}} [f^{N-1}(x, u)+\min_{u \in U_N} f^N(x, u)]$$

и т. д. и, наконец, находим $\omega_0(x)$. Функции $g_1^{(t)}, g_2^{(t)}, \ldots, g_n^{(t)}, f^t$, $t=1, 2, \ldots N$, непрерывны, а множества $U_1, U_2, \ldots, U_N$ компактны, поэтому минимум в правой части (4.16) всегда достигается при каждом фиксированном $x \in \mathbf{R}^n$. Следовательно, определение функций $\omega_N, \omega_{N-1}, \ldots, \omega_0$ корректно в том смысле, что равенства (4.16) действительно определяют указанные функции, заданные на всем пространстве $\mathbf{R}^n$.

Динамическим программированием называют метод решения задач оптимального управления, основанный на следующих необходимых и достаточных условиях оптимальности.

Теорема 4.2. *Пусть* $\overline{\mathfrak{U}}=\{\overline{u}^{(1)}, \overline{u}^{(2)}, \ldots, \overline{u}^{(N)}\}$ *— некоторая последовательность допустимых управлений, а* $\overline{\mathfrak{X}}=\{\overline{x}^{(0)}, \overline{x}^{(1)}, \ldots, \overline{x}^{(N)}\}$ *— соответствующая этой последовательности управлений траектория. Для того чтобы последовательность управлений* $\overline{\mathfrak{U}}$ *и траектория* $\overline{\mathfrak{X}}$ *были оптимальными для дискретной системы* (4.14) *с критерием оптимальности* (4.15), *необходимо и достаточно, чтобы для каждого* $t=1, 2, \ldots, N$ *выполнялось равенство*

$$f^t(\overline{x}^{(t-1)}, \overline{u}^{(t)})+\omega_t(g^{(t)}(\overline{x}^{(t-1)}, \overline{u}^{(t)}))=$$

$$=\min_{u \in U_t} [f^t(\overline{x}^{(t-1)}, u)+\omega_t(g^{(t)}(\overline{x}^{(t-1)}, u))]. \tag{4.17}$$

□ Необходимость. Пусть последовательности $\overline{\mathfrak{U}}$ и $\overline{\mathfrak{X}}$ являются оптимальными. Докажем справедливость равенства (4.17) для всех $t=1, 2, \ldots, N$. Предположим противное: для некоторого непустого множества индексов $\Gamma \subset \{1, 2, \ldots, N\}$ равенство (4.17) не имеет места. Обозначим через k наибольший элемент множества Γ. Согласно предположению, существует управление $u^{(k)} \in U_k$, для которого верно неравенство

$$f^k(\overline{x}^{(k-1)}, u^{(k)})+\omega_k(g^{(k)}(\overline{x}^{(k-1)}, u^{(k)}))<$$

$$< f^k(\overline{x}^{(k-1)}, \overline{u}^{(k)})+\omega_k(g^{(k)}(\overline{x}^{(k-1)}, \overline{u}^{(k)})). \tag{4.18}$$

Положим

$$x^{(t)}=\overline{x}^{(t)},\ t=1, 2, \ldots, k-1,\ u^{(t)}=\overline{u}^{(t)},\ t=1, 2, \ldots, k-1. \tag{4.19}$$

Введем вектор $x^{(k)}=g^{(k)}(\bar{x}^{(k-1)}, u^{(k)})$. По определению функции ω_k имеем

$$\omega_k(g^{(k)}(\bar{x}^{(k-1)}, u^{(k)}))=\omega_k(x^{(k)})=\min_{u\in U_{k+1}}[f^{k+1}(x^{(k)}, u)+$$

$$+\omega_{k+1}(g^{(k+1)}(x^{(k)}, u))].$$

Множество U_{k+1} компактно, а функция от u, стоящая под знаком минимума, непрерывна. По теореме Вейерштрасса, существует управление $u^{(k+1)}\in U_{k+1}$, на котором минимум достигается:

$$\omega_k(x^{(k)})=f^{k+1}(x^{(k)}, u^{(k+1)})+\omega_{k+1}(g^{(k+1)}(x^{(k)}, u^{(k+1)}).$$

Теперь введем $x^{(k+1)}=g^{(k+1)}(x^{(k)}, u^{(k+1)})$ и т. д. Аналогично построим последовательности $\mathfrak{U}=\{u^{(1)}, u^{(2)}, \ldots, u^{(N)}\}$, $\mathfrak{X}=\{\bar{x}^{(0)}, x^{(1)}, \ldots, x^{(N)}\}$, для которых выполняются соотношения (4.14) и, кроме того, на основании (4.18) верно неравенство

$$\sum_{t=k}^{N} f^t(x^{(t-1)}, u^{(t)})<f^t(\bar{x}^{(k-1)}, \bar{u}^{(k)})+\omega_k(g^{(k)}(\bar{x}^{(k-1)}, \bar{u}^{(k)})). \quad (4.20)$$

Поскольку k — максимальный элемент множества Г, равенства (4.17) имеют место для $t>k$:

$$f^N(\bar{x}^{(N-1)}, \bar{u}^{(N)})=\omega_{N-1}(\bar{x}^{(N-1)}),$$

$$f^{N-1}(\bar{x}^{(N-2)}, \bar{u}^{(N-1)})+\omega_{N-1}(g^{(N-1)}(\bar{x}^{N-2}, \bar{u}^{(N-1)}))=\omega_{N-2}(\bar{x}^{(N-2)}),$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$f^{k+1}(\bar{x}^{(k)}, \bar{u}^{(k+1)})+\omega_{k+1}(g^{(k+1)}(\bar{x}^{(k)}, \bar{u}^{(k+1)}))=\omega_k(\bar{x}^{(k)}).$$

Складывая почленно эти равенства, получаем

$$\omega_k(\bar{x}^{(k)})=\sum_{t=k+1}^{N} f^t(\bar{x}^{(t-1)}, \bar{u}^{(t)}).$$

Теперь, подставляя значение $\omega_k(\bar{x}^{(k)})$ в (4.20) и учитывая (4.19), приходим к неравенству

$$\sum_{t=1}^{N} f^t(x^{(t-1)}, u^{(t)})<\sum_{t=1}^{N} f^t(\bar{x}^{(t-1)}, \bar{u}^{(t)}).$$

Следовательно, для последовательностей $\mathfrak{U}$ и $\mathfrak{X}$ значение критерия оптимальности меньше, чем для последовательностей $\bar{\mathfrak{U}}$, $\bar{\mathfrak{X}}$. Это противоречит начальному допущению об оптимальности последовательностей $\bar{\mathfrak{U}}$, $\bar{\mathfrak{X}}$.

Достаточность. Пусть для последовательностей $\bar{\mathfrak{U}}$, $\bar{\mathfrak{X}}$ равенство (4.17) верно при каждом $t=1, 2, \ldots, N$. Установим оптимальность данных последовательностей. Для этого рассмотрим произвольную последовательность допустимых управлений $\mathfrak{U}=$

$=\{u^{(1)}, u^{(2)}, \dots, u^{(N)}\}$ и соответствующую ей, согласно (4.14), траекторию $\mathfrak{X}=\{x^{(0)}, x^{(1)}, \dots, x^{(N)}\}$. Из (4.16) при $x=x^{(t-1)}$ имеем

$$\omega_{t-1}(x^{(t-1)}) \leqslant f^t(x^{(t-1)}, u^{(t)}) + \omega_t(g^{(t)}(x^{(t-1)}, u^{(t)})), \ t=1, 2, \dots, N.$$

Складывая почленно эти неравенства и учитывая, что $g^{(t)}(x^{(t-1)}, u^{(t)})=x^{(t)}$, получаем

$$\omega_0(x^{(0)}) \leqslant \sum_{t=1}^{N} f^t(x^{(t-1)}, u^{(t)}). \tag{4.21}$$

С другой стороны, записав равенство (4.17) в виде

$$f^t(\bar{x}^{(t-1)}, \bar{u}^{(t)}) + \omega_t(g^{(t)}(\bar{x}^{(t-1)}, \bar{u}^{(t)})) = \omega_{t-1}(\bar{x}^{(t-1)})$$

и просуммировав по $t=1, 2, \dots, N$, найдем

$$\omega_0(\bar{x}^{(0)}) = \sum_{t=1}^{N} f^t(\bar{x}^{(t-1)}, \bar{u}^{(t)}). \tag{4.22}$$

Начальное состояние у траекторий $\bar{\mathfrak{X}}$ и $\mathfrak{X}$ одно и то же, поэтому $\bar{x}^{(0)}=x^{(0)}$. Следовательно, из (4.21) и (4.22) вытекает неравенство

$$\sum_{t=1}^{N} f^t(\bar{x}^{(t-1)}, \bar{u}^{(t)}) \leqslant \sum_{t=1}^{N} f^t(x^{(t-1)}, u^{(t)}),$$

которое благодаря произвольности выбора $\mathfrak{A}$ означает оптимальность последовательностей управлений $\bar{\mathfrak{A}}$ и состояний $\bar{\mathfrak{X}}$. ■

Для оптимальных последовательностей $\bar{\mathfrak{A}}$ и $\bar{\mathfrak{X}}$, согласно доказанной теореме, выполняется равенство (4.17) при каждом $t=1, 2, \dots, N$. Суммируя это равенство по $t=1, 2, \dots, N$, придем к равенству (4.22), где в правой части стоит оптимальное значение критерия I. Обозначая начальное состояние $\bar{x}^{(0)}$ через $x^{(0)}$, приходим к следующему утверждению.

Следствие 4.1. *Минимальное значение критерия оптимальности* (4.15) *для дискретной системы* (4.14) *с начальным состоянием* $x^{(0)}$ *равно* $\omega_0(x^{(0)})$.

Следствие 4.1 приводит к следующему способу отыскания минимального значения $\omega_0(x^{(0)})$. Сначала, согласно (4.16), следует найти функцию $\omega_{N-1}(x)$. Затем можно определить функции $\omega_{N-2}(x)$ и т. д. и, наконец, $\omega_0(x)$. Вычислив значение функции $\omega_0(x)$ в точке $x^{(0)}$, получим минимальное значение критерия оптимальности. Заметим, что, найдя функцию $\omega_0(x)$, можно легко получить минимальное значение не только для $x^{(0)}$, но и для другого произвольного начального состояния. В простейший случаях, как показывает приведенный ниже пример, указанный подход удается реализовать аналитически.

3. Пример использования теоремы 4.2. Рассмотрим дискретную управляемую систему, в которой $n=2$, $r=1$, $N=10$ и

$$\begin{cases} x_1^{(t)} = x_1^{(t-1)} + x_2^{(t-1)}, \\ x_2^{(t)} = \left(x_1^{(t-1)} - x_2^{(t-1)}\right) u^{(t)}, \end{cases}$$

$$u^{(t)} \in U = [-1,\ 1], \quad t = 1, 2, \ldots, 10, \quad x^{(0)} = (-2,\ 1)^{\mathrm{T}},$$

с критерием оптимальности в виде функции конечного состояния

$$I = x_2^{(10)} = \left(x_1^{(9)} - x_2^{(9)}\right) u^{(10)}.$$

Решим эту задачу терминального управления, используя результаты, полученные в п. 2.

По определению (4.16), верно равенство $\omega_{10}(x_1,\ x_2) = 0$. Далее имеем

$$\omega_9(x_1,\ x_2) = \min_{u \in [-1,1]} [f^{10}(x_1,\ x_2,\ u) + 0] = \min_{u \in [-1,1]} (x_1 - x_2) u.$$

При фиксированных x_1 и x_2 выражение $(x_1 - x_2) u$ представляет собой линейную функцию переменной u. Следовательно, среди точек минимума этой функции обязательно будет хотя бы одна граничная точка множества U, т. е. $u = -1$ или $u = 1$. Поэтому можно записать

$$\omega_9(x_1,\ x_2) = \min_{u \in \{-1;\ 1\}} (x_1 - x_2) u = \min\{x_2 - x_1;\ x_1 - x_2\}.$$

Минимум в правой части равенства достигается при $u = 1$, когда $x_1 - x_2 \leqslant 0$, и при $u = -1$, если $x_1 - x_2 \geqslant 0$. На основании (4.17) нетрудно определить оптимальное управление:

$$u^{(10)} = \begin{cases} 1, \text{ если } x_1^{(9)} - x_2^{(9)} \leqslant 0, \\ -1, \text{ если } x_1^{(9)} - x_2^{(9)} \geqslant 0. \end{cases}$$

В частности, при $x_1^{(9)} - x_2^{(9)} = 0$ оптимальным является любое из управлений 1 или -1.

Теперь можно определить следующую функцию:

$$\omega_8(x_1,\ x_2) = \min_{u \in [-1,1]} [0 + \omega_9(g^{(9)}(x_1,\ x_2,\ u))] =$$

$$= \min_{u \in [-1,1]} \min\{(x_1 - x_2) u - x_1 - x_2;\ x_1 + x_2 - (x_1 - x_2) u\}.$$

Очевидно, функция $\min\{\ldots\}$ переменной u является вогнутой при произвольных фиксированных x_1 и x_2. Значит, среди ее точек минимума обязательно будет хотя бы одна граничная точка множества U. На основании этого имеем

$$\omega_8(x_1,\ x_2) = \min_{u \in \{-1;\ 1\}} \min\{(x_1 - x_2) u - x_1 - x_2;$$

$$x_1 + x_2 - (x_1 - x_2) u\} = \min\{-2x_1;\ -2x_2;\ 2x_1;\ 2x_2\},$$

причем если $x_1 \geqslant x_2$, $x_1 \leqslant -x_2$ или $x_1 \leqslant x_2$, $x_1 \geqslant -x_2$, то минимум достигается в точке $u = 1$, а если $x_1 \geqslant x_2$, $x_1 \geqslant -x_2$ или $x_1 \leqslant x_2$, $x_1 \leqslant$

$\leqslant -x_2$, то в точке $u=-1$. Согласно (4.17), это означает, что оптимальное управление $u^{(9)}$ имеет вид

$$u^{(9)}=\begin{cases} 1, \text{ если } x_1^{(8)}\geqslant x_2^{(8)},\ x_1^{(8)}\leqslant -x_2^{(8)} \text{ или } x_1^{(8)}\leqslant x_2^{(8)},\ x_1^{(8)}\geqslant -x_2^{(8)};\\ -1, \text{ если } x_1^{(8)}\geqslant x_2^{(8)},\ x_1^{(8)}\geqslant -x_2^{(8)} \text{ или } x_1^{(8)}\leqslant x_2^{(8)},\ x_1^{(8)}\leqslant -x_2^{(8)}.\end{cases}$$

Далее аналогично имеем

$$\omega_7(x_1, x_2)=\min_{u\in[-1,1]} \min[0+\omega_8(g^{(8)}(x_1, x_2, u))]=$$

$$=\min_{u\in[-1,1]} \min\{-2(x_1+x_2);\ -2(x_1-x_2)u;\ 2(x_1+x_2);\ 2(x_1-x_2)u\}=$$

$$=\min_{u\in\{-1;1\}} \min\{-2(x_1+x_2);\ -2(x_1-x_2)u;\ 2(x_1+x_2);\ 2(x_1-x_2)u\}=$$

$$=\min\{-2x_1-2x_2;\ 2x_1+2x_2;\ 2x_2-2x_1;\ 2x_1-2x_2\};$$

причем последний минимум достигается на любом из управлений $u=-1$ или $u=1$. Следовательно, в качестве оптимального может быть взято управление $u^{(8)}=-1$ или $u^{(8)}=1$.

Аналогично, в качестве оптимального управления $u^{(7)}$ можно взять как управление $u^{(7)}=-1$, так и $u^{(7)}=1$, причем $\omega_6(x_1, x_2)=\min\{-4x_1; -4x_2; 4x_1; 4x_2\}$ и т. д. На каждом последующем шаге также оптимальным является любое из управлений -1 или 1. При этом $\omega_0(x_1, x_2)=\min\{-32x_1; -32x_2; 32x_1; 32x_2\}$. Находим оптимальное значение:

$$\omega_0(-2,1)=\min\{64;\ -32;\ -64;\ 32\}=-64.$$

Теперь восстановим оптимальную последовательность управлений. Поскольку начиная с $u^{(8)}$ в качестве оптимального можно брать любое из чисел -1 или 1, для определенности примем $u^{(1)}=u^{(2)}=\ldots$ $\ldots=u^{(8)}=1$. Находим соответствующую этим управлениям часть оптимальной траектории: $x^0=(-2,\ 1)^T$, $x^{(1)}=(-1,\ 3)^T$, $x^{(2)}=(2,\ -4)^T$, $x^{(3)}=(-2,\ 6)^T$, $x^{(4)}=(4,\ -8)^T$, $x^{(5)}=(-4,\ 12)^T$, $x^{(6)}=(8,\ -16)^T$, $x^{(7)}=(8,\ -24)^T$, $x^{(8)}=(-16,\ 32)^T$. Так как $x_1^{(8)}\leqslant x_2^{(8)}$, $x_1^{(8)}\geqslant -x_2^{(8)}$, то $u^{(9)}=1$. В этом случае $x^{(9)}=(16,\ -48)^T$. Следовательно, $x_1^{(9)}-x_2^{(9)}\geqslant 0$ и поэтому $u^{(10)}=-1$. Запишем окончательно оптимальную последовательность управлений и соответствующую траекторию, на которых критерий оптимальности имеет минимальное значение, равное -64:

$$\mathfrak{A}=\{1,\ 1,\ 1,\ 1,\ 1,\ 1,\ 1,\ 1,\ 1,\ -1\},$$

$$\mathscr{X}=\left\{\begin{pmatrix}-2\\1\end{pmatrix}, \begin{pmatrix}-1\\3\end{pmatrix}, \begin{pmatrix}2\\-4\end{pmatrix}, \begin{pmatrix}-2\\6\end{pmatrix}, \begin{pmatrix}4\\-8\end{pmatrix}, \begin{pmatrix}-4\\12\end{pmatrix}, \begin{pmatrix}8\\-16\end{pmatrix},\right.$$

$$\left.\begin{pmatrix}8\\-24\end{pmatrix}, \begin{pmatrix}-16\\32\end{pmatrix}, \begin{pmatrix}16\\-48\end{pmatrix}, \begin{pmatrix}-32\\-64\end{pmatrix}\right\}.$$

По построению данные последовательности удовлетворяют равенству (4.17) при $t=1, 2, \ldots, 10$, записанному применительно к рассматриваемой задаче оптимального управления.

4. Принцип оптимальности. Разобьем весь промежуток времени от 0 до N, где $N \geqslant 2$, на два периода $T_1=\{0, 1, \ldots, k\}$ и $T_2=\{k+1, \ldots, N\}$, где k — некоторое натуральное число вида $1 \leqslant k \leqslant N-1$. Предположим, что в результате применения допустимых управлений $u^{(1)}, u^{(2)}, \ldots, u^{(k)}$ дискретная система (4.14) из начального состояния $x^{(0)}$ перейдет в некоторое состояние $x^{(k)}$. Это состояние можно рассматривать как начальное состояние для второго периода управления T_2. При этом первому периоду соответствует критерий оптимальности $I_1=\sum\limits_{t=1}^{k} f^t(x^{(t-1)}, u^{(t)})$, а второму — $I_2=\sum\limits_{t=k+1}^{N} f^t(x^{(t-1)}, u^{(t)})$. Установим связь между оптимальной последовательностью управлений и подпоследовательностью управлений, отвечающей второму периоду T_2.

Обозначим через $\bar{u}^{(1)}, \bar{u}^{(2)}, \ldots, \bar{u}^{(N)}$ оптимальную последовательность управлений и через $\bar{x}^{(0)}, \bar{x}^{(1)}, \ldots, \bar{x}^{(N)}$ — соответствующую оптимальную траекторию. Имеет место следующее утверждение.

Принцип оптимальности. *Оптимальная последовательность управлений $\bar{u}^{(1)}, \bar{u}^{(2)}, \ldots, \bar{u}^{(N)}$ для системы* (4.14) *с критерием оптимальности* (4.15) *обладает тем свойством, что для произвольного $k \in \{1, 2, \ldots, N-1\}$ подпоследовательность управлений $\bar{u}^{(k+1)}, \ldots, \bar{u}^{(N)}$ является оптимальной последовательностью во втором периоде времени относительно критерия оптимальности I_2 и состояния, в котором оказалась система после первого периода.*

Для доказательства предположим противное: найдется такой номер $k \in \{1, 2, \ldots, N-1\}$ и подпоследовательность допустимых управлений $u^{(k+1)}, \ldots, u^{(N)}$ такая, что

$$\sum_{t=k+1}^{N} f^t(x^{(t-1)}, u^{(t)}) < \sum_{t=k+1}^{N} f^t(\bar{x}^{(t+1)}, \bar{u}^{(t)}), \tag{4.23}$$

где через $x^{(k)}, x^{(k+1)}, \ldots, x^{(N-1)}$ обозначена последовательность состояний, отвечающая $u^{(k+1)}, \ldots, u^{(N)}$, т. е. $x^{(k)}=\bar{x}^{(k)}$, $x^{(t)}=g^{(t)}(x^{(-1)}, u^{(t)})$, $t=k+1, \ldots, N$. На основании последних равенств последовательность управлений $\bar{u}^{(1)}, \bar{u}^{(2)}, \ldots, \bar{u}^{(k)}, u^{(k+1)}, \ldots, u^{(N)}$ является допустимой в исходной задаче и ей отвечает траектория $\bar{x}^{(0)}, \ldots, \bar{x}^{(k+1)}, x^{(k)}, \ldots, x^{(N-1)}$, причем в силу (4.23) имеет место неравенство

$$\sum_{t=1}^{k} f^t(\bar{x}^{(t-1)}, \bar{u}^{(t)}) + \sum_{t=k+1}^{N} f^t(x^{(t-1)}, u^{(t)}) < \sum_{t=1}^{N} f^t(\bar{x}^{(t-1)}, \bar{u}^{(t)}),$$

которое противоречит оптимальности последовательности управлений $\bar{u}^{(1)}, \bar{u}^{(2)}, \ldots, \bar{u}^{(N)}$. Полученное противоречие завершает доказательство.

Принцип оптимальности — это необходимое условие оптимальности. Согласно этому условию, подпоследовательность управлений, не являющаяся оптимальной в каком-либо втором периоде, не может входить ни в одну оптимальную последовательность относительно всего временно́го промежутка от 0 до N.

Для критерия оптимальности I, отличного от (4.15), принцип оптимальности может нарушаться. В этом можно убедиться на следующем простом примере. Пусть $n=1$, $r=1$, $N=3$, $x^{(0)}=1$, $x^{(t)}=x^{(t-1)}+u^{(t)}$, $t=1,\ 2,\ 3$; $u^{(t)}\in U=[-1,\ 1]$, $t=1,\ 2,\ 3$, и критерий оптимальности имеет вид

$$I=\max_{t=1,2,3}|x^{(t-1)}|.$$

Ясно, что $I\geqslant 1$ и минимальное значение критерия, равное 1, достигается, например, для последовательности управлений $\bar{u}^{(1)}=-1$, $\bar{u}^{(2)}=-1$, $\bar{u}^{(3)}=1$. Однако подпоследовательность $\bar{u}^{(2)}$, $\bar{u}^{(3)}$ не является оптимальной относительно критерия $\max\limits_{t=2,3}|x^{(t-1)}|$ и начального состояния $\bar{x}^{(1)}=0$. Этим требованиям удовлетворяет последовательность управлений $\tilde{u}^{(2)}=0$, $\tilde{u}^{(3)}=0$.

§ 4.3. ВЫЧИСЛИТЕЛЬНЫЕ АСПЕКТЫ ДИНАМИЧЕСКОГО ПРОГРАММИРОВАНИЯ

1. Формальное описание численного метода. В предыдущем параграфе было показано, что метод динамического программирования для решения задачи оптимального управления состоит в пошаговом построении функций $\omega_{N-1}(x)$, ..., $\omega_0(x)$ согласно равенствам (4.16). Найдя последнюю функцию ω_0 и вычислив ее значение в начальном состоянии, определяют искомое минимальное значение критерия оптимальности. В процессе вычислений на каждом шаге t для каждого состояния необходимо запоминать соответствующее оптимальное управление (точнее, «условно-оптимальное» управление). На основе этой информации «восстанавливается» оптимальная последовательность управлений и оптимальная траектория. Пример подобных вычислений приведен в п. 3 § 4.2.

Однако на практике, как правило, аналитически решить задачи оптимального управления не удается и приходится использовать численную процедуру, простейшая версия которой приводится ниже.

Выделим в пространстве состояний x, т. е. в $\mathbf{R}^n$, прямоугольное множество вида

$$X=\{x\in\mathbf{R}^n|a_i\leqslant x_i\leqslant b_i,\quad i=1,2,\ldots,n\},$$

в котором содержатся все возможные состояния дискретной системы, т. е. такое множество, что выполняется условие $x^{(t)}\in X$, $t=0,\ 1,\ \ldots,\ N$, для всех допустимых последовательностей управлений. Отметим, что точное определение границ a_i и b_i часто сопряжено с трудностями. Каждый отрезок $[a_i,\ b_i]$ разделим на части, а именно выберем число $\Delta x_i>0$, характеризующее точность вычислений, и положим

$$\hat{x}_i^{(k)}=a_i+(k-1)\Delta x_i,\ k=1,2,\ldots,A_i,\ A_i=\left[\frac{b_i-a_i}{\Delta x_i}\right]+1,$$

где [ ] — символ целой части числа. Тем самым в пространстве состояний образуется *сетка* (рис. 4.3), узлами которой являются точки вида $(\hat{x}_1^{(k_1)}, \hat{x}_2^{(k_2)}, \ldots, \hat{x}_n^{(k_n)})^T$, где $k_1 \leqslant A_1$, $k_2 \leqslant A_2$, ..., $k_n \leqslant A_n$. Эту сетку можно рассматривать в различные моменты времени $t=0, 1, \ldots, N$. Тогда в пространстве (x, t) будут заданы сетки, которые обозначим соответственно через $C_0, C_1, \ldots, C_N$ (рис. 4.4).

Аналогично выбираем некоторую сетку в пространстве управлений $\mathbf{R}^r$ для каждого $t=1, 2, \ldots, N$ и через U_t^0 обозначаем множество узлов этой сетки, содержащихся в множе-

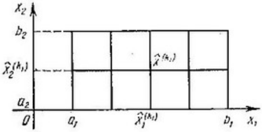

Рис. 4.3

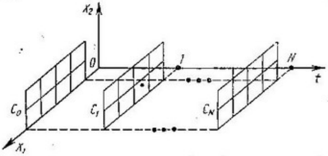

Рис. 4.4

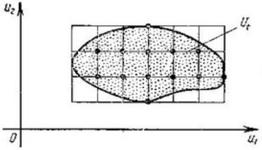

Рис. 4.5

стве допустимых управлений U_t (рис. 4.5; множество U_t^0 отмечено кружками).

Теперь опишем численный метод. Сначала прямым перебором вычисляют и заносят в память ЭВМ значения функции

$$\omega_{N-1}(x) = \min_{u \in U_N^0} f^N(x, u)$$

для всех x, являющихся узлами сетки C_{N-1}. Условимся далее узлы сеток в пространстве состояний обозначать через $\hat{x}^{(k)}$, а узлы сеток в пространстве управлений — через $\hat{u}^{(l)}$ и т. д., опуская индекс t. Для каждого узла $\hat{x}^{(k)}$ следует запоминать также управление $\hat{u}^{(l)} = \hat{u}^N(\hat{x}^{(k)}) \in U_N^0$, реализующее минимум в правой части равенства. В результате в памяти ЭВМ будут храниться все значения функции $\omega_{N-1}(x)$ и управлений $u^{(N)}(x)$, соответствующие узлам сетки C_{N-1}. На следующем шаге вычисляют и запоминают значения функции

$$\omega_{N-2}(x) = \min_{u \in U_{N-1}^0} [f^{N-1}(x, u) + \omega_{N-1}(g^{(N-1)}(x, u))]$$

в узлах сетки C_{N-2}. Для каждого узла следует запомнить и управление $\hat{u}^{(N-1)}(x)$ на котором достигается минимум в правой части равенства. Поскольку функция ω_{N-1} задана в виде таблицы, могут возникнуть трудности при вычислении значений функции ω_{N-2}. Дело в том, что состояние $g^{(N-1)}(\hat{x}^{(k)}, \hat{u}^{(l)})$ при фиксированных $\hat{x}^{(k)}$ и $\hat{u}^{(l)}$ (где $t=N-2$ для $\hat{x}^{(k)}$ и $t=N-1$ для $\hat{u}^{(l)}$), находясь внутри сетки C_{N-1}, может не совпадать ни с одним из ее узлов и поэтому из таблицы для ω_{N-1} нельзя определить значение $\omega_{N-1}(g^{(N-1)}(\hat{x}^{(k)}, \hat{u}^{(l)}))$. В таких случаях состояние $g^{(N-1)}(\hat{x}^{(k)}, \hat{u}^{(l)})$ можно отождествить с ближайшим узлом сетки C_{N-1} и требуемое значение определить из имеющейся таблицы для функции ω_{N-1}. При этом точность вычислений уменьшается. После занесения в память ЭВМ таблиц значений функции $\omega_{N-2}(x)$ и соответствующих управлений $u^{(N-1)}(x)$ таблица значений для ω_{N-1} может быть «забыта».

Дальнейшие шаги аналогичны. Процесс вычислений продолжают до тех пор, пока не будет получена таблица функции ω_0 в узлах сетки C_0. Минимальное значение критерия оптимальности равно значению функции ω_0, вычисленному в $x^{(0)}$ (или в узле, ближайшем к $x^{(0)}$). Приближенную оптимальную последовательность управлений и соответствующую траекторию по хранящимся в памяти ЭВМ таблицам функций $u^{(1)}(x), u^{(2)}(x), \ldots, u^{(N)}(x)$ находят в следующем порядке:

$$
\begin{array}{ll}
\bar{u}^{(1)} = u^{(1)}(x^{(0)}), & \bar{x}^{(1)} = g^{(1)}(x^{(0)}, \bar{u}^{(1)}), \\
\bar{u}^{(2)} = u^{(2)}(\bar{x}^{(1)}), & \bar{x}^{(2)} = g^{(2)}(\bar{x}^{(1)}, \bar{u}^{(2)}), \\
\ldots\ldots\ldots & \ldots\ldots\ldots\ldots\ldots \\
\bar{u}^{(N)} = u^{(N)}(\bar{x}^{(N-1)}), & \bar{x}^{(N)} = g^{(N)}(\bar{x}^{(N-1)}, \bar{u}^{(N)}).
\end{array}
$$

2. Анализ численного метода. Отметим достоинства описанного метода.

1) Метод не предполагает использования каких-либо аналитических свойств всех функций, участвующих в постановке задачи оптимального управления, причем множества U_t могут иметь произвольную структуру.

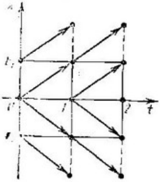

Рис. 4.6

2) Метод позволяет построить приближение к глобальному минимуму критерия оптимальности.

3) Таблица последней функции ω_0 получается не только для одного заданного состояния $x^{(0)}$, а для всех узлов сетки C_0. Это позволяет проанализировать «чувствительность» решения задачи к изменению начальных данных.

4) Метод можно использовать при решении задач с дополнительными фазовыми ограничениями $x^{(t)} \in X_t$, $t=1, 2, \ldots, N$. Для этого следует производить вычисления не для всех узлов сеток C_t, а лишь для множества «допустимых» узлов Y_t. Множество Y_t, $1 \leqslant t < N$, получается из множества узлов сетки C_t удалением узлов, не входящих в X_t, а также узлов $\hat{x}^{(k)}$, для которых состояние $g^{(t+1)}(\hat{x}^{(k)}, u)$ не принадлежит множеству X_{t+1} ни для каких $u \in U_{t+1}$; Y_N — это совокупность узлов сетки C_N, входящих в X_N. Оказывается, что при добавлении фазовых ограничений эффектив-

ность метода лишь возрастает, поскольку уменьшается количество переборов при отыскании минимума на каждом шаге.

Основной недостаток численного метода динамического программирования заключается в непомерно большом требовании к объему памяти ЭВМ. Уже для $n=4$; 5 и $A_i=100$ эти требования превышают возможности современных ЭВМ. Приходится уменьшать число узлов, выбирая «крупные» ячейки сеток (т. е. увеличивая Δx_i) или занижая границы a_i и b_i множества X. Однако в этом случае снижается точность вычислений, а при заниженных значениях a_i и b_i процесс вычислений не всегда удается довести до конца (рис. 4.6).

§ 4.4. ПРИКЛАДНЫЕ ЗАДАЧИ

1. Задача о кратчайшем пути. Пусть имеется некоторая система, которая может находиться в одном из конечного числа состояний. Переход из одного состояния в другое осуществляется по определенному правилу за определенное время. Требуется из заданного начального состояния перевести систему в желаемое состояние за минимальное время.

Для наглядности будем интерпретировать эту задачу как задачу нахождения кратчайшего пути в сети. Предварительно напомним понятия, связанные с сетью. *Ориентированная сеть* состоит из непустого конечного *множества вершин* V и подмножества X множества $V\times V$: $X\subset V\times V$. Элементы множества X представляют собой упорядоченные пары вершин и называются *дугами* сети. Занумеруем вершины сети числами натурального ряда 1, 2, ..., N. Наличие в множестве X упорядоченной пары (i, j) означает, что из вершины с номером i исходит дуга, которая входит в вершину с номером j. Каждой дуге (i, j) поставлено в соответствие некоторое неотрицательное число t_{ij}, которое будем интерпретировать как *длину данной дуги*. *Путем* называется конечная последовательность вершин, обозначаемая $(i_1, i_2, ..., i_n)$ и такая, что из вершины i_k исходит дуга, которая входит в вершину i_{k+1}, $k=1, 2, ..., n-1$. *Длиной пути* называется сумма длин входящих в него дуг. Путь, в котором начальная и конечная вершины совпадают, т. е. $i_1=i_n$, $n\geqslant 2$, называется *циклом*. Сеть, не содержащая циклов, называется *ациклической*.

Вершины ориентированной ациклической цепи всегда можно занумеровать так, чтобы для каждой дуги (i, j) было справедливо неравенство $i<j$. Поэтому далее будем считать, что такая нумерация произведена. Рассмотрим ациклическую сеть, имеющую 10 вершин, которая изображена на рис. 4.7. Вершины изображены в виде кружков, а дуги — стрелками. Возле каждой стрелки указана длина данной дуги. Внимательно просматривая данную сеть, можно выделить кратчайший путь из вершины *1* в конечную вершину *10*. Однако если бы сеть содержала достаточно большое чис-

ло вершин, то, используя метод просмотра, справиться с задачей было бы не просто.

Построим на данном простом примере алгоритм решения задачи, основанный на идеях динамического программирования и пригодной для сетей с большим числом вершин.

Начнем искать оптимальный путь «с конца». Из вершин *9* и *8* движение в вершину *10* определено однозначно. Присвоим указанным вершинам числа, соответствующие длинам дуг, т. е. 13 и 18. Из вершины *7* исходят две дуги: в вершину *10* и в вершину *9*. Поскольку $t_{7,10}=t_{7,9}+t_{9,10}=20$, присваиваем вершине *7* число 20. Из вершины *6* исходят три дуги, причем оптимальным движением из вершины *6* является перемещение по дуге *(6, 10)*, длина которой равна *15*. Приписываем это число вершине *6*. Из вершины *5* можно попасть либо в вершину *6* либо в вершину *8*. Сравниваем длины: $12+18<16+15$; следовательно, вершине *5* следует приписать число $12+18=30$. Поступая аналогично, придем в вершину *1*, которой будет приписано число 32 — длина искомого кратчайшего пути. Нетрудно восстановить и сам кратчайший путь: *1, 2, 4, 6, 10*.

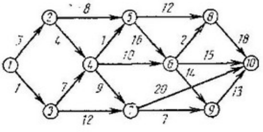

Рис. 4.7

Теперь сформулируем в общем виде *алгоритм нахождения кратчайшего пути*. Будем считать, что требуется найти кратчайший путь из вершины *1* в вершину N.

Шаг 1. Положить $\omega_N=0$ и $i=N-1$, где N — число вершин данной сети.

Шаг 2. Положить $\omega_i=\min(t_{ij}+\omega_j)$, где минимум вычисляется для всех $j>i$, для которых существует дуга (i, j). Запомнить путь, на котором реализуется указанный минимум. Если минимум достигается сразу на нескольких путях, то можно запомнить любой из них.

Шаг 3. Если $i=1$, то вычисления закончены. В противном случае уменьшить i на единицу и вернуться к шагу 2.

Вершины сети должны быть занумерованы таким образом, что номер вершины, из которой дуга исходит, должен быть меньше номера вершины, в которую она входит. Это предполагается выполненным для каждой дуги.

2. Задача об оптимальном распределении времени пребывания реагентов в каскаде реакторов идеального смешения [7]. В химических реакторах в результате химических превращений из исходного сырья получают необходимые продукты. Реактор идеального смешения представляет собой идеализированную модель реактора (сосуда), снабженного перемешивающим устройством, причем перемешивание происходит настолько интенсивно, что обеспечивается равномерность состава и температуры смеси в объеме реактора.

Пусть в реакторе происходит превращение реагента A в реагент B. Концентрации реагента A на входе в идеальный реактор $x_{вх}$ и на выходе из него $x_{вых}$ связаны [7] зависимостью

$$x_{вых} = x_{вх}/(1 + k\tau), \tag{4.24}$$

где τ — время пребывания реагента A в реакторе, k — константа скорости химической реакции (зависящая от температуры). При этом скорость реакции прямо пропорциональна концентрации реагента.

Соединим выход первого реактора с входом второго реактора, а выход второго — с входом третьего реактора и т. д. В результате получим каскад реакторов (многоступенчатый реактор). Обозначим через N число реакторов в каскаде. Заданы концентрация $x^{(0)}$ реагента A на входе каскада и концентрация $x^{(N)}$ на выходе последнего реактора. Требуется определить минимальное время пребывания реагента в каскаде и распределение этого времени по реакторам, если температура во всех реакторах одинакова.

Обозначим через $x^{(t)}$ концентрацию реагента на выходе реактора с номером t. Используя зависимость (4.24), можно записать

$$x^{(t)} = x^{(t-1)}/(1 + k_t\tau_t), \quad t = 1, 2, \ldots, N,$$

где τ_t — время пребывания реагента A в реакторе с номером t. Температура во всех реакторах считается одинаковой, поэтому можно принять $k_t = k$, $t = 1, 2, \ldots, N$, и вводя «обобщенное время» $u^{(t)} = k\tau_t$, записать

$$x^{(t)} = x^{(t-1)}/(1 + u^{(t)}), \quad t = 1, 2, \ldots, N. \tag{4.25}$$

Эти соотношения определяют дискретную управляемую систему с начальным состоянием $x^{(0)}$ и фиксированным конечным состоянием $x^{(N)}$. На управление наложено ограничение $u^{(t)} > 0$, $t = 1, 2, \ldots, N$. Критерий оптимальности, согласно условиям задачи, имеет вид

$$I = \sum_{t=1}^{N} u^{(t)}$$

и явно от $x^{(t)}$ не зависит. Его значение с точностью до множителя k совпадает с общим временем пребывания реагента в каскаде реакторов. Заметим, что в данной задаче параметр t, который в предыдущих пунктах играл роль времени означает номер реактора каскада.

Из (4.25) при $t = N$ находим $u^{(N)} = (x^{(N-1)}/x^{(N)}) - 1$. Следовательно, величина $u^{(N)}$ однозначно определяется заданным значением $x^{(N)}$ и величиной $x^{(N-1)}$. Поэтому

$$\omega_{N-1}(x^{(N-1)}) = \min_{u^{(N)} > 0} u^{(N)} = (x^{(N-1)}/x^{(N)}) - 1.$$

Далее, используя равенство (4.25) при $t=N-1$, имеем

$$\omega_{N-2}^{*}(x^{(N-2)})=\min_{u^{(N-1)}>0}\left(u^{(N-1)}+\frac{x^{(N-1)}}{x^{(N)}}-1\right)=$$

$$=\min_{u^{(N-1)}>0}\left(u^{(N-1)}+\frac{x^{(N-2)}}{x^{(N)}(1+u^{(N-1)})}-1\right).$$

Продифференцируем функцию, стоящую под знаком минимума, по $u^{(N-1)}$ и приравняем производную нулю. Имеем

$$1-\frac{x^{(N-2)}}{x^{(N)}\,(1+u^{(N-1)})^2}=0.$$

Отсюда находим единственный положительный корень $u^{(N-1)}=$ $=(x^{(N-2)}/x^{(N)})-1$. Если повторно продифференцировать функцию под знаком минимума, то полученное выражение будет положительным для любого $u^{(N-1)}$. Значит, найденный корень доставляет минимум. Таким образом, окончательно получаем

$$\omega_{N-2}(x^{(N-2)})=2\,[(x^{(N-2)}/x^{(N)})^{1/2}-1].$$

Аналогично находим:

$$\omega_{N-3}(x^{(N-3)})=\min_{u^{(N-2)}>0}\{u^{(N-2)}+2\,[(x^{(N-2)}/x^{(N)})^{1/2}-1]\},$$

$$u^{(N-2)}=(x^{(N-3)}/x^{(N)})^{1/3}-1,\quad \omega_{N-3}(x^{(N-3)})=3\,[(x^{(N-3)}/x^{(N)})^{1/3}-1].$$

В общем случае имеем

$$u^{(N-k)}=(x^{(N-k-1)}/x^{(N)})^{\frac{1}{k+1}}-1,\quad \omega_{N-k-1}(x^{(N-k-1)})=$$

$$=(k+1)\,[(x^{(N-k-1)}/x^{(N)})^{\frac{1}{k+1}}-1]. \tag{4.26}$$

В частности, при $k=N-1$

$$u^{(1)}=(x^{(0)}/x^{(N)})^{1/N}-1,\quad \omega_0(x^{(0)})=N\,[(x^{(0)}/x^{(N)})^{1/N}-1]. \tag{4.27}$$

Здесь $\omega_0(x^{(0)})$ — оптимальное значение «обобщенного времени».

Восстановим последовательность оптимальных управлений. Для этого в (4.25) при $t=1$ подставим $u^{(1)}$ из формулы (4.27). Имеем

$$x^{(1)}=[(x^{(0)})^{N-1}\,x^{(N)}]^{1/N}.$$

Используя найденное выражение, из (4.26) при $k=N-2$ получаем

$$u^{(2)}=(x^{(1)}/x^{(N)})^{\frac{1}{N-1}}-1=(x^{(0)}/x^{(N)})^{\frac{1}{N}}-1,$$

т. е. время пребывания реагента A во втором реакторе совпадает с временем пребывания в первом реакторе.

Аналогично приходим к равенствам

$$u^{(1)}=u^{(2)}=\ldots=u^{(N)}=(x^{(0)}/x^{(N)})^{1/N}-1.$$

Следовательно, оптимальным является распределение времени, одинаковое для всех реакторов каскада, а минимальное время пребывания реагента в каскаде равно $N\,(x^{(0)}/x^{(N)})^{1/N}-N$.

3. Задача регулирования скорости истечения жидкости. В верхнюю часть резервуара, имеющего форму кругового цилиндра, подается жидкость, а из нижней ее части жидкость вытекает. Заданы начальный уровень жидкости h и требуемый расход жидкости Q_t^* в данные моменты времени $t=1, 2, \ldots, N$. Регулируя скорость поступления жидкости, нужно добиться наименьшего отклонения расхода жидкости Q_t от заданных значений Q_t^*, $t=1, 2, \ldots, N$.

Отклонение набора величин $(Q_1, Q_2, \ldots, Q_N)$ от заданного набора $(Q_1^*, Q_2^*, \ldots, Q_N^*)$ будем измерять как расстояние между точками N-мерного евклидова пространства. Тогда минимизация отклонения сведется к минимизации выражения $\sqrt{\sum\limits_{t=1}^{N}(Q_t-Q_t^*)^2}$, что равносильно минимизации функции

$$I=\sum_{t=1}^{N}(Q_t-Q_t^*)^2. \tag{4.28}$$

Обозначим теперь через q_t скорость поступления жидкости в емкость, а через h_t — уровень жидкости в момент времени t. Запишем соотношение материального баланса в момент времени t:

$$Q_t=q_t-S\,(h_t-h_{t-1}),\quad t=1, 2,\ldots, N, \tag{4.29}$$

где S (м^2) — площадь поперечного сечения емкости. Будем считать, что скорость истечения жидкости пропорциональна $\sqrt{h_t}$:

$$Q_t=k\sqrt{h_t},\quad t=1, 2,\ldots, N. \tag{4.30}$$

Переменные q_t играют роль управлений. Согласно равенствам (4.30), состояние могут описывать как переменные Q_t, так и h_t. Для определенности выберем в качестве переменных состояния Q_t. Выразив из (4.30) величину h_t через Q_t и подставив найденное выражение в (4.29), получим квадратное уравнение относительно Q_t. Найдем единственный положительный корень этого уравнения:

$$Q_t=\frac{k^2}{2S}\left(\sqrt{1+\frac{4S^2Q_{t-1}^2}{k^4}+\frac{4Sq_t}{k^2}}-1\right),\quad t=1, 2,\ldots, N. \tag{4.31}$$

Теперь сформулируем задачу оптимального управления. Имеется дискретная управляемая система (4.31) с начальным состоянием $Q_0=k\sqrt{h}$. Требуется найти последовательность управлений $q_1, q_2, \ldots, q_N$ и соответствующую последовательность состояний

$Q_1, Q_2, \ldots, Q_N$, доставляющие минимальное значение критерию оптимальности (4.28).

Кроме естественных ограничений на выбор управлений $q_t > 0$, $t=1, 2, \ldots, N$, в сформулированной задаче могут быть и дополнительные ограничения, например вида

$$h_* \leqslant h_t \leqslant h^*, \quad q_* \leqslant q_t \leqslant q^*, \quad Q_* \leqslant Q_t \leqslant Q^*,$$

где $t=1, 2, \ldots, N$ и $h_*, h^*, q_*, q^*, Q_*, Q^*$ — заданные числа.

Решить полученную задачу оптимального управления при конкретных значениях входящих в нее параметров можно с помощью процедуры, описанной в § 4.3.

В заключение отметим, что если скорость истечения жидкости регулируется насосом, то равенство (4.30) можно заменить равенством $Q_t = Kh_t$ и вместо (4.31) управляемая система будет описываться более простыми соотношениями вида

$$Q_t = \frac{K}{K+S}\left(q_t + \frac{SQ_{t-1}}{K}\right), \quad t=1, 2, \ldots, N.$$

Глава 5
НЕЛИНЕЙНОЕ ПРОГРАММИРОВАНИЕ

В этой главе рассматриваются численные методы решения задач минимизации нелинейной функции как при ограничениях, так и при их отсутствии.

Задачи и соответствующие методы рассматриваются в данной главе по схеме «от простого — к сложному». Сначала излагаются наиболее распространенные методы минимизации функции одной переменной. Далее разобраны методы, пригодные для решения задач без ограничений с целевой функцией нескольких переменных. Наиболее сложные задачи оптимизации с ограничениями и методы их решения рассмотрены в § 5.3.

Универсального метода (алгоритма), с помощью которого можно было бы успешно решать разнообразные задачи оптимизации, не существует. Поэтому для решения каждого конкретного класса задач используют (и, как правило, не один) «свой» численный метод. Следует помнить о том, что эффективность численного решения зависит от того, насколько полно и точно отражается в применяемом методе специфика данной задачи, ее «индивидуальные» особенности.

Изложение в основном сводится к описанию и обсуждению методов. С доказательствами приводимых утверждений можно ознакомиться в [8, 17, 19, 21].

Заметим, что в этой главе рассмотрены наиболее простые численные методы, разработанные к настоящему времени.

§ 5.1. МИНИМИЗАЦИЯ ФУНКЦИИ ОДНОЙ ПЕРЕМЕННОЙ

1. Классический метод минимизации функции одной переменной. Рассмотрим задачу нахождения минимума функции f одной переменной на отрезке $[a, b]$. Классический метод подробно излагается в соответствующем разделе курса математического анализа. Напомним его содержание. Предполагается, что непрерывная функция f имеет непрерывную производную на всем отрезке $[a, b]$, за исключением, быть может, конечного числа точек. Согласно классическому методу, вычисляют производную $f'(x)$ и определяют критические точки, т. е. такие внутренние точки отрезка $[a, b]$, в которых производная обращается в нуль или не существует. Далее, в окрестности каждой критической точки исследуют знак производной и отбирают те из них, при переходе через которые производная меняет знак с $-$ на $+$ (это точки локального минимума). Наконец, в каждой из выделенных точек (включая концы отрезка $[a, b]$) вычисляют значения целевой функции. Сравнивая найденные значения, определяют минимальное. Точка, соответствующая этому минимальному значению, является точкой глобального минимума функции f на отрезке $[a, b]$.

Если функция f имеет непрерывную производную в каждой внутренней точке отрезка $[a, b]$, то процедура упрощается: находят все точки, в которых производная равна нулю, и, вычислив значения функции в найденных точках (включая концы отрезка), выделяют точку глобального минимума.

Основным недостатком классического метода является узкая область его применимости. Так, если значения целевой функции определены из наблюдений или в результате проведения экспериментов, то получить аналитическое выражение для ее производной трудно. Но даже если производная найдена, то отыскание корней уравнения $f'(x)=0$ может составить сложную вычислительную задачу, для решения которой потребуется неоправданно много времени и средств.

Ниже изложены методы минимизации, применение которых не требует знания производных и в которых, кроме того, объем вычислений значений целевой функции в определенном смысле является наименьшим.

2. Унимодальные функции. Выделим класс функций, обладающих с вычислительной точки зрения важным свойством. А именно: функция f называется *унимодальной* на отрезке $[a, b]$, если она имеет на этом отрезке единственную точку глобального минимума $x_{\min}$ и слева от этой точки является строго убывающей, а справа — строго возрастающей. Другими словами, функция f унимодальна, если точка $x_{\min}$ существует и единственна, причем для любых двух точек $x_1, x_2 \in [a, b]$ таких, что $x_1 < x_2$ из неравенства $x_1 > x_{\min}$ всегда следует $f(x_1) < f(x_2)$, а из неравенства $x_2 < x_{\min}$ необходимо вытекает неравенство $f(x_1) > f(x_2)$.

На рис. 5.1 приведен пример графика унимодальной функции.

Рассмотрим две произвольные точки x_1, $x_2 \in [a, b]$ такие, что $a < x_1 < x_2 < b$. Имеет место следующее важное свойство унимодальной функции: *если* $f(x_1) \leqslant f(x_2)$, *то* $x_{\min} < x_2$; *если же* $f(x_1) \geqslant f(x_2)$, *то* $x_{\min} > x_1$.

В самом деле, пусть выполняется неравенство $f(x_1) \leqslant f(x_2)$. Предположим противное: $x_{\min} \geqslant x_2$. Если $x_{\min} > x_2$, по определению унимодальной функции получим $f(x_1) > f(x_2)$, что невозможно. Если же $x_{\min} = x_2$, то, поскольку $f(x_1) \leqslant f(x_2)$, имеем две различные точки глобального минимума x_1 и x_2. Это также противоречит определению унимодальной функции. В случае $f(x_1) \geqslant f(x_2)$ доказательство аналогично.

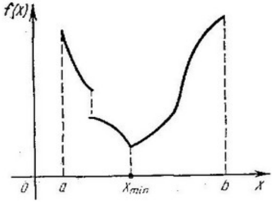

Рис. 5.1

Рассмотренное свойство унимодальной функции иллюстрирует рис. 5.2.

3. Метод «золотого сечения». Будем искать точку глобального минимума унимодальной функции

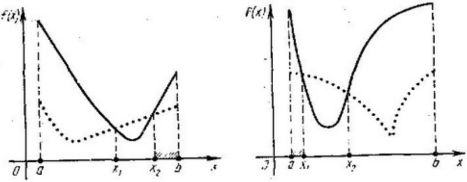

Рис. 5.2

f на отрезке $[a, b]$ так, чтобы количество вычислений значений целевой функции, необходимое для обеспечения заданной точности, было по возможности меньшим.

Рассмотрим на исходном отрезке точку x_1 и вычислим $f(x_1)$. Зная значение целевой функции в одной точке, невозможно сузить область поиска точки $x_{\min}$. Поэтому выберем вторую точку x_2 так, чтобы $a < x_1 < x_2 < b$, и вычислим $f(x_2)$. Возможен один из следующих двух случаев: $f(x_1) \leqslant f(x_2)$ или $f(x_1) > f(x_2)$. Согласно свойству унимодальной функции, установленному в п. 2, в первом случае искомая точка $x_{\min}$ не может быть на отрезке $[x_2, b]$, а во втором — на отрезке $[a, x_1]$ (на рис. 5.2 эти отрезки отмечены штриховкой). Следовательно, теперь область поиска сужается и следующую точку x_3 следует брать в одном из укороченных отрезков $[a, x_2]$ или $[x_1, b]$.

Установим, где на исходном отрезке лучше всего выбирать точки x_1 и x_2. Так как первоначально ничего не известно о положении точки $x_{\min}$, то оба указанных выше случая равновозможны, т. е. «лишним» может оказаться любой из отрезков $[x_2, b]$ и $[a, x_1]$. Отсюда ясно, что точки x_1 и x_2 должны быть расположены симметрично относительно середины отрезка $[a, b]$. Далее, для того чтобы максимально сузить область поиска, эти точки должны быть «поближе» к середине исходного отрезка. Однако слишком близкими их брать не следует, поскольку мы хотим построить алгоритм, для реализации которого необходимо минимальное количество вычислений значений функции. Это имеет место тогда, когда на втором этапе сужения области поиска потребуется вычислить лишь о д н о значение $f(x_3)$, которое будем сравнивать с уже имеющимся значением $f(x_1)$ или $f(x_2)$ в зависимости от того, какой из двух случаев реализовался. Поэтому если точки x_1 и x_2 взять рядом с серединой исходного отрезка, то на втором этапе сужение области поиска будет незначительным (рис. 5.3). Таким образом, с од-

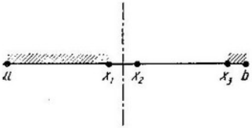

Рис. 5.3

Рис. 5.4

ной стороны, точки x_1 и x_2 следует выбирать рядом с серединой отрезка, а с другой — слишком близкими их брать нельзя. Для того чтобы найти «золотую середину», рассмотрим для простоты вместо $[a, b]$ отрезок $[0, 1]$ единичной длины (рис. 5.4).

Для того чтобы точка B была «выгодной» как на данном, так и на следующем этапе, она должна делить отрезок AD в таком же отношении, как и AC: $AB/AD = BC/AC$. При этом в силу симметрии аналогичным свойством будет обладать и точка C. В терминах координаты x записанная пропорция принимает вид

$$\frac{x}{1} = \frac{1-2x}{1-x}.$$

Уравнение имеет один корень, меньший 1, и равный $3-\sqrt{5})/2 \approx 0{,}382$. О точке, которая расположена на расстоянии $3-\sqrt{5})/2$% длины от одного из концов отрезка, говорят, что она осуществляет *золотое сечение* данного отрезка. Очевидно, каждый отрезок имеет две такие точки, расположенные симметрично относительно середины.

Итак, точки x_1 и x_2 должны осуществлять золотое сечение начального отрезка $[a, b]$. Отбрасывая ту часть $[a, b]$, в которой $x_{\min}$

заведомо быть не может, к укороченному отрезку (обозначим его через $[a_1, b_1]$) применяем аналогичные рассуждения с той лишь разницей, что здесь уже имеется одна внутренняя точка, осуществляющая золотое сечение. Поэтому к ней симметрично добавляем точку x_3 и сравниваем значения целевой функции в указанных двух точках. В результате приходим к еще более короткому отрезку (который обозначаем через $[a_2, b_2]$), содержащему $x_{\min}$. Аналогично строим последовательность точек $x_1, x_2, \ldots$, сходящуюся к $x_{\min}$. При этом последовательность отрезков $[a, b]$, $[a_1, b_1]$, $[a_2, b_2]$, ..., стягивается в точку $x_{\min}$. Рис. 5.5 иллюстрирует построение четырех членов последовательности, соответствующей неравенствам $f(x_1) \geqslant f(x_2)$, $f(x_2) \leqslant f(x_3)$, $f(x_4) \geqslant f(x_2)$.

Рис. 5.5

Приведем строгое описание алгоритма метода «золотого сечения». Для удобства введем следующие обозначения: $a_0=a$, $b_0=b$.

Ш а г 0. Вычисляем координаты точек, осуществляющих золотое сечение исходного отрезка: $y_0=a_0+\dfrac{3-\sqrt{5}}{2}(b_0-a_0)$, $z_0=a_0+b_0-y_0$. Кроме того, вычисляем значения $f(y_0)$ и $f(z_0)$.

Ш а г k $(k\geqslant 1)$. В результате предыдущего шага известны величины y_{k-1}, z_{k-1}, a_{k-1}, b_{k-1}, $f(y_{k-1})$, $f(z_{k-1})$. Сравниваем значения $f(y_{k-1})$ и $f(z_{k-1})$. Если $f(y_{k-1})\leqslant f(z_{k-1})$, то полагаем $a_k=a_{k-1}$, $b_k=z_{k-1}$, $z_k=y_{k-1}$ и, используя эти числа, вычисляем координату симметричной точки (слева от имеющейся) $y_k=a_k+b_k-z_k$ и значение $f(y_k)$.

Если же выполняется противоположное неравенство $f(y_{k-1})>f(z_{k-1})$, то следует принять $a_k=y_{k-1}$, $b_k=b_{k-1}$, $y_k=z_{k-1}$ и вычислить координату симметричной точки (справа от имеющейся) $z_k=a_k+b_k-y_k$ вместе со значением $f(z_k)$.

Независимо от того, какой из этих двух случаев реализуется, нужно вычислить длину следующего $(k+1)$-го отрезка, т. е. величину $\Delta_{k+1}=b_k-y_k=z_k-a_k$.

Шаги алгоритма осуществляют до тех пор, пока не будет выполнено неравенство $\Delta_{k+1}\leqslant\varepsilon$, где $\varepsilon>0$ — заданная точность вычислений. Затем выбирают наименьшее из чисел $f(y_k)$ и $f(z_k)$ — приближенный минимум. А точка, которая ему соответствует, дает приближение к $x_{\min}$. При этом отклонение точки приближенного минимума от точки $x_{\min}$ действительного минимума не превышает ε.

В результате будет известна максимальная возможная величина отклонения ε точки приближенного минимума от точки $x_{\min}$, а об отклонении значения приближенного минимума от $f(x_{\min})$ по-прежнему ничего не известно.

Установим, что отрезки $[a_k, b_k]$, $k=1, 2, \ldots$, построенные с помощью метода золотого сечения, действительно стягиваются в точку $x_{\min}$. На k-м шаге значение длины отрезка $[a_k,\ b_k]$ составляет

$$b_k - a_k = \frac{\sqrt{5}-1}{2}(b_{k-1} - a_{k-1}),\ k = 1,\ 2, \ldots .$$

Поэтому

$$b_k - a_k = \left(\frac{\sqrt{5}-}{2}\right)^k (b_0 - a_0) < 0{,}7^k (b_0 - a_0) \underset{k\to\infty}{\to} 0.$$

Следовательно, отрезок $[a_k,\ b_k]$, в котором находится точка $x_{\min}$, при неограниченном увеличении k стягивается в точку, принадлежащую одновременно всем отрезкам $[a_k,\ b_k]$, $k = 1, 2, \ldots$. Функция f унимодальная, поэтому такой точкой может быть лишь $x_{\min}$.

Еще раз отметим, в методе золотого сечения на нулевом шаге вычисляют два значения целевой функции, а на каждом последующем шаге — лишь одно.

4. Метод Фибоначчи. На практике количество вычислений значений целевой функции часто бывает ограничено некоторым числом n (тем самым ограничено и число шагов вычислений по методу золотого сечения; оно не превышает $n-1$). Метод Фибоначчи отличается от метода золотого сечения лишь выбором первых двух симметричных точек и гарантирует более точное приближение к точке $x_{\min}$ за $n-1$ шаг, чем метод золотого сечения за то же количество шагов.

Согласно методу Фибоначчи, на нулевом шаге координаты первых двух симметричных точек вычисляют по формулам

$$y_0 = a_0 + \frac{F_n}{F_{n+2}}(b_0 - a_0),$$

$$z_0 = a_0 + b_0 - y_0,$$

где через F_{n+2} обозначено $(n+2)$-е число Фибоначчи, определяемое рекуррентной формулой

$$F_{n+2} = F_n + F_{n+1},\ n = 1,\ 2,\ 3, \ldots;\ F_1 = F_2 = 1.$$

Запишем первые десять чисел Фибоначчи: $F_1 = F_2 = 1$, $F_3 = 2$, $F_4 = 3$, $F_5 = 5$, $F_6 = 8$, $F_7 = 13$, $F_8 = 21$, $F_9 = 34$, $F_{10} = 55$.

Дальнейшие вычисления по методу Фибоначчи совпадают с соответствующими шагами метода золотого сечения. Отличие состоит в том, что нахождение величин Δ_k становится лишним, так как при $k = n-1$ процесс вычислений заканчивают и y_{n-1} принимают за приближенное значение $x_{\min}$.

Рассмотрим подробнее ситуацию, возникающую при $k = n-1$. Используя математическую индукцию и определение чисел Фибоначчи, можно доказать, что

$$\Delta_k = b_k - a_k = \frac{F_{n-k+2}}{F_{n+2}}(b_0 - a_0), \quad k = 1,\ 2, \ldots, n-1. \tag{5.1}$$

С помощью этих формул для симметричных точек y_k и z_k запишем следующие выражения:

$$y_k=a_k+\Delta_{k+2}=a_k+\frac{F_{n-k}}{F_{n+2}}(b_0-a_0),$$

$$z_k=a_k+\Delta_{k+1}=a_k+\frac{F_{n-k+1}}{F_{n+2}}(b_0-a_0).$$

Отсюда видно, что при $k=n-1$ две симметричные точки сольются в одну (так как $y_{n-1}=z_{n-1}$) и разделят отрезок $[a_{n-1}, b_{n-1}]$ на две равные части. Согласно формуле (5.1), длина этого отрезка равна $\Delta_{n-1}=(2/F_{n+2})(b_0-a_0)$. Следовательно, принимая y_{n-1} за приближенное значение $x_{\min}$, имеем следующую оценку отклонения этого значения от истинного:

$$|y_{n-1}-x_{\min}|\leqslant\frac{b_0-a_0}{F_{n+2}}. \tag{5.2}$$

Таким образом, погрешность вычислений по методу Фибоначчи при фиксированном n не превышает значения правой части неравенства (5.2).

На практике заранее может быть задано не число n вычислений значений целевой функции, а погрешность $\varepsilon>0$. В этом случае, для того чтобы определить необходимое для обеспечения заданной точности число n, можно воспользоваться неравенством (5.2). В самом деле, если $(b_0-a_0)/F_{n+2}\leqslant\varepsilon$, то требуемая точность будет достигнута. Отсюда вытекают условия для определения n:

$$F_{n+1}<\frac{b_0-a_0}{\varepsilon}\leqslant F_{n+2}.$$

Можно доказать [8], что $\lim\limits_{n\to\infty}F_n/F_{n+2}=(3-\sqrt{5})/2$, поэтому при достаточно больших n вычисления согласно методу Фибоначчи и методу золотого сечения начинают практически из одной и той же пары симметричных точек.

В обосновании методов золотого сечения и Фибоначчи важную роль играет свойство унимодальности минимизируемой функции. Это не означает, что рассмотренные методы неприменимы к функциям, которые не обладают этим свойством. Унимодальность, в частности, влечет совпадение локального и глобального минимумов. А оба рассмотренных метода носят «локальный характер». Поэтому если целевая функция не является унимодальной, то численная реализация обоих методов приведет, вообще говоря, лишь в окрестность точки локального минимума. При этом само значение локального минимума может оказаться весьма далеким от значения глобального минимума (рис. 5.6).

5. Метод равномерного перебора. В идейном отношении (и при реализации на ЭВМ) наиболее простым методом отыскания глобального минимума является метод равномерного перебора. Согласно этому методу, фиксируют величину шага $h>0$, вычисляют значения целевой функции в точках $x_1=a$ (или в точке, «близкой» к a справа), $x_2=x_1+h$ и полученные значения сравнивают. Запоминают меньшее из этих двух значений. Далее вычисляют значение функции f при $x_3=x_2+h$ и сравнивают его с тем значением, которое хранится в памяти. Снова запоминают меньшее значение. Таким образом последовательно перебирают значения f в точках $x_k=x_{k-1}+h$, $k=4, 5, \ldots$, до тех пор, пока при некотором $k=n$ очередная точка x_{n+1} «не покинет» отрезок $[a, b]$. Значение целевой функции, которое останется в памяти после остановки (это $\min\limits_{i=1,2,\ldots,n} f(x_i)$), полагают приближенным значением глобального минимума.

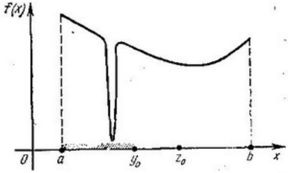

Рис. 5.6

При практической реализации такого подхода основная проблема заключается в установлении величины шага h. Простые примеры (рис. 5.6) убеждают, что даже при сравнительно малом h существует возможность «проскочить» глобальный минимум. В общем случае невозможно решить вопрос о том, насколько малым следует выбирать h, чтобы значение приближенного минимума отличалось от истинного не более чем на $\varepsilon>0$:

$$\left| \min_{i=1,2,\ldots,n} f(x_i) - \min_{x\in[a,b]} f(x) \right| \leqslant \varepsilon. \tag{5.3}$$

Однако существует достаточно широкий класс функций, для которых можно решить эту проблему.

Говорят, что *функция f удовлетворяет условию Липшица* с константой $L\geqslant 0$, если

$$| f(x)-f(x') | \leqslant L\, | x-x' | \text{ для всех } x, x' \in [a, b]. \tag{5.4}$$

Геометрически неравенство (5.4) означает, что тангенс угла наклона любой секущей (прямой, проходящей через две точки графика) к положительному направлению оси абсцисс не может превышать L или же может быть меньше чем $-L$.

Выполнение неравенства (5.4) влечет непрерывность f на $[a, b]$. Значит, функция, удовлетворяющая условию Липшица, достигает на $[a, b]$ своего наименьшего (и наибольшего) значения (при этом может быть несколько локальных минимумов и максимумов).

Пусть функция f удовлетворяет условию Липшица с константой L и $x_1=a+h/2$, $x_k=x_{k-1}+h$, $k=2, 3, \ldots, n-1$, $x_n=\min\{x_{n-1}+h;\ b\}$, где n подбирается так, чтобы было выполнено условие $h/2<b-x_{n-1}\leqslant 3h/2$. Тогда выбор шага по формуле $h=2\varepsilon/L$ гарантирует выполнение неравенства (5.3).

□ Заметим, что система отрезков $[x_1-h/2,\ x_1+h/2]$, $[x_2-h/2,\ x_2+h/2]$, ..., $[x_n-h/2,\ x_n+h/2]$ покрывает весь исходный отрезок

$[a, b]$. Поэтому точка глобального минимума $x_{\min}$ принадлежит одному из них, т. е. $x_m - h/2 \leqslant x_{\min} \leqslant x_m + h/2$ при некотором $m \in \{1, 2, ..., n\}$. Используя неравенство (5.4) для точек x_m и $x_{\min}$, имеем

$$|f(x_m) - f(x_{\min})| \leqslant L\, |x_m - x_{\min}| \leqslant \frac{Lh}{2} = \varepsilon.$$

Отсюда, так как $f(x_m) \geqslant \min\limits_{i=1,2,...,n} f(x_i)$, вытекает требуемое неравенство (5.3). ■

На практике константа L, как правило, не известна. Поэтому вычисляют угловые коэффициенты нескольких секущих графика целевой функции и в качестве L берут значение, наибольшее по абсолютной величине. Следует отметить, что если взять слишком большое значение L, то это приведет к слишком малой величине шага h, а «заниженное» значение может привести к нарушению неравенства (5.3).

§ 5.2. ЧИСЛЕННЫЕ МЕТОДЫ В ЗАДАЧАХ БЕЗ ОГРАНИЧЕНИЙ

1. Общая схема методов спуска. Будем рассматривать задачу безусловной минимизации, т. е. задачу минимизации целевой функции f на всем пространстве $\mathbf{R}^n$. Сущность всех методов приближенного решения этой задачи, излагаемых в данном параграфе, состоит в построении последовательности точек $x^{(0)}, x^{(1)}, x^{(2)}, ..., x^{(k)}, ...,$ монотонно уменьшающих значение целевой функции:

$$f(x^{(0)}) \geqslant f(x^{(1)}) \geqslant f(x^{(2)}) \geqslant ... \geqslant f(x^{(k)}) \geqslant ... . \qquad (5.5)$$

Такие методы (алгоритмы) называют *методами спуска*. При использовании этих методов применяют следующую схему. Пусть на k-й итерации имеется точка $x^{(k)}$. Тогда выбирают направление спуска $p^{(k)} \in \mathbf{R}^n$ и длину шага вдоль этого направления $\alpha_k > 0$. Следующую точку последовательности вычисляют по формуле

$$x^{(k+1)} = x^{(k)} + \alpha_k p^{(k)}, \quad k = 0, 1, 2, ...$$

Согласно этой формуле, величина продвижения из точки $x^{(k)}$ в $x^{(k+1)}$ зависит как от α_k, так и от $p^{(k)}$. Однако α_k традиционно называют длиной шага.

Формально различные методы спуска отличаются друг от друга способом выбора числа α_k и вектора $p^{(k)}$. Если для определения α_k и $p^{(k)}$ требуется вычислять только значения целевой функции, соответствующие методы называют *методами нулевого порядка* или *методами поиска*. *Методы первого порядка* требуют, кроме того, вычисления первых производных целевой функции. Если же метод предполагает использование и вторых производных, то его называют *методом второго порядка* и т. д.

С помощью методов нулевого порядка можно решать задачи более широкого класса, чем с помощью методов первого или второго порядка. Однако методы нулевого порядка, как правило, требуют бóльших вычислений для достижения заданной точности, поскольку использование только значений целевой функции не позволяет достаточно точно определить направление на точку минимума.

Важнейшей характеристикой любых методов спуска является их *сходимость*. Сходимость здесь понимается в том смысле, что последовательность $\{x^{(k)}\}$ должна сходиться к точке глобального (локального) минимума. Однако точки минимума могут составлять целое множество и многие алгоритмы позволяют построить последовательность $\{x^{(k)}\}$, которая сама не является сходящейся, но любая ее сходящаяся подпоследовательность имеет в качестве предельной некоторую точку минимума (рис. 5.7). В этом случае говорят, что каждая предельная точка последовательности $\{x^{(k)}\}$ является точкой минимума. С помощью подобных алгоритмов можно строить последовательности точек, сколь угодно близко приближающихся ко множеству точек минимума.

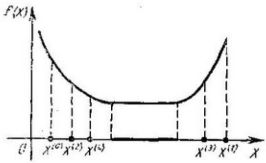

Рис. 5.7

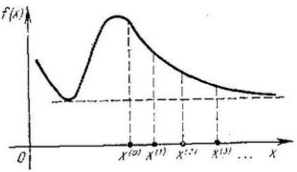

Рис. 5.8

Возможен случай, когда ничего определенного сказать и о сходимости последовательностей нельзя, однако известно, что соответствующая последовательность значений функции $\{f(x^{(k)})\}$ сходится к минимальному значению (локальному или глобальному минимуму). Тогда говорят, что последовательность $\{x^{(k)}\}$ *сходится* к минимуму *по функции* (рис. 5.8). Кроме того, существуют еще более слабые типы сходимости, когда, например, последовательность $\{x^{(k)}\}$ (каждая ее подпоследовательность) имеет в качестве предельной стационарную точку (т. е. точку, в которой градиент равен нулевому вектору), являющуюся лишь «подозрительной» на оптимальную.

Как правило, тип сходимости одного и того же метода зависит от конкретного вида целевой функции, т. е. в разных задачах метод может сходиться по-разному. При достаточно жестких требованиях к функции f с помощью метода можно строить последовательность, сходящуюся в точке глобального минимума. Если же этот метод применить к функциям, не удовлетворяющим этим требованиям, то может быть получена последовательность, сходящаяся только по функции, либо последовательность, не являющаяся сходящейся ни в каком смысле.

Методы спуска в силу условия монотонности (5.5) обычно не приводят к точке локального (глобального) максимума.

Отметим, что даже в тех случаях, когда нет сходимости ни в одном смысле, последовательное уменьшение значения целевой функции может представлять практический интерес.

2. Метод покоординатного спуска. Согласно этому методу, направление спуска выбирают параллельным координатным осям. Сначала производят спуск вдоль первой оси Ox_1, затем — вдоль второй — Ox_2 и т. д. до последней оси Ox_n.

Обозначим i-й орт пространства $\mathbf{R}^n$ через $e^{(i)}$, т. е. вектор, у которого все координаты нулевые, кроме i-й, равной единице. Пусть $x^{(0)}$ — начальная точка и α_0 — некоторое положительное число. Точку $x^{(1)}$ определяют следующим образом. Вычисляют значение функции $f(x)$ при $x=x^{(0)}+\alpha_0 e^{(1)}$ и проверяют выполнение неравенства

$$f(x^{(0)}+\alpha_0 e^{(1)})<f(x^{(0)}). \tag{5.6}$$

Если это неравенство справедливо, то вдоль направления оси Ox_1 значение функции f уменьшилось и поэтому полагают

$$x^{(1)}=x^{(0)}+\alpha_0 e^{(1)},\ \ \alpha_1=\alpha_0.$$

Если (5.6) не имеет места, то делают шаг в противоположном направлении, т. е. проверяют неравенство

$$f(x^{(0)}-\alpha_0 e^{(1)})<f(x^{(0)}). \tag{5.7}$$

В случае выполнения этого неравенства полагают

$$x^{(1)}=x^{(0)}-\alpha_0 e^{(1)},\ \ \alpha_1=\alpha_0.$$

Возможно, что оба неравенства (5.6) и (5.7) окажутся невыполненными. Тогда следует считать $x^{(1)}=x^{(0)}$, $\alpha_1=\alpha_0$.

Второй шаг производят вдоль координатной оси Ox_2: если $f(x^{(1)}+\alpha_1 e^{(2)})<f(x^{(1)})$, то полагают $x^{(2)}=x^{(1)}+\alpha_1 e^{(2)}$, $\alpha_2=\alpha_1$, если же последнее неравенство не имеет места, то проверяют неравенство $f(x^{(1)}-\alpha_1 e^{(2)})<f(x^{(1)})$ и в случае его выполнения считают $x^{(2)}=x^{(1)}-\alpha_1 e^{(2)}$, $\alpha_2=\alpha_1$, Если ни одно из двух неравенств не выполняется, то полагают $x^{(2)}=x^{(1)}$, $\alpha_2=\alpha_1$. Так перебирают все n направлений координатных осей. На этом первая итерация закончена; на n-м шаге будет получена некоторая точка $x^{(n)}$. Если при этом $x^{(n)}\neq x^{(0)}$, то, аналогично, начиная с $x^{(n)}$ осуществляют вторую итерацию. Если же $x^{(n)}=x^{(0)}$ (это имеет место в том случае, когда на каждом шаге ни одно из пары проверяемых неравенств не окажется выполненным), то величину шага следует уменьшить, взяв, например, $\alpha_{n+1}=\alpha_n/2$, и в следующей итерации использовать новое значение величины шага.

Последующие итерации производят аналогично. На практике вычисления продолжают до тех пор, пока не выполнится некоторое условие окончания счета. Часто используют следующие условия:

$$\| x^{(k+1)} - x^{(k)} \| \leqslant \delta \text{ или } | f(x^{(k+1)}) - f(x^{(k)}) | \leqslant \varepsilon, \tag{5.8}$$

где δ и ε — некоторые положительные числа, характеризующие точность решения исходной задачи минимизации.

На рис. 5.9 изображены несколько линий уровня некоторой функции двух переменных и проиллюстрирован описанный метод.

Сходимость метода покоординатного спуска устанавливает следующее утверждение.

Пусть функция f определена, выпукла и непрерывно дифференцируема на $\mathbb{R}^n$, причем начальная точка $x^{(0)}$ выбрана так, что множество $\{x \in \mathbb{R}^n \mid f(x) \leqslant f(x^{(0)})\}$ ограничено. Тогда каждая предельная точка последовательности, построенной по методу покоординатного спуска, является точкой глобального минимума.

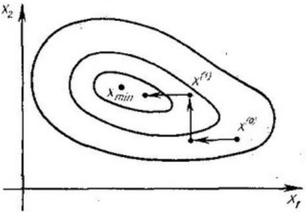

Рис. 5.9

Заметим, что сходимость гарантируется, если начальная точка $x^{(0)}$ выбрана надлежащим образом.

Рассматриваемый метод относится к классу методов нулевого порядка и для его реализации не требуется вычислять производные. Однако в условиях сформулированного утверждения имеется требование непрерывной дифференцируемости f. Примеры показывают [8], что если метод покоординатного спуска применять к функциям, не удовлетворяющим этому требованию, то точка минимума может быть не получена.

Иногда, стремясь ускорить сходимость метода, величину α_k подбирают так, чтобы при переходе от $x^{(k)}$ к $x^{(k+1)}$ вдоль направления спуска обеспечивалось наибольшее возможное убывание целевой функции. Другими словами, α_k находят из условия минимума функции $f(x^{(k)} + \alpha e^{(i)})$ одной переменной α:

$$f(x^{(k)} + \alpha_k e^{(i)}) = \min_{-\infty < \alpha < +\infty} f(x^k + \alpha e^{(i)}),$$

где номер $i \in \{1, 2, \ldots, n\}$ определяется номером шага k. Такой выбор α_k, вообще говоря, приводит к тому, что для достижения заданной точности потребуется меньшее число шагов. Однако выполнение каждого шага будет сопряжено с решением задачи минимизации функции одной переменной, что приведет к дополнительным вычислениям. Кроме того, нахождение точного значения α_k в этой задаче не всегда возможно; если же вместо точного значения использовать приближенное, то может нарушаться условие убывания $f(x^{(k+1)}) \leqslant f(x^{(k)})$.

3. Градиентные методы. Как отмечалось в гл. 2, ненулевой антиградиент $-\nabla f(x^{(0)})$ указывает направление, небольшое перемещение вдоль которого из $x^{(0)}$ приводит к значению функции f меньшему, чем $f(x^{(0)})$. Это замечательное свойство антиградиента лежит в основе *градиентных методов*, согласно которым на k-й итерации полагают $p^{(k)}=-\nabla f(x^{(k)})$, т. е.

$$x^{(k+1)}=x^{(k)}-\alpha_k\nabla f(x^{(k)}),\ \alpha_k>0,\ k=0,\ 1,\ 2,\dots .$$

Эти методы отличаются друг от друга способом выбора величины шага α_k. Достаточно малый шаг α_k обеспечивает убывание целевой функции

$$f(x^{(k+1)})=f(x^{(k)}-\alpha_k\nabla f(x^{(k)}))<f(x^{(k)}), \tag{5.9}$$

но может привести к слишком большому количеству итераций для достижения требуемой точности. С другой стороны, выбор большого шага может привести к нарушению неравенства (5.9).

На практике нередко в качестве величины шага выбирают некоторое $\alpha>0$, одинаковое для всех итераций. При этом если условие (5.9) (при $\alpha_k=\alpha$) нарушится, то для текущей итерации величину α уменьшают до тех пор, пока указанное неравенство не станет выполненным.

Часто величину α_k рекомендуют выбрать так, чтобы имело место более жесткое условие убывания, чем (5.9):

$$f(x^{(k)})-f(x^{(k)}-\alpha_k\nabla f(x^{(k)}))\geqslant\varepsilon\alpha_k\,\|\nabla f(x^{(k)})\|, \tag{5.10}$$

где $0<\varepsilon<1$ — некоторая фиксированная константа. Здесь также сначала фиксируют некоторое $\alpha_k=\alpha>0$ (например, $\alpha_k=1$), одинаковое для всех итераций, а затем при необходимости уменьшают его до тех пор, пока не будет выполнено неравенство (5.10).

В *методе наискорейшего спуска* величину $\alpha_k>0$ определяют в результате минимизации функции $\varphi_k(\alpha)=f(x^{(k)}-\alpha\nabla f(x^{(k)}))$ одной переменной α:

$$\varphi_k(\alpha_k)=\min_{\alpha\geqslant 0}\varphi_k(\alpha). \tag{5.11}$$

Функция $\varphi_k(\alpha)$ в точке $\alpha_k>0$ достигает минимума, поэтому ее производная в этой точке равна нулю:

$$\left.\frac{d\varphi_k(\alpha)}{d\alpha}\right|_{\alpha=\alpha_k}=\frac{d}{d\alpha}f\left(x_1^{(k)}-\alpha\frac{\partial f(x^{(k)})}{\partial x_1},\dots,x_n^{(k)}-\alpha\frac{\partial f(x^{(k)})}{\partial x_n}\right)\Bigg|_{\alpha=\alpha_k}=$$

$$=-\sum_{i=1}^{n}\frac{\partial f(x^{(k+1)})}{\partial x_i}\frac{\partial f(x^{(k)})}{\partial x_i}=-\langle\nabla f(x^{(k+1)}),\ \nabla f(x^{(k)})\rangle=0.$$

Последнее равенство означает, что направление спуска на $(k+1)$-й итерации ортогонально направлению спуска на предыдущей k-й итерации. Таким образом, кривая движения по методу наискорейшего спуска представляет собой ломаную, соседние звенья которой взаимно ортогональны, причем звено, соединяющее $x^{(k)}$ и $x^{(k+1)}$, лежит в гиперплоскости, касательной к поверхности уровня $f(x)=f(x^{(k+1)})$ (рис. 5.10).

При реализации градиентных методов в качестве критериев окончания счета кроме (5.8) используют также условие вида

$$\|\nabla f(x^{(k)})\| \leqslant \gamma,$$

где $\gamma>0$ — фиксированная точность вычислений.

Точный смысл сходимости градиентных методов раскрывает следующее утверждение.

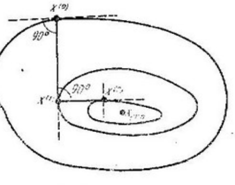

Рис. 5.10

Пусть функция f ограничена снизу, непрерывно дифференцируема на $\mathbf{R}^n$ и ее градиент $\nabla f(x)$ удовлетворяет условию Липшица:

$$\|\nabla f(x)-\nabla f(x')\| \leqslant L\|x-x'\| \quad \text{для всех } x,\ x'\in\mathbf{R}^n,$$

где $L\geqslant 0$ — некоторая фиксированная константа. Кроме того, пусть выбор величины шага α_k производится на основе условия (5.10) *или* (5.11). *Тогда, какова бы ни была начальная точка $x^{(0)}$, оба градиентных метода приводят к построению последовательности $\{x^{(k)}\}$, обладающей свойством $\lim\limits_{k\to\infty}\|\nabla f(x^{(k)})\|=0$.*

*Если, кроме того *), функция f дважды непрерывно дифференцируема и существуют такие числа $M\geqslant m>0$, что*

$$m\|y\|^2 \leqslant \langle \nabla^2 f(x)y,\ y\rangle \leqslant M\|y\|^2 \quad \text{для всех } x,\ y, \tag{5.12}$$

то для обоих указанных градиентных методов последовательность $\{x^{(k)}\}$ будет сходиться к точке глобального минимума.

Первая часть утверждения гарантирует сходимость лишь в смысле $\lim\limits_{k\to\infty}\|\nabla f(x^k)\|=0$, т. е. сходимость по функции либо к точной нижней грани $\inf\limits_{x\in\mathbf{R}^n} f(x)$, либо к значению функции f в некоторой стационарной точке x^*. При этом сама точка x^* не обязательно является точкой локального минимума; она может быть точкой седлового типа. Однако на практике подобная ситуация маловероятна и применение градиентных методов, как правило, позволяет получить приближенное значение минимума целевой функции (вообще говоря, локального).

Сравнивая рис. 5.10 и 5.11, можно сделать вывод, что для целевой функции, линии уровня которой близки к окружностям, требуемая точность будет достигнута довольно быстро, тогда как если

---

*) В этой части утверждения требование ограниченности снизу функции f на самом деле является лишним, так как если функция f удовлетворяет (5.12), то она заведомо ограничена снизу [21].

линии уровня сильно вытянуты в окрестности оптимальной точки, то градиентные методы приведут к медленному зигзагообразному продвижению в направлении на оптимальную точку. О функции, поверхности уровня которой сильно вытянуты, говорят, что она имеет «овражный» характер (в случае двух переменных график такой функции действительно напоминает овраг). О степени «овражности» функции f можно получить представление, зная минимальное (m) и максимальное (M) собственные числа матрицы Гессе $\nabla^2 f$, вычисленной в оптимальной точке: чем меньше отношение m/M, тем больше «овражность» данной функции. Применение градиентных методов к таким функциям приводит к спуску на «дно оврага», после чего, поскольку направление спуска почти перпендикулярно «линии дна», точки последовательности $\{x^{(k)}\}$ будут поочередно находиться то на одном «склоне оврага», то на другом и продвижение к оптимальной точке сильно замедляется.

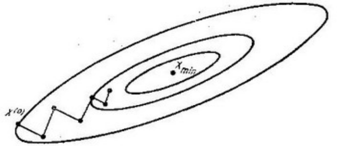

Рис. 5.11

Иногда при минимизации функции «овражного» типа удается использовать следующую процедуру. Выбирают две близкие начальные точки $\bar{x}^{(0)}$ и $\bar{x}^{(1)}$. Из этих точек совершают градиентный спуск на «дно оврага», в результате получают соответственно точки $x^{(0)}$ и $x^{(1)}$. Далее полагают

$$\bar{x}^{(2)} = x^{(1)} - h_1 \frac{x^{(1)} - x^{(0)}}{\| x^{(1)} - x^{(0)} \|} \operatorname{sign}(f(x^{(1)}) - f(x^{(0)})),$$

где $h_1 > 0$ — сравнительно большой (овражный) шаг. Затем из $\bar{x}^2$ вновь совершают градиентный спуск (он может состоять из нескольких итераций) на «дно оврага» и получают точку $x^{(2)}$. Следующую точку $\bar{x}^{(3)}$ находят по формуле, аналогичной приведенной выше:

$$\bar{x}^{(3)} = x^{(2)} - h_2 \frac{x^{(2)} - x^{(1)}}{\| x^{(2)} - x^{(1)} \|} \operatorname{sign}(f(x^{(2)}) - f(x^{(1)}))$$

и т. д. (рис. 5.12). Эффективность описанного метода зависит от величин овражных шагов $h_1, h_2, \ldots$. Если овражный шаг слишком велик, то можно довольно далеко отойти от линии «дна оврага»; если же этот шаг мал, то продвижение будет незначительным. Точных методов выбора овражного шага не существует; его величину подбирают эмпирически, учитывая известные свойства минимизируемой функции.

4. Метод Ньютона. Это метод второго порядка. Одно из преимуществ этого метода по сравнению с градиентными методами состоит в том, что он «не реагирует» на овражный характер минимизируемой функции. Метод Ньютона заключается в применении формулы

$$x^{(k+1)}=x^{(k)}-[\nabla^2 f(x^{(k)})]^{-1}\nabla f(x^{(k)}),\ k=0,\ 1,\dots, \tag{5.13}$$

где $[\nabla^2 f(x^{(k)})]^{-1}$ — матрица, обратная матрице Гессе (матрице вторых производных), вычисленной в точке $x^{(k)}$. Предполагается, что $\nabla^2 f(x^{(k)})$ — невырожденная матрица, поэтому обратная ей матрица существует.

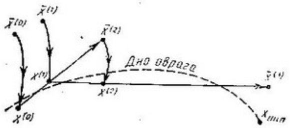

Рис. 5.12

Для того чтобы установить геометрический смысл формулы (5.13), будем считать, что $\nabla^2 f(x^{(k)})$ — положительно-определенная матрица, и рассмотрим квадратичную аппроксимацию функции f в окрестности точки $x^{(k)}$:

$$f_k(x)=f(x^{(k)})+\langle\nabla f(x^{(k)}),\ x-x^{(k)}\rangle+\frac{1}{2}\ \langle\nabla^2 f(x^{(k)})(x-x^{(k)}),$$

$$x-x^{(k)}\rangle.$$

Вычисляя градиент этой функции и приравнивая его нулю, получаем

$$\nabla f_k(x)=\nabla f(x^{(k)})+\nabla^2 f(x^{(k)})(x-x^{(k)})=0_n.$$

Отсюда, умножая равенство на матрицу $[\nabla^2 f(x^{(k)})]^{-1}$ слева, находим стационарную точку функции $f_k(x)$:

$$x=x^{(k)}-[\nabla^2 f(x^{(k)})]^{-1}\nabla f(x^{(k)}). \tag{5.14}$$

Матрица $\nabla^2 f(x^{(k)})$ положительно определена, поэтому точка x вида (5.14) является точкой глобального минимума квадратичной функции $f_k(x)$ (см. п. 2 § 2.5). Правые части равенств (5.13) и (5.14) идентичны. Следовательно, на каждой итерации метода Ньютона очередную точку $x^{(k+1)}$ выбирают такой, чтобы она доставляла глобальный минимум квадратичной аппроксимации $f_k(x)$ (рис. 5.13). Градиентные методы, по существу, используют линейную аппроксимацию целевой функции и поэтому менее точно определяют направление на точку минимума. Таким образом, метод Ньютона позволяет достигнуть заданной точности за меньшее число итераций, чем градиентные методы. Однако каждая

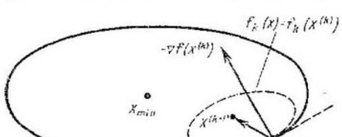

Рис. 5.13

итерация метода Ньютона связана с вычислением матрицы $\nabla^2 f(x^{(k)})$ и последующим ее обращением, что требует большего объема вычислений по сравнению с одной итерацией градиентного метода.

Направление движения из точки $x^{(k)}$ по методу Ньютона определяется вектором

$$x^{(k+1)} - x^{(k)} = -[\nabla^2 f(x^{(k)})]^{-1} \nabla f(x^{(k)}).$$

Если матрица $\nabla^2 f(x^{(k)})$ положительно определена, то и обратная матрица также положительно определена. * Поэтому

$$\langle \nabla f(x^{(k)}),\ x^{(k+1)} - x^{(k)} \rangle = -\langle [\nabla^2 f(x^{(k)})]^{-1} \nabla f(x^{(k)}),\ \nabla f(x^{(k)}) \rangle < 0,$$

а это, согласно лемме 2.1, означает, что небольшое перемещение вдоль направления $x^{(k+1)} - x^{(k)}$ должно привести к уменьшению значения целевой функции. Однако в методе Ньютона величина перемещения не регулируется; она однозначно определяется точками $x^{(k)}$ и $x^{(k+1)}$ (и равна $\|x^{(k+1)} - x^{(k)}\|$). Поэтому если начальная точка $x^{(0)}$ выбрана сравнительно «далеко» от точки минимума, то метод Ньютона может приводить к «большим» перемещениям, что повлечет нарушение условия убывания $f(x^{(k+1)}) < f(x^{(k)})$. Это положение иллюстрирует пример функции одной переменной, изображенной на рис. 5.14. Здесь вторая производная имеет в точке $x^{(0)}$ положительное значение, но оно слишком мало, так как в окрестности $x^{(0)}$ функция f близка к линейной. В результате значение величины, обратной второй производной, довольно большое и точка $x^{(1)}$ оказывается дальше от $x_{\min}$, чем $x^{(0)}$. По тем же причинам и $x^{(2)} - x_{\min} > x^{(0)} - x_{\min}$ и т. д. Таким образом, в данном случае использование метода Ньютона приводит к построению последовательности, удаляющейся от точки минимума.

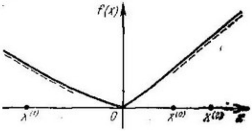

Рис. 5.14

Имеет место следующее утверждение.

Пусть функция f дважды непрерывно дифференцируема на $\mathbb{R}^n$, x^* — стационарная точка f и матрица Гессе $\nabla^2 f(x^*)$ невырожденная. Тогда существует окрестность точки x^* такая, что для любого начального приближения $x^{(0)}$ из этой окрестности, последовательность $\{x^{(k)}\}$, построенная согласно (5.13), *сходится к x^*.*

Таким образом, если начальное приближение $x^{(0)}$ выбрано достаточно близким к точке минимума, то последовательность, построенная по методу Ньютона, будет сходиться к этой точке. Однако, не зная положения точки минимума, такой выбор $x^{(0)}$ практически

* Так как матрица $\nabla^2 f(x^{(k)}) = A$ предполагается невырожденной, то любой вектор $y \neq 0_n$ можно представить в виде $y = Ax$, $x \in \mathbb{R}^n$. Положительная определенность матрицы A означает, что $\langle Ax, x \rangle > 0$ для всех $x \neq 0_n$. Отсюда при $x = A^{-1}y$ получаем $\langle AA^{-1}y, A^{-1}y \rangle = \langle y, A^{-1}y \rangle = \langle A^{-1} y, y \rangle > 0$ для любого $y \neq 0_n$, что означает положительную определенность обратной матрицы A^{-1}.

осуществить невозможно. Поэтому были разработаны модификации метода Ньютона, которые обеспечивают сходимость независимо от начального приближения. Это *методы Ньютона с регулировкой шага:*

$$x^{(k+1)}=x^{(k)}-\alpha_k[\nabla^2 f(x^{(k)})]^{-1}\nabla f(x^{(k)}),\ k=0,\ 1,\dots . \tag{5.15}$$

Величину α_k здесь выбирают следующим образом. Сначала (как в методе Ньютона) полагают $\alpha=1$ и проверяют условие убывания

$$f(x^{(k)}+\alpha p^{(k)})-f(x^{(k)})\leqslant \varepsilon\alpha\langle\nabla f(x^{(k)}),\ p^{(k)}\rangle, \tag{5.16}$$

где $p^{(k)}=-[\nabla^2 f(x^{(k)})]^{-1}\nabla f(x^{(k)})$ — направление спуска, а $0<\varepsilon<{}^1/_2$ — некоторое число. Если неравенство (5.16) выполнено, то полагают $\alpha_k=\alpha=1$ и производят следующую итерацию. В противном случае берут меньшее значение α (например, $\alpha={}^1/_2$) и снова проверяют выполнение неравенства (5.16). Таким образом, шаг делится до тех пор, пока не будет выполнено неравенство (5.16). Как только это произойдет, текущее значение α принимают за величину шага α_k.

Другой способ выбора α_k такой же, как и в методе наискорейшего спуска,— из условия минимизации функции в направлении движения, т. е. из условия

$$f(x^{(k)}-\alpha_k[\nabla^2 f(x^{(k)})]^{-1}\nabla f(x^{(k)}))=$$
$$=\min_{\alpha\geqslant 0}[f(x^{(k)}-\alpha[\nabla^2 f(x^{(k)})]^{-1}\nabla f(x^{(k)}))].$$

Пусть функция f дважды непрерывно дифференцируема и существуют числа $M\geqslant 0$ и $m>0$ такие, что выполняется условие (5.12). *Тогда последовательность* (5.15), *где α_k выбирают одним из двух указанных выше способов, сходится к точке глобального минимума независимо от выбора начального приближения.*

5. Метод сопряженных направлений. Рассмотрим сначала метод сопряженных направлений первого порядка для квадратичной целевой функции, а затем обобщим его на случай функции более общего вида.

Пусть целевая функция f является квадратичной, т. е.

$$f(x)=\frac{1}{2}\langle Qx,\ x\rangle+\langle c,\ x\rangle, \tag{5.17}$$

причем Q — симметричная неотрицательно-определенная матрица размера $n\times n$ и $c\in\mathbf{R}^n$.

Говорят, что линейно независимая система n векторов $\{p^{(k)}\}_{k=0}^{n-1}$ является *системой сопряженных направлений* (относительно матрицы Q), если $\langle Qp^{(i)},\ p^{(j)}\rangle=0$ для всех $i,\ j=0,\ 1,\ \dots,\ n-1,\ i\neq j$.

В частном случае, когда Q — единичная матрица, система сопряженных направлений превращается в систему попарно ортогональных векторов. В этом смысле первую систему можно рассматривать как определенное обобщение второй системы.

Зафиксируем некоторую систему сопряженных направлений $\{p^{(k)}\}_{k=0}^{n-1}$ (о ее построении сказано ниже). Вычисления согласно методу сопряженных направлений проводятся по обычной схеме

$$x^{(k+1)}=x^{(k)}+\alpha_k p^{(k)},\ k=0,\ 1,\ldots, \tag{5.18}$$

где величину шага α_k выбирают из условия минимума функции $\varphi_k(\alpha)=f(x^{(k)}+\alpha p^{(k)})$. Найдем конкретное выражение для α_k. В результате несложных преобразований получаем

$$\frac{d\varphi_k}{d\alpha}=\alpha\,\langle Qp^{(k)},\ p^{(k)}\rangle+\langle Qx^{(k)},\ p^{(k)}\rangle+\langle c,\ p^{(k)}\rangle. \tag{5.19}$$

Приравняем эту производную нулю и найдем искомое значение $\alpha=\alpha_k$. Имеем

$$\alpha_k=-\frac{\langle \nabla f(x^{(k)}),\ p^{(k)}\rangle}{\langle Qp^{(k)},\ p^{(k)}\rangle},\ k=0,\ 1,\ldots,n-1. \tag{5.20}$$

Здесь предполагается, что знаменатель дроби отличен от нуля. Если же $\langle Qp^{(k)},\ p^{(k)}\rangle=0$ для некоторого k и, кроме того, $\langle \nabla f(x^{(k)}),\ p^{(k)}\rangle\neq 0$, то из (5.19) следует, что производная $d\varphi_k/d\alpha$ отлична от нуля в каждой точке $\alpha\in\mathbf{R}$. Это означает, что целевая функция не может достигать минимума на луче $\{x\in\mathbf{R}^n \mid x=x^{(k)}+\alpha p^{(k)},\ \alpha\in\mathbf{R}\}$. А значит, она не достигает минимума и на $\mathbf{R}^n$, т. е. исходная задача решения не имеет.

Как следует из формулируемого ниже утверждения, метод сопряженных направлений при минимизации квадратичной функции является конечным. Этим он выгодно отличается от уже рассмотренных методов.

Пусть $\{p^{(k)}\}_{k=0}^{n-1}$ — некоторая система сопряженных направлений (относительно Q). Тогда с помощью формул (5.18), (5.20) *метода сопряженных направлений для произвольно выбранного начального приближения $x^{(0)}$ точка минимума x^* квадратичной функции* (5.17) *(если она существует) вычисляется не более чем за n шагов, т. е. $x^*=x^{(n-1)}$.*

Доказательство. Поскольку $\{p^{(k)}\}$ — линейно независимая система из n векторов, она является базисом в $\mathbf{R}^n$, а значит, вектор $x^*-x^{(0)}$ можно представить в виде линейной комбинации векторов этой системы. Таким образом, найдутся такие числа α_0^*, $\alpha_1^*,\ldots,\alpha_{n-1}^*$, что

$$x^*=x^{(0)}+\sum_{k=0}^{n-1}\alpha_k^* p^{(k)}. \tag{5.21}$$

Умножая это равенство на матрицу Q, а затем скалярно на вектор $p^{(i)}$, получаем равенство

$$\langle Qx^*,\ p^{(i)}\rangle=\langle Qx^{(0)},\ p^{(i)}\rangle+\sum_{k=0}^{n-1}\alpha_k^*\,\langle Qp^{(k)},\ p^{(i)}\rangle.$$

Отсюда, используя сопряженность системы направлений $\{p^{(i)}\}$, а также равенство $Qx^*=-c$ (см. п. 2 § 2.5), имеем

$$\alpha_i^* \langle Qp^{(i)}, p^{(i)}\rangle = -\langle p^{(i)}, Qx^{(0)}+c\rangle,$$

следовательно,

$$\alpha_i^* = -\frac{\langle \nabla f(x^{(0)}), p^{(i)}\rangle}{\langle Qp^{(i)}, p^{(i)}\rangle}, \quad i=0, 1,\dots, n-1. \tag{5.22}$$

Теперь рассмотрим соотношения (5.18). Имеем

$$x^{(n)}=x^{n-1}+\alpha_{n-1}p^{(n-1)}=x^{(n-2)}+\alpha_{n-2}p^{(n-2)}+\alpha_{n-1}p^{(n-1)}=\dots,$$
$$\dots=x^{(0)}+\sum_{k=0}^{n-1}\alpha_k p^{(k)}.$$

Сравнивая это равенство с (5.21), видим, что для завершения доказательства остается убедиться в том, что $\alpha_k=\alpha_k^*$, $k=0, 1, \dots, n-1$. Последние равенства вытекают из формул (5.20), (5.22) и равенств

$$\langle \nabla f(x^{(k)}), p^{(k)}\rangle = \langle Qx^{(k)}+c, p^{(k)}\rangle = \langle Qx^{(k-1)}+\alpha_{k-1}Qp^{(k-1)}+$$
$$+c, p^{(k)}\rangle = \langle Qx^{(k-1)}+c, p^{(k)}\rangle = \dots = \langle Qx^{(0)}+c, p^{(k)}\rangle =$$
$$=\langle \nabla f(x^{(0)}), p^{(k)}\rangle, \quad k=0, 1,\dots, n-1.$$

Укажем способ построения системы сопряженных направлений. Для этого используется итерационный процесс вида

$$p^{(0)}=-\nabla f(x^{(0)}), \tag{5.23}$$
$$p^{(k)}=-\nabla f(x^{(k)})+\beta_{k-1}p^{(k-1)}, \quad k=1, 2,\dots, n-1,$$

где

$$\beta_{k-1}=\frac{\langle \nabla f(x^{(k)}), \nabla f(x^{(k)})\rangle}{\langle \nabla f(x^{(k-1)}), \nabla f(x^{(k-1)})\rangle}, \quad k=1, 2,\dots, n-1. \tag{5.24}$$

Сопряженные направления (доказательство сопряженности системы направлений (5.23) см., например, в [17]) вычисляют последовательно одно за другим по мере того, как в этом возникает необходимость при вычислениях по формуле (5.18).

Если целевая функция нелинейная и не является квадратичной, метод сопряженных направлений уже не является конечным, хотя можно использовать те же формулы (5.18), (5.23) и (5.24), причем α_k также определять из условия минимума

$$\varphi_k(\alpha_k)=f(x^{(k)}+\alpha_k p^{(k)})=\min_{\alpha\geqslant 0}\varphi_k(\alpha).$$

Числа β_{k-1} можно выбирать в виде

$$\beta_{k-1}=\frac{\langle \nabla f(x^{(k)}), \nabla f(x^{(k)})-\nabla f(x^{(k-1)})\rangle}{\langle \nabla f(x^{(k-1)}), \nabla f(x^{(k-1)})\rangle}, \quad k=1, 2,\dots, n-1.$$

После нахождения точки $x^{(n)}$, если условие окончания счета $\|\nabla f(x^{(n)})\|<\varepsilon$ еще не выполнено, процесс можно повторить, взяв точку $x^{(n)}$ в качестве начальной. Если и следующие n шагов не приведут к требуемой точности, вычисления продолжают.

Следующее утверждение определяет характер сходимости метода сопряженных направлений в случае неквадратичных функций, градиент которых удовлетворяет условию Липшица.

Пусть функция f ограничена снизу и такова, что для некоторой константы $L>0$ неравенство $\|\nabla f(x')-\nabla f(x'')\|\leqslant L\|x-x'\|)$ выполняется при всех x', $x''\in \mathbf{R}^n$. Тогда для последовательности точек $\{x^{(k)}\}$, построенной указанным выше способом, справедливо равенство $\lim\limits_{k\to\infty}\|\nabla f(x^{(k)})\|=0$.

6. Замечания. На практике часто комбинируют описанные методы. Сначала обычно используют метод покоординатного спуска, поскольку он не требует для определения направления спуска сложных вычислений. Далее переходят к градиентным методам, так как направление антиградиента точнее указывает на точку минимума, чем направление координатных осей. Наконец, «попав в окрестность точки минимума», для получения более высокой точности используют метод Ньютона или его модификации.

§ 5.3. ЧИСЛЕННЫЕ МЕТОДЫ В ЗАДАЧАХ С ОГРАНИЧЕНИЯМИ

1. Метод покоординатного спуска. Изложенный в п. 2 § 5.2 метод покоординатного спуска можно распространить на задачи с ограничениями, если эти ограничения достаточно просты:

$$X=\{x\in \mathbf{R}^n \mid a_i\leqslant x_i\leqslant b_i,\ i=1,\ 2,\dots,\ n\}, \tag{5.25}$$

где a_i, b_i — заданные числа, $i=1,\ 2,\ \dots,\ n$.

Опишем этот метод. Пусть известны точка $x^{(k)}\in X$ (k-е приближение) и величина шага $\alpha_k>0$ при некотором $k\geqslant 0$. Направление спуска выбирается по формулам

$$p^{(k)}=\mathrm{e}^{(i_k)},\ i_k=k-n\left[\frac{k}{n}\right]+1,$$

где [ ] — символ целой части числа. Этот выбор аналогичен выбору направления спуска в п. 2 § 5.2 и состоит в циклическом переборе единичных ортов $\mathrm{e}^{(1)}$, $\mathrm{e}^{(2)}$, ..., $\mathrm{e}^{(n)}$ пространства R^n. Вычисляют новую точку $x^{(k)}+\alpha_k p^{(k)}$ и проверяют условия допустимости и убывания:

$$x^{(k)}+\alpha_k p^{(k)}\in X,\ f(x^{(k)}+\alpha_k p^{(k)})<f(x^{(k)}).$$

Если оба условия выполнены, то очередное приближение определяют по формуле

$$x^{(k+1)}=x^{(k)}+\alpha_k p^{(k)},\ \alpha_{k+1}=\alpha_k;$$

в противном случае пытаются сделать шаг в противоположном направлении, т. е. проверяют условия $x^{(k)}-\alpha_k p^{(k)}\in X$, $f(x^{(k)}-\alpha_k p^{(k)})<$

$<f(x^{(k)})$. В случае выполнения этих двух условий полагают $x^{(k+1)}=$ $=x^{(k)}-\alpha_k p^{(k)}$, $\alpha_{k+1}=\alpha_k$. Если же хотя бы одно из указанных условий не имеет места, то очередное приближение определяют по формулам

$$x^{(k+1)}=x^{(k)},\ \alpha_{k+1}=\begin{cases}\lambda\alpha_k & \text{при } i_k=n,\ x^{(k)}=x^{k-n+1},\\ \alpha_k & \text{при } i_k\neq n \text{ или } x^{(k)}\neq x^{(k-n+1)},\\ & \text{или } 0\leqslant k\leqslant n-1,\end{cases}$$

где λ — заданный параметр метода, $0<\lambda<1$ (например $\lambda=1/2$). Анализ последних формул показывает, что очередное приближение в задаче с ограничениями строится аналогично задачам без ограничений.

Если функция f выпукла на множестве X вида (5.25) *и непрерывно дифференцируема, то при любом выборе начальной точки $x^{(0)}\in X$ и начального значения $\alpha_0>0$ каждая предельная точка последовательности, построенной описанным методом, является точкой глобального условного минимума.*

2. Метод условного градиента. Этот метод находит применение при решении задач минимизации нелинейной функции на таком выпуклом компактном множестве X, на котором можно без особого труда решить задачу минимизации линейной функции.

Пусть выбрано начальное приближение $x^{(0)}\in X$. Через $x^{(k)}\in X$, как и раньше, обозначим k-е приближение, $k\geqslant 0$. Запишем линейную аппроксимацию функции f в точке $x^{(k)}$:

$$f_L(x)=f(x^{(k)})+\langle\nabla f(x^{(k)}),\ x-x^{(k)}\rangle.$$

Множество X предполагается компактным, поэтому линейная функция $f_L(x)$ на этом множестве достигает своего наименьшего значения. Обозначим через $\bar{x}^{(k)}$ такую произвольную точку минимума. Точку $\bar{x}^{(k)}$ обычно находят в результате решения более простой эквивалентной задачи

$$\langle\nabla f(x^{(k)}),\ x\rangle\to\min_{x\in X}.$$

Точку $\bar{x}^{(k)}$ используют для определения направления спуска: $p^{(k)}=$ $=\bar{x}^{(k)}-x^{(k)}$. Следующее приближение находят по формуле $x^{(k+1)}=$ $=x^{(k)}+\alpha_k p^{(k)}$, где $0<\alpha_k\leqslant 1$. В силу выпуклости множества X всегда выполняется условие допустимости $x^{(k+1)}\in X$.

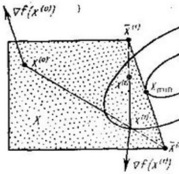

Рис. 5.15

На рис. 5.15 проиллюстрирован метод условного градиента.

Отметим, что если X задано конечной системой линейных неравенств и равенств, то вспомогательная задача нахождения точки $\bar{x}^{(k)}$ является задачей линейного программирования и для ее решения может быть использован симплекс-метод.

Часто точное решение во вспомогательной задаче найти не удается и приходится использовать приближенное решение. В этом случае точка $\bar{x}^{(k)}$ определяется с точностью до ε_k:

$$\langle \nabla f(x^{(k)}), \bar{x}^{(k)} \rangle \leqslant \min_{x \in X} \langle \nabla f(x^{(k)}), x \rangle + \varepsilon_k, \; \varepsilon_k \geqslant 0,$$

причем $\lim\limits_{k \to \infty} \varepsilon_k = 0$. Последнее равенство представляет собой требование с увеличением номера итерации k все более точно решать вспомогательную задачу.

При разных способах выбора длины шага α_k получаются различные варианты метода условного градиента. Рассмотрим некоторые из них.

1) Величину α_k выбирают из условия минимизации целевой функции на отрезке, соединяющем точки $x^{(k)}$ и $\bar{x}^{(k)}$, т. е. из условия

$$f_k(\alpha_k) = \min_{0 \leqslant \alpha \leqslant 1} f_k(\alpha),$$

где $f_k(\alpha) = f(x^{(k)} + \alpha \cdot (\bar{x}^{(k)} - x^{(k)}))$. Для решения этой задачи могут быть использованы методы минимизации функции одной переменной, изложенные в § 5.1. Как правило, точное решение здесь найти невозможно; следовательно, можно ограничиться приближенным значением величины α_k, удовлетворяющим следующим условиям:

$$f_k(\alpha_k) \leqslant \min_{0 \leqslant \alpha \leqslant 1} f_k(\alpha) + \delta_k, \; \delta_k \geqslant 0, \; \sum_{k=0}^{\infty} \delta_k < +\infty.$$

Здесь условие $\sum\limits_{k=0}^{\infty} \delta_k < +\infty$ означает, что δ_k должны стремиться к нулю при $k \to \infty$ «достаточно быстро». Например, $(k+1)^{-1} \to 0$ при $k \to \infty$, но для $\delta_k = (k+1)^{-1}$ получаем гармонический ряд, который, как известно, расходится: $\sum\limits_{k=0}^{\infty} (k+1)^{-1} = +\infty$.

2) Величину α_k можно выбирать и вне всякой связи с целевой функцией. Требуется лишь, чтобы было выполнено условие

$$\lim_{k \to \infty} \alpha_k = 0, \; \sum_{k=0}^{\infty} \alpha_k = +\infty, \; 0 < \alpha_k \leqslant 1.$$

Этим требованиям удовлетворяет, например, следующий выбор: $\alpha_k = (k+1)^{-1}$, $k = 0, 1, \ldots$. Достоинством данного способа выбора α_k является простота его реализации, однако метод условного градиента с таким выбором α_k не обязательно является методом спуска. Нетрудно привести пример, где требование монотонного убывания $f(x^{(k+1)}) < f(x^{(k)})$ для некоторых k нарушается.

Пусть $X \subset \mathbf{R}^n$ — выпуклое компактное множество, функция f выпукла, непрерывно дифференцируема и ее градиент удовлетворяет условию Липшица, т. е.

$$\|\nabla f(x) - \nabla f(x')\| \leqslant L \|x - x'\| \; \textit{для всех} \; x, \; x' \in X,$$

где $L \geqslant 0$ — некоторая константа. Тогда, какова бы ни была начальная точка $x^{(0)} \in X$, описанные выше варианты метода условного градиента приводят к построению последовательности $\{x^{(k)}\}$, каждая предельная точка которой является точкой глобального минимума функции f на множестве X.

3. Метод барьерных функций. Этот метод предназначен для решения задач нелинейного программирования с ограничениями в форме неравенств:

$$f(x) \to \min_{x \in X}; \quad X = \{x \in \mathbf{R}^n | g_j(x) \leqslant 0, \ j = 1, 2, \ldots, m\}. \tag{5.26}$$

Введем множества точек:

$$X_1 = \{x \in \mathbf{R}^n | g_j(x) < 0, \quad j = 1, 2, \ldots, m\},$$

$$X_2 = \{x \in \mathbf{R}^n | g_j(x) = 0 \text{ для некоторого } j\}.$$

Очевидно, $X_1 \cup X_2 = X$ и $X_1 \cap X_2 = \varnothing$. Будем предполагать, что $X_1 \neq \varnothing$. Это предположение запрещает присутствие в задании множества X ограничений-равенств.

На множестве X_1 определим функцию $B(x)$, которая называется *барьерной функцией*. Барьерные функции могут быть двух типов:

$$B(x) = -\sum_{j=1}^{m} (1/g_j(x)),$$

$$B(x) = \sum_{j=1}^{m} \max\{-\ln(-g_j(x)); \ 0\}.$$

Отметим следующие очевидные свойства барьерных функций:

1°. $B(x) \geqslant 0$ для всех $x \in X_1$.

2°. $\lim\limits_{k \to \infty} B(x^{(k)}) = +\infty$ для любой последовательности точек $\{x^{(k)}\}$ из X_1 и такой, что $\lim\limits_{k \to \infty} x^{(k)} \in X_2$.

Составим функцию:

$$F_k(x) = f(x) + r_k B(x), \quad k = 1, 2, \ldots, \tag{5.27}$$

где $\{r_k\}_{k=1}^{\infty}$ — монотонно убывающая, сходящаяся к нулю последовательность положительных чисел.

Лемма 5.1. *Пусть X — замкнутое ограниченное множество, у которого $X_1 \neq \varnothing$. Пусть также $f(x)$, $g_1(x)$, ..., $g_m(x)$ — непрерывные на X функции. Тогда задача $F_k(x) \to \min\limits_{x \in X_1}$ имеет решение при каждом $k = 1, 2, \ldots$,*

□ Возьмем произвольное натуральное k. Обозначим $F^* = \inf\limits_{x \in X_1} F_k(x) < +\infty$. По определению точной нижней грани существует минимизирующая последовательность $\{x^{(s)}\}$ такая, что $x^{(s)} \in X_1$ и $\lim\limits_{s \to \infty} F_k(x^{(s)}) = F^*$. В силу компактности множества X из этой последовательности можно выделить сходящуюся подпоследовательность. Без ограничения общности можно считать, что сходится сама эта последовательность: $\lim x^{(s)} = \bar{x} \in X$. Включение $\bar{x} \in X_2$ невозможно согласно свойству 2° барьерной функции. Следовательно, $\bar{x} \in X_1$. Наконец, используя непрерывность функции $F_k(x)$, получаем $\lim\limits_{s\to\infty} F_k(x^{(s)}) = F_k(\bar{x}) = F^*$. ■

Рассмотрим метод барьерных функций. Выберем последовательность $\{r_k\}$, обладающую указанными выше свойствами. Например, можно положить $r_k = k^{-1}$ или $r_k = 10^{-k}$, $k = 1, 2, \ldots$. Очередное приближение $x^{(k)}$ находим в результате минимизации функции $F_k(x)$ вида (5.27) на открытом множестве X_1. Существование такой точки $x^{(k)}$ обеспечивается леммой 5.1. Точное решение возможно здесь редко, поэтому в качестве $x^{(k)}$ возьмем приближенное решение, т. е. будем исходить из следующих условий:

$$F_k(x^{(k)}) \leqslant \min_{x \in X_1} F_k(x) + \varepsilon_k, \quad \varepsilon_k \geqslant 0, \quad \lim_{k\to\infty} \varepsilon_k = 0.$$

Для минимизации $F_k(x)$ на множестве X_1 можно использовать, например, градиентные методы безусловной оптимизации, изложенные в предыдущем разделе. Это возможно, поскольку X_1 — открытое множество. Нужно лишь дополнительно после каждой итерации градиентного метода проверять, принадлежит ли очередное приближение множеству X_1, и в случае выхода за его пределы уменьшать величину шага.

Для иллюстрации описанного метода рассмотрим следующий простейший пример:

$$x \to \min_{x \in X}; \quad X = \{x \in \mathrm{R} \mid x \geqslant 0\}.$$

Здесь $n = 1$, $m = 1$, $f(x) = x$, $g_1(x) = -x$, $X_1 = \{x \mid x > 0\}$, $X_2 = \{0\}$. Очевидно, решением этой задачи служит граничная точка $x_{\min} = 0$. В качестве барьерной функции возьмем, например, $B(x) = -1/x$ при $x > 0$ и положим $r_k = k^{-1}$, $k = 1, 2, \ldots$. В этом случае функция (5.27) принимает вид $F_k(x) = x + (kx)^{-1}$. Нетрудно найти ее точку минимума $x^{(k)} = 1/\sqrt{k}$ на множестве положительных чисел. Имеем $\lim\limits_{k\to\infty} x^{(k)} = x_{\min}$. На рис. 5.16 дана геометрическая иллюстрация к процессу решения примера при $k = 1, 2, 3$. При увеличении k и приближении точки $x^{(k)}$ к границе допустимого множества линия графика функции $F_k(x)$ становится все более крутой, как бы воздвигая «барьер» и не давая возможности точке $x^{(k)}$ выйти за пределы допустимого множества.

Пусть $X \subset \mathrm{R}^n$ — компактное множество, у которого $X_1 \neq \varnothing$, функции f, g_1, g_2, ..., g_m непрерывны на X и

$$\inf_{x \in X_1} f(x) = \min_{x \in X} f(x). \tag{5.28}$$

Тогда каждая предельная точка последовательности, построенной с помощью метода барьерных функций, является точкой глобального минимума в задаче (5.26).

Если равенство (5.28) не выполняется, то метод барьерных функций может и не приводить к точке минимума. Так, например, в случае, изображенном на рис. 5.17, к точке минимума линейной функции $\langle c, x\rangle$ подобраться «изнутри» невозможно.

Достоинством метода барьерных функций является широта его применения. Как показывает приведенное выше утверждение, тре-

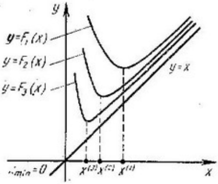

Рис. 5.16

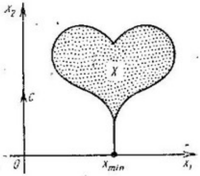

Рис. 5.17

буется лишь непрерывность функций, фигурирующих в постановке исходной задачи. Отметим еще одно важное свойство метода: его можно использовать и в локальном варианте. А именно: если в качестве $x^{(k)}$ выбирать точку л о к а л ь н о г о минимума функции $F_k(x)$, то получающаяся последовательность $\{x^{(k)}\}$ даст приближение к точке л о к а л ь н о г о минимума целевой функции.

Недостатком метода, затрудняющим непосредственное применение алгоритмов безусловной минимизации $F_k(x)$ для определения точек $x^{(k)}$, является ярко выраженная «овражная» структура барьерной функции при малых значениях r_k. Кроме того, напомним, что метод неприменим к задачам, в которых среди ограничений имеется хотя бы одно равенство.

4. Метод штрафных функций. Основная идея метода штрафных функций состоит в сведении исходной задачи минимизации с ограничениями (как в форме неравенств, так и в форме равенств) к последовательности задач без ограничений, для решения которых используются соответствующие методы безусловной минимизации. Внешне этот метод напоминает метод барьерных функций.

Будем рассматривать общую задачу нелинейного программирования

$$f(x) \to \min_{x \in X}; \tag{5.29}$$

$$X = \{x \in \mathbf{R}^n | g_j(x) \leqslant 0,\ j = 1, 2, \ldots, m;\ g_j(x) = 0,\ j = m+1, \ldots, s\},$$

в которой все функции f, g_j считаются непрерывными на всем пространстве $\mathbf{R}^n$.

Функция $P(x)$, определенная и непрерывная на $\mathbf{R}^n$, называется *штрафной функцией*, если выполняются следующие условия:

1) $P(x) = 0$ для всех $x \in X$;
2) $P(x) > 0$ для всех $x \notin X$.

Образно говоря, функция $P(x)$ назначает положительный «штраф» за выход за пределы допустимого множества X, тогда как для точек из X «штраф» отсутствует.

Указанными свойствами обладает, например, функция вида

$$P(x)=\sum_{j=1}^{m}(\max\{g_j(x);\ 0\})^p+\sum_{j=m+1}^{s}|g_j(x)|^p, \tag{5.30}$$

где p — произвольное фиксированное натуральное число. Если функции g_j непрерывно дифференцируемы (дважды непрерывно дифференцируемы), то и функция штрафа (5.30) непрерывно дифференцируема (дважды непрерывно дифференцируема) при $p\geqslant 2$ (соответственно $p\geqslant 3$). При $p=1$ функция (5.30) непрерывна, но дифференцируемой на $\mathbf{R}^n$ может и не быть.

Введем вспомогательную функцию ($k=1, 2, \ldots$)

$$\Phi_k(x)=f(x)+l_kP(x), \tag{5.31}$$

где $l_k>0$, и перейдем к описанию метода штрафных функций. Выберем некоторую последовательность положительных чисел $\{l_k\}$, обладающую следующими свойствами: $l_1<l_2<\ldots<l_k<\ldots$ и $\lim\limits_{k\to\infty} l_k=+\infty$. Например, можно положить $l_k=k$ или $l_k=10^k$, $k=1, 2, \ldots$. Для $k=1, 2, \ldots$ последовательно решаем задачи безусловной минимизации $\Phi_k(x)\to\min\limits_{k\in\mathbf{R}^n}$, где функция $\Phi_k(x)$ имеет вид (5.31). Обозначим через $x^{(k)}$ произвольное решение этой вспомогательной задачи. При определенных условиях последовательность точек $\{x^{(k)}\}$ дает приближение к решению задачи (5.29). Точное решение вспомогательной задачи безусловной минимизации возможно крайне редко (оно может и вообще отсутствовать при некоторых k), поэтому можно использовать приближение к нижней грани, т. е. исходить из условий

$$\Phi_k(x^{(k)})\leqslant\inf_{x\in\mathbf{R}^n}\Phi_k(x)+\varepsilon_k,\quad \varepsilon_k\geqslant 0;\quad \lim_{k\to\infty}\varepsilon_k=0.$$

При этом можно применять известные методы безусловной минимизации. Заканчивая описание метода, отметим, что точки последовательности $\{x^{(k)}\}$ не обязательно будут принадлежать допустимому множеству X.

Проиллюстрируем метод штрафных функций на примере задачи из п. 3. Рассмотрим в качестве штрафной функцию вида (5.30) при $p=2$, т. е. $P(x)=(\max\{-x;\ 0\})^2$. Положим $l_k=k$, $k=1, 2, \ldots$. Тогда $\Phi_k(x)=x+k(\max\{-x;\ 0\})^2$. Для определения минимума функции $\Phi_k(x)$ вычисляем ее производную по x и приравниваем нулю:

$$\Phi'_k(x)=1-2k\max\{-x;\ 0\}=0.$$

Отсюда находим $x^{(k)}=-(2k)^{-1}$, $k=1, 2, \ldots$. Очевидно, полученная последовательность точек сходится к решению исходной задачи:

$\lim\limits_{k\to\infty} x^{(k)}=0=x_{\min}$. Обратим внимание на то, что $x^{(k)}<0$, $k=1, 2, \ldots$, т. е. ни одна точка $x^{(k)}$ не является допустимой для исходной задачи. Геометрическая иллюстрация примера дана на рис. 5.18.

Прежде чем формулировать условия сходимости для метода штрафных функций, введем вспомогательное множество

$$\Omega=\{x\in \mathbf{R}^n | \Phi_1(x)\leqslant f(x^{(0)})+l_1\},$$

где $x^{(0)}$ — некоторое решение исходной задачи (5.29).

Пусть все функции f, g_i, участвующие в задаче (5.29), *непрерывны на* $\mathbf{R}^n$, *задача* (5.29) *имеет хотя бы одно решение и множество Ω не пусто и ограничено. Тогда каждая предельная точка последовательности $\{x^{(k)}\}$, построенной с помощью метода штрафных функций, является решением задачи* (5.29). Доказательство этого утверждения можно найти в книге Ю. Г. Евтушенко «Методы решения экстремальных задач и их применение в системах оптимизации» (М.: Наука, 1982).

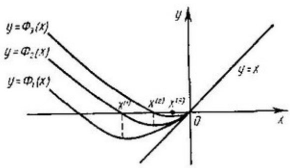

Рис. 5.18

Отметим, что функция $\Phi_k(x)$ при достаточно больших l_k становится овражной. Таким образом, недостатки метода барьерных функций, связанные с овражностью $F_k(x)$, в полной мере присущи и методу штрафных функций. Еще один недостаток метода штрафных функций состоит в том, что точки последовательности $\{x^{(k)}\}$, как правило, оказываются недопустимыми. В таких случаях, прекращая вычисления на некотором шаге, получают приближенное решение, которое не удовлетворяет ограничениям исходной задачи.

5. Конечный алгоритм решения простейшей задачи квадратичного программирования. В п. 5 § 5.2 приведен алгоритм метода сопряженных направлений, который является конечным в случае квадратичной целевой функции. Метод сопряженных направлений используется на промежуточных этапах алгоритмов решения задачи квадратичного программирования, причем свойство конечности удается сохранить.

Рассмотрим задачу квадратичного программирования

$$f(x)=\frac{1}{2}\langle Qx,\ x\rangle+\langle c,\ x\rangle\to\min;\ \ x_1,\ x_2,\ldots,\ x_n\geqslant 0,$$

в которой линейные ограничения общего вида отсутствуют, но имеется условие неотрицательности переменных. Матрицу Q считаем симметричной и неотрицательно-определенной. Для этой простейшей задачи квадратичного программирования конечный алгоритм формулируется следующим образом.

Пусть имеется k-е приближение — точка $x^{(k)}$, $x^{(k)} \geqq 0_n$ $(k \in \in \{0, 1, ...\})$. Если выполняется равенство

$$\partial f(x^{(k)})/\partial x_i = 0 \text{ для всех } i \text{ вида } x_i^{(k)} > 0,$$

а также неравенство

$$\partial f(x^{(k)})/\partial x_i \geqslant 0 \text{ для всех } i \text{ вида } x_i^{(k)} = 0,$$

то $x^{(k)}$ — решение задачи и вычисления окончены.

Это утверждение является прямым следствием теоремы 2.7, поскольку приведенные выше условия — это соотношения (2.32), записанные применительно к простейшей задаче квадратичного программирования.

В противном случае полагаем

$$J_k = \{j \in \{1, 2, ..., n\} \mid x_j^{(k)} > 0 \text{ или } \partial f(x^{(k)})/\partial x_j < 0,\ x_j^{(k)} = 0\}$$

и минимизируем квадратичную функцию f с помощью метода сопряженных направлений по переменным x_j, $j \in J_k$, принимая $x_j = 0$, $j \notin J_k$. При минимизации надо следить за сохранением неотрицательности переменных. Это делают следующим образом. Обозначают через $\tilde{x}$ вектор тех переменных, по которым ведется минимизация. Согласно методу сопряженных направлений, на шаге l вычисляют

$$\alpha_{k_l} = -\frac{\langle \nabla f(\tilde{x}^{(k_l)}),\ \tilde{p}^{(k_l)} \rangle}{\langle \tilde{Q}\tilde{p}^{(k_l)},\ \tilde{p}^{(k_l)} \rangle},$$

где $\tilde{p}^{(k_l)}$ и $\tilde{Q}$ имеют аналогичный вектору $\tilde{x}$ смысл. Кроме того, необходимо найти

$$\bar{\alpha}_{k_l} = \min\{-\tilde{x}_j^{(k_l)}/\tilde{p}_j^{(k_l)} \mid j \in J_k,\quad \tilde{p}_j^{(k_l)} < 0\},$$

а если $\tilde{p}_j^{(k_l)} \geqslant 0$ имеет место для всех $j \in J_k$, то полагают $\bar{\alpha}_{k_l} = +\infty$. Теперь следует сравнить α_{k_l} и $\bar{\alpha}_{k_l}$. Если $\alpha_{k_l} < \bar{\alpha}_{k_l}$, то очередную точку $\tilde{x}^{(k_l+1)}$ строят по формуле $\tilde{x}^{(k_l+1)} = \tilde{x}^{(k_l)} + \alpha_{k_l}\tilde{p}^{(k_l)}$; если же $\alpha_{k_l} \geqslant \bar{\alpha}_{k_l}$, то используют формулу $\tilde{x}^{(k_l+1)} = \bar{\alpha}_{k_l}\tilde{p}^{(k_l)}$. Если при минимизации переменная x_j, $j \in J_k$, обращается в нуль, то номер j из множества J_k нужно исключить. Полученная в результате применения метода сопряженных направлений точка с включением отброшенных ранее нулевых компонент дает $x^{(k+1)}$.

6. Метод линеаризации. Этот метод предназначен для решения задач нелинейного программирования, и его суть в общих чертах состоит в следующем. На каждом шаге ограничения задачи линеаризуют (т. е. вместо исходных ограничений используют их линейные аппроксимации в очередной точке), а целевую функцию также линеаризуют, но, для того чтобы получающаяся линеаризованная задача имела решение, к целевой функции добавляют некоторое квадратичное слагаемое. Решение этой задачи квадратичного программирования принимают за следующее приближение.

Будем рассматривать задачу нелинейного программирования

$$f_0(x) \to \min; \quad f_i(x) \leqslant 0, \quad i \in I = \{1, 2, \dots, m\},$$

в которой все функции предполагаем дифференцируемыми. Введем функцию

$$F(x) = \max\{f_i(x) \mid i \in I\}$$

и будем считать, что $F(x) \geqslant 0$ для всех x (это допустимо, так как в число функций в ограничениях задачи всегда можно включить функцию, тождественно равную нулю). Введем множество

$$I_\delta(x) = \{i \in I \mid f_i(x) \geqslant F(x) - \delta\}, \quad \delta \geqslant 0.$$

Пусть $x^{(0)}$ — начальная точка, ε $(0 < \varepsilon < 1)$ — фиксированное число и $x^{(k)}$ — k-е приближение. Кроме того, пусть N — положительная константа, обладающая тем свойством, что множество вида

$$\{x \in \mathbf{R}^n \mid f_0(x) + NF(x) \leqslant f_0(x^{(0)}) + NF(x^{(0)})\}$$

ограничено.

Очередное приближение строят следующим образом. При $x = x^{(k)}$ решают задачу квадратичного программирования

$$\langle \nabla f_0(x), p \rangle + \frac{1}{2}\|p\|^2 \to \min_p;$$

$$\langle \nabla f_i(x), p \rangle + f_i(x) \leqslant 0, \quad i \in I_\delta(x).$$

Проще всего это выполнить, перейдя к соответствующей двойственной задаче, которая в данном случае состоит в максимизации по неотрицательным λ_i, $i \in I_\delta(x)$, функции вида

$$\min_{p \in \mathbf{R}^n} \left[\langle \nabla f_0(x), p \rangle + \frac{1}{2}\|p\|^2 + \sum_{i \in I_\delta} \lambda_i (\langle \nabla f_i(x), p \rangle + f_i(x)) \right]. \quad (5.32)$$

Вектор p, реализующий здесь минимум, легко найти, дифференцируя по p выражение в квадратных скобках и приравнивая полученный результат нулю:

$$p = -\nabla f_0(x) - \sum_{i \in I_\delta(x)} \lambda_i f_i(x). \quad (5.33)$$

Подставив найденное выражение для p в (5.32), получим, что двойственная задача сводится к максимизации по неотрицательным λ_i, $i \in I_\delta(x)$, квадратичной функции

$$-\frac{1}{2}\left\| \nabla f_0(x) + \sum_{i \in I_\delta(x)} \lambda_i \nabla f_i(x) \right\|^2 + \sum_{i \in I_\delta(x)} \lambda_i f_i(x).$$

Это уже простейшая задача квадратичного программирования, и для ее решения может быть применен конечный алгоритм предыдущего пункта. Обозначим решение указанной задачи через $\lambda_i^{(k)}$. Подставляя $\lambda_i=\lambda_i^{(k)}$ в (5.33), найдем соответствующий вектор $p=p^{(k)}$, который определит направление движения из точки $x^{(k)}$. Для установления длины шага последовательно проверяем неравенства

$$f_0(x^{(k)}+2^{-i}p^{(k)})+NF(x^{(k)}+2^{-i}p^{(k)})\leqslant f_0(x^{(k)})+$$
$$+NF(x^{(k)})-2^{-i}\varepsilon\|p^{(k)}\|^2$$

для $i=0, 1, 2, \ldots$ и фиксируем первое значение $i=i_0$, при котором неравенство выполнится. Окончательно следует положить

$$\alpha_k=2^{-i_0},\ x^{(k+1)}=x^{(k)}+\alpha_k p^{(k)}.$$

Константу N подбирают таким образом, чтобы выполнялось неравенство $N\geqslant\gamma=\sum\limits_{i\in I_\delta(x^{(k)})}\lambda_i$. Очень большим N выбирать не следует, поскольку это приводит к слишком малой длине шага α_k. Если на некотором шаге оказывается $N\leqslant\gamma$, то N заменяют на $N=2\gamma$. Величину δ нужно уменьшать в том случае, если вспомогательная задача квадратичного программирования не имеет решения. Вообще же значение δ желательно иметь по возможности бóльшим.

При довольно общих предположениях [21] последовательность $\{x^{(k)}\}$, построенная методом линеаризации, обладает следующим свойством: *каждая ее предельная точка удовлетворяет исходным ограничениям* $f_j(x)\leqslant 0,\ j=1, 2, \ldots, m$, *и является стационарной точкой функции Лагранжа* $f_0(x)+\sum\limits_{i\in I_\delta(x)}\lambda_i f_i(x)$ *при некоторых неотрицательных* λ_i.

Метод линеаризации универсален, и его можно использовать при решении задач нелинейного программирования, в которых кроме ограничений-неравенств имеются и ограничения равенства.

§ 5.4. ПРИКЛАДНЫЕ ЗАДАЧИ

1. Задача определения параметров функции. Пусть некоторый процесс характеризуется следующей зависимостью:

$$y=a+bx+e^{-cx}, \tag{5.34}$$

где a, b, c — параметры, подлежащие определению. В результате проведенных измерений установлено несколько значений искомой функции при определенных значениях переменной (табл. 5.1).

Требуется определить значения параметров a, b, c, при которых отклонение значений функции (5.34) от значений, приведенных в табл. 5.1, было бы наименее возможным.

Для того чтобы сформулировать соответствующую задачу оптимизации, необходимо указать «меру рассогласования», т. е. выбрать функцию, с помощью которой вычисляются отклонения фактических значений функции (5.34) от значений, приведенных в таблице. Для определенности условимся отклонение одного вектора из $\mathbf{R}^5$ от другого измерять как евклидово расстояние. Тогда, ис-

Таблица 5.1

x	0,5	2,4	8	-1	-3
y	$-1,5$	12	50	8	8077

пользуя эквивалентность задач $\sqrt{h(x)} \to \min$ и $h(x) \to \min$, приходим к следующей задачей безусловной минимизации:

$$f(a, b, c) = (a + 0{,}5b + e^{-0{,}5c} + 1{,}5)^2 + (a + 2{,}4b + e^{-2{,}4c} - 12)^2 +$$
$$+ (a + 8b + e^{-8c} - 50)^2 + (a - b + e^{c} - 8)^2 +$$
$$+ (a - 3b + e^{3c} - 8077)^2 \to \min_{a,b,c}.$$

Это задача нелинейного (невыпуклого) программирования. Используем для ее решения численные методы, изложенные в § 5.2. В качестве начального приближения выберем начало координат, т. е. положим $x^{(0)} = (0, 0, 0)$.

Результаты, полученные в случае применения **метода покоординатного спуска** с шагом единичной длины, приведены в табл. 5.2.

Как видим, на протяжении 15 итераций величину шага делить не пришлось. Значение целевой функции уменьшилось от 65224353 до 25,86, т. е. почти в 2522210 раза. Поскольку целевая функция не может принимать отрицательные значения, минимум находится в пределах от 0 до 25,86, а значит, можно считать, что в точке $(-3, 7{,}3)$ значение функции близко к минимальному.

Можно проверить, что $f(-5, 7{,}3) \approx 0$. Это означает, что $(-5, 7{,}3)$ — точка приближенного глобального минимума, и с помощью метода покоординатного спуска мы приблизились к ней достаточно близко. Обратим внимание на «непрямолинейный» характер приближения: на протяжении первых шести итераций по переменным a и b имело место удаление от точки глобального минимума, а не приближение к ней. Наконец, отметим, что приближение именно к точке глобального минимума оказалось возможным благодаря «удачному» выбору начальной точки и величины шага.

Перейдем к рассмотрению **метода градиентного спуска**. Найдем частные производные целевой функции и вычислим градиент в начальной точке $x^{(0)} = (0, 0, 0)$:

$$\nabla f(0, 0, 0) = (-16281;\ 47636;\ -47636).$$

Таблица 5.2

Номер итерации	a	b	c	f(a, b, c)
1	0	0	0	65224353
	1	0	0	65208077
	1	1	0	65255810
	1	—1	0	65157363
	1	—1	1	64852935
2	2	—1	1	64836597
	2	0	1	64884101
	2	—2	1	64789395
	2	—2	2	58365374
3	3	—2	2	58749843
	3	—1	2	58794785
	3	—3	2	58705114
	3	—3	3	7091,15
4	4	—3	3	7040,5
	4	—2	3	5911,09
	4	—2	4	$\approx 2\cdot 10^{10}$
	4	—2	2	58765100
5	5	—2	3	5632,08
	5	—1	3	4424,76
	5	—1	4	$\approx 2\cdot 10^{10}$
	5	—1	2	58764006
6	6	—1	3	4421,74
	6	0	3	3388
	6	0	4	$\approx 2\cdot 10^{10}$
	6	0	2	64819273
7	7	0	3	3408,77
	5	0	3	3377,21
	5	1	3	2489,67
	5	1	4	$\approx 2\cdot 10^{10}$
	5	1	2	58853010
8	6	1	3	2544,26
	4	1	3	2475,09
	4	2	3	1733,78
	4	2	4	$\approx 2\cdot 10^{10}$
	4	2	2	58915328
9	5	2	3	1762,17
	3	2	3	1715,4
	3	3	3	1120,31
	3	3	4	$\approx 2\cdot 10^{10}$
	3	3	2	58930372
10	4	3	3	1152,45
	2	3	3	1006,13
	2	4	3	649,26
	2	4	4	$\approx 2\cdot 10^{10}$
	2	4	2	59037577
11	3	4	3	685,24
	1	4	3	623,28
	1	5	3	320,63
	1	5	4	$\approx 2\cdot 10^{10}$
	1	5	2	59098842
12	2	5	3	360,42
	0	5	3	290,84
	0	6	3	134,41
	0	6	4	$\approx 2\cdot 10^{10}$
	0	6	2	59160396
13	1	6	3	207,99
	—1	6	3	100,83
	—1	7	3	90,62
	—1	7	4	$\approx 2\cdot 10^{10}$
	—1	7	2	59221893
14	0	7	3	138
	—2	7	3	53,25
	—2	8	3	189,26
	—2	6	3	77,26
	—2	7	4	$\approx 2\cdot 10^{10}$
	—2	7	2	59237287
15	—1	7	3	90,62
	—3	7	3	25,86

Далее воспользуемся формулой

$$x^{(1)}=(0,\ 0,\ 0)-\alpha_0\nabla f\,(0,\ 0,\ 0)=-\alpha_0(-16281;\ 47636;\ -47636).$$

Проверка показывает, что при $\alpha_0 \geqslant 10^{-4}$ значение $f(x^{(1)})$ больше, чем $f(x^{(0)}$, а при $\alpha_0 \leqslant 10^{-5}$ продвижение из $x^{(0)}$ в $x^{(1)}$ незначительно. Поэтому выберем среднее, т. е. примем $\alpha_0=5\cdot 10^{-5}$. В результате получаем $x^{(1)}=(0{,}58;\ 2{,}38;\ 2{,}38)$, $f(x^{(1)})=46542509<f(x^{(0)})$.

Вычисляем градиент в полученной точке $\nabla f(x^{(1)})=(-13707;$ $40419;\ -51633754)$. Имеем $x^{(2)}=(0{,}58;\ 2{,}38;\ 2{,}38)-\alpha_1\ (-13707;$ $40419;\ -51633754)$. Здесь величина шага 10^{-7} оказывается слишком большой, поэтому принимаем $\alpha_1=10^{-8}$ и находим $x^{(2)}=(0{,}58;$ $2{,}38;\ 2{,}9)$, $f(x^{(2)})=4330134{,}7<f(x^{(1)})$. Затем определяем $\nabla f(x^{(2)})=$ $=(-4210;\ 11957;\ -74939424)$ и вычисляем $x^{(3)}=(0{,}58;\ 2{,}38;\ 2{,}9)-$ $-\alpha_2\ (-4210;\ 11957;\ -74939424)$. При $\alpha_2=10^{-9}$ (величина 10^{-8} ведет к увеличению значения целевой функции) получаем $x^{(3)}=(0{,}58;$ $2{,}38;\ 2{,}98)$, $f(x^{(3)})=204635<f(x^{(2)})$.

Анализ показывает, что дальнейшие вычисления согласно градиентному методу приведут к приближению третьей переменной c к оптимальному значению, равному 3, тогда как первые две переменные практически не изменятся. Эта ситуация характерна для функций «овражного» типа.

Применим теперь **м е т о д Н ь ю т о н а**. Вычислим матрицу вторых производных целевой функции в начальной точке

$$\nabla^2 f\,(0,\ 0,\ 0)=\begin{pmatrix} 10 & 13{,}8 & -13{,}8 \\ 13{,}8 & 16{,}0 & -160 \\ -13{,}8 & 160 & -151617 \end{pmatrix}$$

и найдем обратную матрицу:

$$[\nabla^2 f\,(0,\ 0,\ 0)]^{-1}=\frac{1}{151732}\begin{pmatrix} 17228 & -1485{,}73 & 0 \\ -1485{,}73 & 1075{,}62 & -1 \\ 0 & -1 & -1 \end{pmatrix}.$$

В соответствии с формулой (5.13) имеем

$$x^{(1)}=x^{(0)}-[\nabla^2 f\,(x^{(0)})]^{-1}\,\nabla f\,(x^{(0)})=(2315;\ -497{,}42;\ 0)^{\mathrm{T}},$$

причем $f(x^{(1)})=34540556<f(x^{(0)})$. Далее находим

$$\nabla^2 f(x^{(1)})=\begin{pmatrix} 10 & 13{,}8 & -13{,}8 \\ 13{,}8 & 160 & -160 \\ -13{,}8 & -160 & -278618{,}25 \end{pmatrix};$$

$$[\nabla^2 f\,(x^{(1)}]^{-1}=\frac{1}{278536}\begin{pmatrix} 31643{,}75 & -2728{,}93 & 0 \\ -2728{,}93 & 1976{,}49 & -1 \\ 0 & -1 & -1 \end{pmatrix};$$

$$\nabla f(x^{(1)}) = (4{,}88;\ -14{,}13;\ 14{,}13)^{т};$$

$$x^{(2)} = x^{(1)} - [\nabla^2 f(x^{(1)})]^{-1} \nabla f(x^{(1)}) = (2314{,}33;\ -497{,}27;\ 0)^{т},$$

причем $f(x^{(2)}) = 34540531 < f(x^{(1)})$. Можно проверить, что частные производные целевой функции в точке $x^{(2)}$ близки к нулю. Следовательно, применение метода Ньютона приводит в окрестность некоторой стационарной точки, которая находится на значительном удалении от точки глобального минимума.

2. Задача распределения воды по отстойникам. Пусть в системе сооружений, предназначенной для биологической очистки сточных вод, имеется n отстойников для удаления компонентов загрязнения с удельным весом бо́льшим, чем у воды. Тяжелые частицы оседают на дно и с помощью специального насоса периодически удаляются, а верхний слой сливается и проходит следующие стадии очистки. Пусть в систему очистных сооружений подается количество воды Q. Требуется это количество воды распределить по отстойникам таким образом, чтобы на выходе суммарное количество взвешенных частиц было минимальным, т. е. чтобы очистка была наиболее полной. Эта задача имеет смысл, даже если все отстойники имеют одинаковую конструкцию, поскольку они, как правило, обладают различными характеристиками по удалению взвешенных частиц.

Обозначим через Q_i количество воды, поступающей в i-й отстойник, а через c_i — концентрацию взвешенных частиц на выходе i-го отстойника, $i=1, 2, \ldots, n$. Задача сводится к минимизации функции

$$\sum_{i=1}^{n} c_i Q_i$$

при ограничениях

$$\sum_{i=1}^{n} Q_i = Q,\quad Q_i \geqslant 0,\quad i=1,2,\ldots, n.$$

Будем считать, что концентрация c_i взвешенных частиц пропорциональна количеству воды Q_i, подаваемому в этот отстойник, т. е. примем $c_i = \alpha_i Q_i$, где α_i $(\alpha_i > 0)$ — коэффициент пропорциональности, характеризующий «качество работы» i-го отстойника. Вводя переменные $x_i = Q_i/Q$, $i=1, 2, \ldots, n$, окончательно приходим к следующей задаче квадратичного программирования:

$$\sum_{i=1}^{n} \alpha_i x_i^2 \to \min;$$

$$\sum_{i=1}^{n} x_i = 1,\quad x_i \geqslant 0,\quad i=1, 2,\ldots, n.$$

Рассмотрим задачу описанного типа со следующими числовыми данными:

$$f(x)=x_1^2+1{,}2x_2^2+1{,}3x_3^2+1{,}5x_4^2+1{,}5x_5^2\to\min;$$

$$\sum_{i=1}^{5} x_i=1,\ \ x_1,\ x_2,\dots,\ x_5\geqslant 0.$$

Для ее решения воспользуемся методом, изложенным в п. 5 § 5.3. Предварительно «избавимся» от ограничения-равенства, выразив переменную x_1 через остальные и подставив полученное выражение в формулу для f:

$$f=(1-x_2-\dots-x_5)^2+1{,}2x_2^2+\dots+1{,}5x_5^2=$$

$$=2{,}2x_2^2+2{,}3x_3^2+2{,}5x_4^2+2{,}5x_5^2+2x_2x_3+2x_2x_4+2x_2x_5+$$

$$+2x_3x_4+2x_3x_5+2x_4x_5-2x_2-2x_3-2x_4-2x_5+1.$$

Для упрощения записи последнее слагаемое далее учитывать не будем и, кроме того, обозначим $x=(x_2,\ x_3,\ x_4,\ x_5)^{\mathrm{T}}$. Находим

$$\nabla f(x)=\begin{pmatrix}4{,}4x_2+2x_3+2x_4+2x_5-2\\ 2x_2+4{,}6x_3+2x_4+2x_5-2\\ 2x_2+2x_3+5x_4+2x_5-2\\ 2x_2+2x_3+2x_4+5x_5-2\end{pmatrix},\quad Q=\begin{pmatrix}4{,}4&2&2&2\\ 2&4{,}6&2&2\\ 2&2&5&2\\ 2&2&2&5\end{pmatrix}.$$

В качестве начальной возьмем точку $x^{(2)}=(0{,}2;\ 0{,}2;\ 0{,}2;\ 0{,}2)^{\mathrm{T}}$. Далее определяем:

$$\nabla f(x^{(0)})=(0{,}08;\ 0{,}12;\ 0{,}2;\ 0{,}2)^{\mathrm{T}},$$

$$p^{(0)}=-\nabla f(x^{(0)})=(-0{,}08;\ -0{,}12;\ -0{,}2;\ -0{,}2)^{\mathrm{T}},$$

$$\alpha_0=-\frac{\langle\nabla f^{(0)},\ p^{(0)}\rangle}{\langle Qp^{(0)},\ p^{(0)}\rangle}=-\frac{-0{,}1}{1{,}012}=0{,}099,$$

$$\bar\alpha_0=\min\{0{,}2/0{,}08;\ 0{,}2/0{,}12;\ 0{,}2/0{,}2\}=1.$$

Так как $\alpha_0<\bar\alpha_0$, то

$$x^{(1)}=x^{(0)}+\alpha_0p^{(0)}=(0{,}192;\ 0{,}188;\ 0{,}18;\ 0{,}18)^{\mathrm{T}}.$$

Далее выполняем следующий шаг:

$$\nabla f(x^{(1)})=(-0{,}059;\ -0{,}031;\ 0{,}02;\ 0{,}02)^{\mathrm{T}},$$

$$\beta_0=\frac{\langle\nabla f(x^{(1)}),\ \nabla f(x^{(1)})\rangle}{\langle\nabla f(x^{(0)}),\ \nabla f(x^{(0)})\rangle}=0{,}053,$$

$$p^{(1)}=-\nabla f(x^{(1)})+\beta_0p^{(0)}=(0{,}055;\ 0{,}025;\ -0{,}03;\ -0{,}03)^{\mathrm{T}},$$

$$\alpha_1=-\frac{\langle\nabla f(x^{(1)}),\ p^{(1)}\rangle}{\langle Qp^{(1)},\ p^{(1)}\rangle}=0{,}347,\quad \bar\alpha_1=0{,}18/0{,}03=0{,}6>\alpha_1,$$

$$x^{(2)}=x^{(1)}+\alpha_1p^{(1)}=(0{,}211;\ 0{,}197;\ 0{,}17;\ 0{,}17)^{\mathrm{T}}.$$

Наконец, имеем

$$\nabla f(x^{(2)}) = (0{,}0024;\ 0{,}008;\ 0{,}006;\ 0{,}006)^{\mathrm{T}},$$

$$\beta_1 = \frac{\langle \nabla f(x^{(2)}),\ \nabla f(x^{(2)}) \rangle}{\langle \nabla f(x^{(1)}),\ \nabla f(x^{(1)}) \rangle} = 0{,}027,$$

$$p^{(2)} = -\nabla f(x^{(2)}) + \beta_1 p^{(1)} = (-0{,}001;\ -0{,}0073;\ -0{,}0068;\ -0{,}0068)^{\mathrm{T}},$$

$$\alpha_2 = -\frac{\langle \nabla f(x^{(2)}),\ p^{(2)} \rangle}{\langle Qp^{(2)},\ p^{(2)} \rangle} = 0{,}117,$$

$$\bar{\alpha}_2 = \min\{0{,}211/0{,}001;\ 0{,}197/0{,}0073;\ 0{,}17/0{,}0068\} = 2{,}5 > \alpha_2,$$

$$x^{(3)} = x^{(2)} + \alpha_2 p^{(2)} = (0{,}211;\ 0{,}196;\ 0{,}169;\ 0{,}169)^{\mathrm{T}}.$$

Учитывая, что $x_1 = 1 - x_2 - \ldots - x_5$, окончательно запишем найденное приближенное решение: $(0{,}255;\ 0{,}211;\ 0{,}196;\ 0{,}169;\ 0{,}169)^{\mathrm{T}}$.

Глава 6
ГЕОМЕТРИЧЕСКОЕ ПРОГРАММИРОВАНИЕ

Геометрическое программирование, как отмечалось в гл. 2, позволяет решать задачи минимизации нелинейного программирования с функциями специального вида, которые называются позиномами. Многие технические задачи оптимизации удается сформулировать именно в терминах геометрического программирования. Решение подобных задач, как правило, можно свести к решению более простых задач максимизации, в которых целевая функция является вогнутой, а ограничения — линейные. Такое сведение обосновывает приведенная в § 6.3 теорема двойственности геометрического программирования.

§ 6.1. ЗАДАЧА ГЕОМЕТРИЧЕСКОГО ПРОГРАММИРОВАНИЯ, ПРЕОБРАЗОВАННАЯ И ДВОЙСТВЕННАЯ ЗАДАЧИ

1. Постановка задачи геометрического программирования. Выражение вида

$$c x_1^{a_1} x_2^{a_2} \ldots x_n^{a_n},$$

где $c > 0$ и $a_i \in \mathbf{R}$, $i = 1, 2, \ldots, n$, называют *одночленным позиномом*. Сумму конечного числа одночленных позиномов, т. е. функцию

$$f(x) = \sum_{i=1}^{m} c_i x_1^{a_{i1}} x_2^{a_{i2}} \ldots x_n^{a_{in}} \quad (x > 0_n)$$

переменных $x_1, x_2, \ldots, x_n$, называют *позиномом* или *позиномиальной функцией*. Все коэффициенты $c_1, c_2, \ldots, c_m$ считают положительными, тогда как показатели степеней a_{ij} $(i = 1, 2, \ldots, m;\ j = 1, 2, \ldots,$

n) — произвольные вещественные числа. Область определения позинома состоит из тех векторов пространства $\mathbb{R}^n$, координаты которых положительны.

Простейшими примерами позиномов являются следующие функции: $f_1(x_1, x_2) = x_1 + 2x_1\sqrt[3]{x_2} + (x_2/x_1)^2$, $f_2(x_1, x_2, ..., x_n) = 1/x_1 + 2/x_2^2 + ... + n/x_n^n$. Очевидно, сумма и произведение любых двух позиномов также являются позиномами.

Задача геометрического программирования — это задача математического программирования, в которой все функции являются позиномиальными.

Для позинома $f(x)$ неравенство $f(x) \leqslant 0$ при $x > 0_n$ решений не имеет, поэтому ограничение с позиномиальной функцией записывают в виде $f(x) \leqslant b$, где $b > 0$. Последнее неравенство эквивалентно неравенству $f(x)/b \leqslant 1$, где $f(x)/b$ — позином. Поэтому можно сразу считать, что ограничения задачи геометрического программирования имеют вид $f_k(x) \leqslant 1$.

В соответствии с этим *задача геометрического программирования* формулируется следующим образом:

$$f_0(x) \to \min;$$

$$f_k(x) \leqslant 1, \; k = 1, 2, ..., p,$$

$$x = (x_1, x_2, ..., x_n)^{\mathrm{T}} > 0_n,$$

где все функции $f_0, f_1, ..., f_p$ — позиномы. Ограничения, состоящие в требовании положительности всех переменных, называют *естественными ограничениями*, а ограничения вида $f_k(x) \leqslant 1$ — *вынужденными*. Естественные ограничения должны быть в каждой задаче геометрического программирования; вынужденных ограничений может и не быть (в этом случае $p = 0$).

Обозначим через m_j число одночленных позиномов в f_j, $j = 0, 1, ..., p$. Пусть $\sum\limits_{j=0}^{p} m_j = m$. Введем следующие множества индексов:

$$J_0 = \{1, 2, ..., m_0\}, \; J_1 = \{m_0 + 1, \; m_0 + 2, ..., m_0 + m_1\}, ...$$

$$..., J_k = \left\{\sum_{j=0}^{k-1} m_j + 1 + , ..., \sum_{j=0}^{k} m_j\right\}, ..., J_p = \left\{\sum_{j=0}^{p-1} m_j + 1, ..., m\right\}.$$

Легко видеть, что $\bigcup\limits_{k=0}^{p} J_k = \{1, 2, ..., m\}$.

Теперь, если все члены позинома f_k занумеровать числами из множества J_k, можно записать

$$f_k(x) = \sum_{i \in J_k} c_i x_1^{a_{i1}} x_2^{a_{i2}} ... x_n^{a_{in}}, \quad k = 0, 1, ..., p. \tag{6.1}$$

Матрицу A, составленную из элементов a_{ij}, называют *матрицей экспонент*. Она содержит m строк и n столбцов.

Запишем задачу геометрического программирования в развернутом виде:

$$f_0(x)=\sum_{i\in J_0} c_i x_1^{a_{i1}} x_2^{a_{i2}} \dots x_n^{a_{in}} \to \min;$$

$$f_1(x)=\sum_{i\in J_1} c_i x_1^{a_{i1}} x_2^{a_{i2}} \dots x_n^{a_{in}} \leqslant 1,$$

$$f_2(x)=\sum_{i\in J_2} c_i x_1^{a_{i1}} x_2^{a_{i2}} \dots x_n^{a_{in}} \leqslant 1,$$

$$\dots\dots\dots\dots\dots\dots\dots\dots$$

$$f_p(x)=\sum_{i\in J_p} c_i x_1^{a_{i1}} x_2^{a_{i2}} \dots x_n^{a_{in}} \leqslant 1,$$

$$x_1,\ x_2, \dots, x_n > 0.$$

Это задача нелинейного программирования, и в общем случае она не является задачей выпуклого программирования.

Любая задача геометрического программирования может быть сформулирована так, что ранг ее матрицы экспонент A будет совпадать с числом неизвестных n. Действительно, если $\operatorname{rang} A<n$, то среди столбцов матрицы A есть линейно зависимые от остальных. Не уменьшая общности, предположим, что последний столбец матрицы A линейно зависит от остальных, т. е. для некоторых чисел $\lambda_1, \lambda_2, \dots, \lambda_{n-1}$ выполняется условие $a_{in}=\sum_{j=1}^{n-1} \lambda_j a_{ij}, \quad i=1,2, \dots, m.$

Вводя новые переменные $\bar{x}_1, \dots, \bar{x}_{n-1}$ с помощью равенств $x_j=\bar{x}_j x_n^{-\lambda_j}$, $j=1, 2, \dots, n-1$ получаем

$$x_1^{a_{i1}} x_2^{a_{i2}} \dots x_n^{a_{in}}=(\bar{x}_1 x_n^{-\lambda_1})^{a_{i1}} (\bar{x}_2 x_n^{-\lambda_2})^{a_{i2}} \dots (\bar{x}_{n-1} x_n^{-\lambda_{n-1}})^{a_{i,n-1}} \times$$
$$\times\ x_n^{\sum_{j=1}^{n-1} \lambda_j a_{ij}}=\bar{x}_1^{a_{i1}} \bar{x}_2^{a_{i2}} \dots \bar{x}_{n-1}^{a_{i,n-1}}, \quad i=1, 2, \dots, m.$$

Это означает, что в каждом одночленном позиноме задачи геометрического программирования множители $x_n^{a_{in}}$ можно вычеркнуть (т. е. положить $x_n=1$) и это не повлияет на оптимальное значение самой задачи. Исключая все столбцы матрицы A, линейно зависимые от остальных, придем к матрице, ранг которой совпадает с числом неизвестных. Эта матрица получается из исходной вычеркиванием столбцов, линейно зависимых от остальных. Выбор таких столбцов не однозначен, поэтому и сокращенная матрица определяется не однозначно. Исходная задача геометрического программирования и задача с сокращенной матрицей экспонент либо обе имеют оптимальное решение, либо обе их не имеют, причем оптимальные значения этих задач (если они существуют) равны. Зная оптимальное решение задачи геометрического программирования с сокращенной матрицей экспонент, всегда можно записать оптимальное решение исходной задачи. Так, если была исключена только переменная x_n и $\bar{x}_1^*, \bar{x}_2^*, \ldots, \bar{x}^*_{n-1}$ — оптимальное решение новой задачи, то $x_i=\bar{x}_i^*$, $i=1, 2, \ldots, n-1$, $x_n=1$ — одно из оптимальных решений исходной задачи геометрического программирования.

Итак, всегда можно предполагать выполненным равенство $\operatorname{rang} A=n$.

2. Двойственная задача геометрического программирования. Рассмотрим функцию вида

$$v(\delta)=\prod_{i=1}^{m}\left(\frac{c_i}{\delta_i}\right)^{\delta_i}\prod_{k=0}^{p}\lambda_k^{\lambda_k}, \quad \delta \geqq 0_m, \tag{6.2}$$

переменных $\delta_1, \delta_2, \ldots, \delta_m$, где

$$\lambda_k=\sum_{i\in J_k}\delta_i, \quad k=0, 1, \ldots, p, \tag{6.3}$$

а $c_1, c_2, \ldots, c_m$ — коэффициенты задачи геометрического программирования. При вычислении значений этой функции считаем, что $\delta_i^{\delta_i}=1$ при $\delta_i=0$ и $\lambda_k^{\lambda_k}=1$ при $\lambda_k=0$. При этих допущениях функция $v(\delta)$ непрерывна при всех $\delta\geqq 0_m$. Ее называют *двойственной функцией,* а переменные $\delta_1, \delta_2, \ldots, \delta_m$ — *двойственными.* Если имеется некоторая задача геометрического программирования, то по коэффициентам $c_1, c_2, \ldots, c_m$, можно однозначно записать двойственную функцию. Каждый множитель $\lambda_k^{\lambda_k}$ $(k=1, 2, \ldots, p)$ в двойственной функции отвечает вынужденному ограничению $f_k(x)\leqslant 1$ задачи геометрического программирования.

Теперь сформулируем *двойственную задачу геометрического программирования:*

$$v(\delta)\to\max;$$

$$A^{\mathrm{T}}\delta=0_n, \quad \sum_{i\in J_0}\delta_i=1, \quad \delta\geqq 0_m,$$

где A — матрица экспонент задачи геометрического программирования. В подробной записи двойственная задача имеет вид

$$v(\delta)=\prod_{i=1}^{m}\left(\frac{c_i}{\delta_i}\right)^{\delta_i}\prod_{k=0}^{p}\lambda_k^{\lambda_k}\to\max;$$

$$\begin{aligned}
&a_{11}\delta_1+a_{21}\delta_2+\ldots+a_{m1}\delta_m=0,\\
&a_{12}\delta_1+a_{22}\delta_1+\ldots+a_{m2}\delta_m=0,\\
&\ldots\ldots\ldots\ldots\ldots\ldots\\
&a_{1n}\delta_1+a_{2n}\delta_2+\ldots+a_{mn}\delta_m=0,\\
&\delta_1+\delta_2+\ldots+\delta_{m_0}=1; \quad \delta_1, \delta_2, \ldots, \delta_m\geqslant 0,
\end{aligned}$$

где λ_k, $k=0, 1, \ldots, p$, определяются равенствами (6.3). Число двойственных переменных m равно общему числу одночленных позиномов в исходной задаче. Сумма первых m_0 двойственных переменных должна быть равна единице (m_0 — число одночленных позиномов целевой функции задачи геометрического программирования).

Для каждой задачи геометрического программирования можно записать (и притом только одну) соответствующую ей двойственную задачу. Рассмотрим, например, задачу геометрического программирования вида

$$\frac{x_1}{x_2x_3}+\frac{x_2}{x_1x_3}+\frac{x_3}{x_1x_2}\to \min;\quad \frac{x_1\sqrt{x_3}}{\sqrt[4]{x_2}}\leqslant 1,\quad x_1,\ x_2,\ x_3>0. \qquad (6.4)$$

Здесь одно вынужденное ограничение. Матрица экспонент содержит четыре строки и три столбца:

$$A=\begin{pmatrix} 1 & -1 & -1\\ -1 & 1 & -1\\ -1 & -1 & 1\\ 1 & -1/4 & 1/2\end{pmatrix},$$

причем $\operatorname{rang} A=3$, т. е. число переменных уменьшить нельзя. Ограничение содержит только одночленный позином, поэтому $\lambda_1=\delta_4$. Учитывая, что $c_1=c_2=c_3=c_4=1$, запишем двойственную функцию:

$$v(\delta)=\delta_1^{-\delta_1}\delta_2^{-\delta_2}\delta_3^{-\delta_3}\delta_4^{-\delta_4}\delta_4^{\delta_4}.$$

Таким образом, задаче геометрического программирования (6.4) соответствует двойственная задача вида

$$\begin{gathered}\delta_1^{-\delta_1}\delta_2^{-\delta_2}\delta_3^{-\delta_3}\to \max;\\ \delta_1-\delta_1-\delta_3+\delta_4=0,\\ -\delta_1+\delta_2-\delta_3-\frac{1}{4}\delta_4=0,\\ -\delta_1-\delta_2+\delta_3+\frac{1}{2}\delta_4=0,\\ \delta_1+\delta_2+\delta_3=1;\quad \delta_1,\ \delta_2,\ \delta_3\ \delta_4\geqslant 0.\end{gathered} \qquad (6.5)$$

3 Преобразованная задача геометрического программирования. Вернемся к задаче геометрического программирования и произведем замену переменных:

$$x_j=e^{z_j},\quad j=1,\ 2,\ \dots,\ n.$$

Такая замена правомерна, поскольку верно неравенство $x_j>0$, $j=1,\ 2\ \dots,\ n$. В результате позиномы (6.1) преобразуются в функции

$$h_k(z)=\sum_{i\in J_k} c_i e^{\sum_{j=1}^{n} a_{ij}z_j},\quad k=0,\ 1,\dots,\ p, \qquad (6.6)$$

и задача геометрического программирования станет *преобразованной задачей геометрического программирования:*

$$h_0(z) \to \min, \ h_k(z) \leqslant 1, \quad k=1, 2, \ldots, p,$$

где компоненты вектора z могут принимать любые вещественные значения, а функции h_k определяются равенствами (6.6).

Оптимальные значения задачи геометрического программирования и соответствующей ей преобразованной задачи совпадают (при условии, что они существуют). Преобразованная задача имеет несомненное преимущество перед исходной задачей геометрического программирования: она является задачей выпуклого программирования. Убедимся в этом. Функция e^x — выпуклая, поэтому для $\lambda \in [0, 1]$, $z^{(1)}, z^{(2)} \in \mathbf{R}^n$, имеем

$$e^{\sum\limits_i a_i\left(\lambda z_i^{(1)} + \left(1-\lambda z_i^{(2)}\right)\right)} = e^{\lambda \sum\limits_i a_i z_i^{(1)} + (1-\lambda) \sum\limits_i a_i z_i^{(2)}} \geqslant$$

$$\geqslant \lambda e^{\sum\limits_i a_i z_i^{(1)}} + (1-\lambda) e^{\sum\limits_i a_i z_i^{(2)}}.$$

Следовательно, функция нескольких переменных типа $e^{\sum\limits_i a_i z_i}$ = выпуклая и линейная положительная комбинация подобных функций, в частности все функции $h_k(z)$ вида (6.6), также выпуклые *.

Анализ преобразованной задачи позволяет сделать важный вывод об исходной задаче геометрического программирования. Если $n=m$, то вектор-столбцы матрицы экспонент A образуют базис в пространстве $\mathbf{R}^n$ (при условии, что $\operatorname{rang} A = n$), а значит, любой вектор из $\mathbf{R}^n$ можно получить как линейную комбинацию этих базисных векторов. В частности, найдутся такие числа $z_j^{(l)}$, что

$$-l = \sum_{j=1}^{n} a_{ij} z_j^{(l)}, \quad i=1, 2, \ldots, n; \quad l=1, 2, \ldots .$$

В этом случае из равенств (6.6) следует, что

$$\lim_{l \to \infty} h_k(z^{(l)}) = 0, \ k=0, 1, \ldots, p, \ \text{ где } z^{(l)} = \left(z_1^{(l)}, z_2^{(l)}, \ldots, z_n^{(l)}\right)^{\mathsf{T}}.$$

Таким образом, множество допустимых решений преобразованной задачи не пусто, но решения, реализующего условный минимум функции h_0, не существует. Следовательно, и задача геометрического программирования, в которой $\operatorname{rang} A = n = m$, бессодержательна и всегда можно полагать, что $\operatorname{rang} A = n < m$.

Заметим, что, если $\operatorname{rang} A = n = m$, система уравнений $A^{\mathsf{T}}\delta = 0_n$ в двойственной задаче имеет только нулевое решение $\delta = 0_n$. Но для

---

* На самом деле e^x строго выпуклая, поэтому и функции типа $e^{\sum\limits_i a_i z_i}$ также строго выпуклые (если среди чисел a_i есть отличные от нуля).

такого решения равенство $\sum_{i \in J_0} \delta_i = 1$ не выполняется, а значит, двойственная задача не имеет допустимых решений.

Рассматривая преобразованную задачу, можно также установить, при каких условиях исходная задача геометрического программирования разрешима, т. е. когда ее оптимальное решение и оптимальное значение существуют.

Лемма 6.1. *Если найдется хотя бы одно допустимое решение x^* задачи геометрического программирования и хотя бы одно допустимое решение δ^* соответствующей двойственной задачи такое, что $\delta^* > 0_m$, то задача геометрического программирования разрешима.*

☐ Если существует допустимое решение задачи геометрического программирования, то, очевидно, существует и допустимое решение z^* соответствующей преобразованной задачи. Поэтому задача

$$h_0(z) \to \inf, \quad h_k(z) \leqslant 1, \quad k = 1, 2, \dots p,$$

где все функции h_k имеют вид (6.6), разрешима. Это означает, что найдется минимизирующая последовательность векторов $\{z^{(l)}\}_{l=1}^{\infty}$ такая, что

$$\lim_{l \to \infty} h_0(z^{(l)}) = \gamma, \tag{6.7}$$

$$h_k(z^{(l)}) \leqslant 1, \quad k = 1, 2, \dots, p; \quad l = 1, 2, \dots, \tag{6.8}$$

где γ — точная нижняя грань функции $h_0(z)$ на множестве допустимых решений преобразованной задачи.

Пусть

$$y_i^{(l)} = \sum_{j=1}^{n} a_{ij} z_j^{(l)}, \quad i = 1, 2, \dots, m; \quad l = 1, 2, \dots \tag{6.9}$$

и $y^{(l)} = (y_1^{(l)}, y_2^{(l)}, \dots, y_m^{(l)})^{\mathrm{T}}$, Так как

$$h_k(z^{(l)}) = \sum_{i \in J_k} c_i e^{y_i^{(l)}}, \quad k = 0, 1, \dots, p,$$

то при $l \to \infty$ компоненты вектора $y^{(l)}$ ограничены сверху. В противном случае нарушается равенство (6.7) (получается, что $\gamma = +\infty$) или же окажется невыполненным по крайней мере одно из неравенств (6.8).

Вектор δ^* является допустимым в двойственной задаче; следовательно, $A^{\mathrm{T}}\delta^* = 0_n$. Поэтому, используя (6.9), получаем

$$\sum_{i=1}^{m} \delta_i^* y_i^{(l)} = \sum_{j=1}^{n} \left(\sum_{i=1}^{m} a_{ij} \delta_i^* \right) z_j^{(l)} = 0, \quad l = 1, 2, \dots .$$

Ранее было установлено, что компоненты вектора $y^{(l)}$ ограничены сверху. Принимая это во внимание, из последних равенств в силу неравенств $\delta_i^*>0$, $i=1, 2, \ldots, m$, можно сделать вывод, что указанные компоненты должны быть ограничены и снизу.

На основании замечания, предшествовавшего лемме, можно считать, что $\operatorname{rang} A=n<m$. Не уменьшая общности, предположим, что первые n строк матрицы A являются линейно независимыми. Обозначим квадратную матрицу размера $n\times n$, которую они образуют, через $\tilde{A}$ и введем вектор $\tilde{y}^{(l)}=(y_1^{(l)}, y_2^{(l)},\ldots, y_n^{(l)})^{\mathrm{T}}$. В этих обозначениях первые n равенств из (6.9) можно записать в виде

$$\tilde{y}^{(l)}=\tilde{A}z^{(l)},\ l=1,\ 2,\ldots .$$

Определитель матрицы $\tilde{A}$ отличен от нуля, поэтому для нее существует обратная матрица $\tilde{A}^{-1}$. Следовательно,

$$z^{(l)}=\tilde{A}^{-1}\tilde{y}^{(l)},\ l=1,\ 2,\ldots .$$

Таким образом, каждая компонента вектора $z^{(l)}$ является линейной комбинацией компонент вектора $y^{(l)}$, которые ограничены сверху и снизу. Следовательно, компоненты векторов $z^{(l)}$, $l=1, 2, \ldots$, также ограничены сверху и снизу.

Из ограниченной последовательности $\{z^{(l)}\}$ в силу соответствующей теоремы математического анализа можно выделить сходящуюся подпоследовательность. Без ограничения общности можно считать сходящейся саму эту последовательность: $\lim\limits_{l\to\infty} z^{(l)}=z^{(0)}$. Функции $h_k(z)$, $k=1, 2, \ldots, p$, непрерывны, поэтому множество допустимых решений преобразованной задачи замкнуто. Следовательно, $h_k(z^{(0)})\leqslant 1$, $k=1, 2, \ldots, p$. Кроме того, в силу непрерывности $h_0(z)$ имеем $h_0(z^{(0)})=\gamma$. Таким образом, преобразованная задача имеет оптимальное решение $z^{(0)}$. В этом случае исходная задача геометрического программирования также имеет оптимальное решение $x_i^{(0)}=\mathrm{e}^{z_i^{(0)}}$, $i=1, 2,\ldots, n$. ■

В задаче вида (6.4), например, $x_1^*=x_2^*=x_3^*=1$ — допустимое решение. С помощью метода исключения Гаусса можно определить единственное допустимое решение двойственной задачи (6.5): $\bar{\delta}_1=1/10$, $\bar{\delta}_2=3/5$, $\bar{\delta}_3=3/10$, $\bar{\delta}_4=4/5$. Все эти числа положительные; следовательно, на основании доказанной леммы можно сделать вывод, что исходная задача (6.4) разрешима.

§ 6.2. ОСНОВНАЯ ЛЕММА ГЕОМЕТРИЧЕСКОГО ПРОГРАММИРОВАНИЯ

1. Вспомогательные неравенства. Согласно неравенству

$$(x_1+x_2+\ldots+x_m)/m\geqslant(x_1x_2\ldots x_m)^{1/m}, \tag{6.10}$$

среднее арифметическое любых m положительных чисел не меньше их среднего геометрического, причем они равны тогда и только тогда, когда $x_1=x_2=\ldots=x_m$.

Более общий результат, связывающий среднее арифметическое и среднее геометрическое чисел с «весами», содержит следующая лемма.

Лемма 6.2. *Для любых чисел* $x_1, x_2, \ldots, x_m > 0$ *и любых весов* $\alpha_1, \alpha_2, \ldots, \alpha_m > 0$ *таких, что* $\alpha_1+\alpha_2+\ldots+\alpha_m=1$, *справедливо неравенство*

$$\sum_{k=1}^{m} \alpha_k x_k \geqslant \prod_{k=1}^{m} x_k^{\alpha_k}, \tag{6.11}$$

причем равенство имеет место тогда и только тогда, когда $x_1=x_2=\ldots=x_m$.

□ Функция $y=\ln x$ строго вогнута на своей области определения, поэтому в соответствии с неравенством Иенсена (см. п. 2 § 1.3) для любых $x_1, x_2, \ldots, x_m > 0$, среди которых есть хотя бы одна пара различных, и любых $\alpha_1, \alpha_2, \ldots, \alpha_m > 0$, $\alpha_1+\alpha_2+\ldots+\alpha_m=1$ выполняется неравенство

$$\ln\left(\sum_{k=1}^{m} \alpha_k x_k\right) > \sum_{k=1}^{m} \alpha_k \ln x_k = \ln \prod_{k=1}^{m} x_k^{\alpha_k}.$$

Отсюда, потенцируя, получаем неравенство

$$\sum_{k=1}^{m} \alpha_k x_k > \prod_{x=1}^{m} x_k^{\alpha_k}.$$

Если среди чисел $x_1, x_2, \ldots, x_m$ нет ни одной пары различных, то $x_1=x_2=\ldots=x_m$. В этом случае, поскольку $\sum\limits_{k=1}^{m} \alpha_k = 1$, обе части неравенства (6.11) одинаковы. Обратно: равенство в (6.11) в силу строгой вогнутости функции $\ln x$ имеет место только в случае $x_1=x_2=\ldots=x_m$. ■

Неравенство (6.10) для среднего арифметического и среднего геометрического получается из неравенства (6.11) при $\alpha_1=\alpha_2=\ldots=\alpha_m=1/m$.

Неравенство (6.11) имеет общий вид, однако нам понадобится подобное неравенство в еще более общей форме, когда весовые коэффициенты не нормированы, т. е. не обязательно подчиняются требованию $\alpha_1+\alpha_2+\ldots+\alpha_m=1$.

Лемма 6.3. *Для любых чисел* $\mathfrak{A}_1, \mathfrak{A}_2, \ldots, \mathfrak{A}_m > 0$ *и любых чисел* $\delta_1, \delta_2, \ldots, \delta_m \geqslant 0$ *выполняется неравенство* *

$$(\mathfrak{A}_1+\mathfrak{A}_2+\ldots+\mathfrak{A}_m)^{\lambda} \geqslant \lambda^{\lambda} \prod_{k=1}^{m} \left(\frac{\mathfrak{A}_k}{\delta_k}\right)^{\delta_k}, \tag{6.12}$$

---

* Считаем, что в неравенстве (6.12) $\lambda^{\lambda}=1$ при $\delta_1=\delta_2=\ldots=\delta_m=0$ и $(\mathfrak{A}_k/\delta_k)^{\delta_k}=1$ при $\delta_k=0$.

где $\lambda=\sum_{k=1}^{m}\delta_k$. *Знак равенства возможен только в одном из следующих двух случаев:*

а) $\delta_1=\delta_2=\ldots=\delta_m=0$;

б) $\delta_1, \delta_2, \ldots, \delta_m>0$ *и* $\mathfrak{A}_1/\delta_1=\mathfrak{A}_2/\delta_2=\ldots=\mathfrak{A}_m/\delta_m$.

□ Если $\lambda=0$, то $\delta_1=\delta_2=\ldots=\delta_m=0$ и обе части неравенства (6.12) совпадают и равны единице. Поэтому пусть $\lambda>0$.

Рассмотрим сначала случай, когда все $\delta_1, \delta_2, \ldots, \delta_m>0$. Для нормированных весов $\alpha_k=\delta_k/\lambda$ и чисел $\mathfrak{A}_k/\delta_k$, $k=1, 2, \ldots, m$, воспользуемся неравенством (6.11):

$$\sum_{k=1}^{m}\frac{\delta_k}{\lambda}\frac{\mathfrak{A}_k}{\delta_k}\geqslant\prod_{k=1}^{m}\left(\frac{\mathfrak{A}_k}{\delta_k}\right)^{\frac{\delta_k}{\lambda}}$$

или

$$\sum_{k=1}^{m}\mathfrak{A}_k\geqslant\lambda\prod_{k=1}^{m}\left(\frac{\mathfrak{A}_k}{\delta_k}\right)^{\frac{\delta_k}{\lambda}},$$

откуда, возводя обе части неравенства в степень λ, получаем неравенство (6.12). Знак равенства возможен только в случае $\mathfrak{A}_1/\delta_1=\mathfrak{A}_2/\delta_2=\ldots=\mathfrak{A}_m/\delta_m$.

Теперь допустим, что среди чисел $\delta_1, \delta_2, \ldots, \delta_m$ имеются равные нулю. Не уменьшая общности, можно считать, что $\delta_1, \delta_2, \ldots, \delta_s>0$ и $\delta_{s+1}=\ldots=\delta_m=0$ $(s>1)$. Положим $\sigma=\delta_1+\delta_2+\ldots+\delta_s$. Применяя только что доказанное неравенство, получаем

$$\left(\sum_{k=1}^{s}\mathfrak{A}_k\right)^{\sigma}\geqslant\sigma^{\sigma}\prod_{k=1}^{s}\left(\frac{\mathfrak{A}_k}{\delta_k}\right)^{\delta_k}.$$

Отсюда, учитывая очевидные соотношения

$$\left(\sum_{k=1}^{m}\mathfrak{A}_k\right)^{\lambda}=\left(\sum_{k=1}^{m}\mathfrak{A}_k\right)^{\sigma}>\left(\sum_{k=1}^{s}\mathfrak{A}_k\right)^{\sigma},$$

$$\prod_{k=1}^{m}\left(\frac{\mathfrak{A}_k}{\delta_k}\right)^{\delta_k}=\prod_{k=1}^{s}\left(\frac{\mathfrak{A}_k}{\delta_k}\right)^{\delta_k},\quad \lambda^{\lambda}=\sigma^{\sigma},$$

имеем

$$\left(\sum_{k=1}^{m}\mathfrak{A}_k\right)^{\lambda}>\lambda^{\lambda}\prod_{k=1}^{m}\left(\frac{\mathfrak{A}_k}{\delta_k}\right)^{\delta_k}. \blacksquare$$

2. Основная лемма геометрического программирования.

Лемма 6.4. (основная лемма геометрического программирования). *Если вектор x удовлетворяет ограничениям задачи геометрического программирования, а вектор δ — ограничениям соответствующей двойственной задачи, то*

$$f_0(x) \geqslant v(\delta). \tag{6.13}$$

Более того, при этих же условиях равенство $f_0(x) = v(\delta)$ имеет место тогда и только тогда, когда

$$\delta_i = \begin{cases} \dfrac{c_i x_1^{a_{i1}} x_2^{a_{i2}} \ldots x_n^{a_{in}}}{f_0(x)} \text{ при } i \in J_0, \\ c_i x_1^{a_{i1}} x_2^{a_{i2}} \ldots x_n^{a_{in}} \cdot \sum\limits_{j \in J_k} \delta_j \text{ при } i \in J_k, \ k = 1, 2, \ldots, p. \end{cases} \tag{6.14}$$

☐ Введем обозначение $u_i(x) = c_i x_1^{a_{i1}} x_2^{a_{i2}} \ldots x_n^{a_{in}}$, $i = 1, 2, \ldots, m$, и применим лемму 6.3 к числам $\mathfrak{A}_i = u_i(x) > 0$ и неотрицательным числам δ_i, $i = 1, 2, \ldots, m$, считая, как и в предыдущем параграфе, что $\sum\limits_{i \in J_k} \delta_i = \lambda_k$. Получаем

$$[f_k(x)]^{\lambda_k} = \left[\sum_{i \in J_k} u_i(x)\right]^{\lambda_k} \geqslant \lambda_k^{\lambda_k} \prod_{i \in J_k} \left(\frac{u_i(x)}{\delta_i}\right)^{\delta_i}, \ k = 0, 1, \ldots, p. \tag{6.15}$$

Отсюда, в частности, при $k = 0$ имеем

$$f_0(x) \geqslant \prod_{i \in J_0} \left(\frac{u_i(x)}{\delta_i}\right)^{\delta_i}, \tag{6.16}$$

так как $\lambda_0 = \sum\limits_{i \in J_0} \delta_i = 1$.

Согласно условию, верно неравенство $1 \geqslant f_k(x)$, $k = 1, 2, \ldots, p$, поэтому из соотношений (6.15) вытекают неравенства

$$1 \geqslant \lambda_k^{\lambda_k} \prod_{i \in J_k} \left(\frac{u_i(x)}{\delta_i}\right)^{\delta_i}, \ k = 1, 2, \ldots, p. \tag{6.17}$$

Умножая почленно неравенство (6.16) на произведение неравенств (6.17), получаем

$$f_0(x) \geqslant \prod_{i=1}^{m} \left(\frac{u_i(x)}{\delta_i}\right)^{\delta_i} \prod_{k=1}^{p} \lambda_k^{\lambda_k}. \tag{6.18}$$

Но

$$\left(\frac{u_i(x)}{\delta_i}\right)^{\delta_i} = \left(\frac{c_i}{\delta_i}\right)^{\delta_i} x_1^{a_{i1}\delta_i} x_2^{a_{i2}\delta_i} \ldots x_n^{a_{in}\delta_i};$$

следовательно,

$$f_0(x) \geqslant v(\delta)\, x_1^{\sum\limits_{i=1}^{m} a_{i1}\delta_i} \cdot \ldots \cdot x_n^{\sum\limits_{i=1}^{m} a_{in}\delta_i}.$$

По условию, вектор $\delta = (\delta_1, \delta_2, \ldots, \delta_m)^{\mathrm{T}}$ является допустимым в двойственной задаче, а значит, удовлетворяет равенствам $\sum\limits_{i=1}^{m} a_{ij}\delta_i = 0$, $j = 1, 2, \ldots, n$. В соответствии с этим последнее неравенство принимает требуемый вид $f_0(x) \geqslant v(\delta)$.

Докажем вторую часть леммы. Пусть $f_0(x) = v(\delta)$. В этом случае неравенство (6.16) и все неравенства (6.17) выполняются как равенства (если это не так, умножая почленно (6.16) на (6.17), мы получили бы неравенство $f_0(x) > v(\delta)$).

В силу леммы 6.3 равенство в (6.16) может иметь место только при

$$\frac{u_1(x)}{\delta_1} = \frac{u_2(x)}{\delta_2} = \ldots = \frac{u_{m_0}(x)}{\delta_{m_0}}.$$

Отсюда, согласно известному свойству пропорций, имеем

$$\frac{u_i(x)}{\delta_i} = \frac{\sum\limits_{i \in J_0} u_i(x)}{\sum\limits_{i \in J_0} \delta_i} = \frac{f_0(x)}{1} = f_0(x)$$

при каждом $i \in J_0$. Следовательно, $\delta_i = u_i(x)/f_0(x)$ при $i \in J_0$, т. е. справедливость первых m_0 равенств в (6.14) проверена.

Установим справедливость остальных равенств. В силу леммы 6.3 равенства в (6.17) возможны лишь в одном из следующих двух случаев: а) $\delta_i = 0$ для всех $i \in J_k$, $k = 1, 2, \ldots, p$; б) $\delta_i > 0$ для всех $i \in J_k$ и $u_i(x)/\delta_i = u_j(x)/\delta_j$ при всех $i, j \in J_k$, $k = 1, 2, \ldots, p$. В первом случае $\sum\limits_{j \in J_k} \delta_j = 0$ и доказываемые равенства в (6.14) при каждом $k = 1, 2, \ldots, p$ превращаются в очевидные: $0 = 0$. Во втором случае, согласно свойству пропорций, получаем

$$\frac{u_i(x)}{\delta_i} = \frac{\sum\limits_{i \in J_k} u_i(x)}{\sum\limits_{i \in J_k} \delta_i} = \frac{f_k(x)}{\lambda_k}, \quad k = 1, 2, \ldots, p, \tag{6.19}$$

при каждом $i \in J_k$. Отсюда

$$\delta_i = \frac{u_i(x)\lambda_k}{f_k(x)} \text{ при } i \in J_k, \ k = 1, 2, \ldots, p. \tag{6.20}$$

В (6.17) неравенства выполняются как равенства, поэтому, подставляя в эти равенства $u_i(x)/\delta_i$ вида (6.19), получаем

$$1=\lambda_k^{\lambda_k}\prod_{i\in J_k}\left(\frac{f_k(x)}{\lambda_k}\right)^{\delta_i}=[f_k(x)]^{\lambda_k},$$

откуда $f_k(x)=1$, $k=1, 2, ..., p$. Следовательно, равенства (6.20) превращаются в доказываемые равенства из (6.14).

Остается показать, что для допустимого решения x задачи геометрического программирования и допустимого решения δ двойственной задачи, связанных равенствами (6.14), выполняется равенство $f_0(x)=v(\delta)$. Из равенств (6.14), согласно лемме 6,3, следует, что неравенство (6.16) должно выполняться как равенство. Более того, все неравенства (6.17) также должны быть равенствами. В самом деле, если $\lambda_k=0$, то справедливо $\delta_i=0$ для всех $i\in J_k$ и, следовательно, правая часть неравенства (6.17) при этом k, так же как и левая, равна единице. Если же $\lambda_k>0$, то из (6.14) получаем равенство $f_k(x)=1$ при этом k и, кроме того, имеем

$$u_i(x)/\delta_i=\lambda_k \text{ при всех } i\in J_k.$$

На основании этого в соответствии с леммой 6.3 делаем вывод, что для этого k неравенство в (6.17) выполняется как равенство. Таким образом, неравенство (6.16) и все неравенства в (6.17) выполняются как равенства. Перемножая почленно все эти равенства, придем к требуемому равенству $f_0(x)=v(\delta)$. ■

Из доказанной леммы следует, что если задача геометрического программирования и двойственная ей задача разрешимы, то оптимальное значение первой задачи не меньше, чем оптимальное значение второй. Как будет показано в следующем параграфе, на самом деле при довольно общих условиях оптимальные значения обеих задач совпадают.

З а м е ч а н и е. Если выполняется равенство $f_0(\bar{x})=v(\bar{\delta})$ при некоторых векторах $\bar{x}$ и $\bar{\delta}$ — допустимых решениях задачи геометрического программирования и соответствующей ей двойственной задачи, то на основании последней леммы $f_0(\bar{x})=v(\bar{\delta})\leqslant f_0(x)$ для любого вектора x, являющегося допустимым решением задачи геометрического программирования, т. е. $\bar{x}$ — оптимальное решение задачи геометрического программирования. Кроме того, согласно той же лемме, справедливо $v(\bar{\delta})=f_0(\bar{x})\geqslant v(\delta)$ для любого вектора δ, являющегося допустимым решением двойственной задачи, а значит, $\bar{\delta}$ — оптимальное решение двойственной задачи геометрического программирования. Таким образом, *равенство $f_0(\bar{x})=v(\bar{\delta})$ является достаточным условием оптимальности решений $\bar{x}$ и $\bar{\delta}$.*

§ 6.3. ТЕОРЕМА ДВОЙСТВЕННОСТИ ГЕОМЕТРИЧЕСКОГО ПРОГРАММИРОВАНИЯ

1. Теорема двойственности. Теорема двойственности, доказываемая на основе основной леммы геометрического программирования и необходимых условий оптимальности выпуклого программирования, представляет собой основной результат главы. Она показывает, что при довольно общих предположениях, решая двойственную задачу, можно определить оптимальное решение и оптимальное значение исходной задачи геометрического программирования.

Теорема 6.1. *Предположим, что ограничения задачи геометрического программирования удовлетворяют условию регулярности Слейтера: найдется такая точка* $\tilde{x} > 0_n$, *что выполнено неравенство* $f_k(\tilde{x}) < 1$, $k = 1, 2, \ldots, p$. *Справедливы следующие два утверждения:*

если существует оптимальное решение $\bar{x}$ *задачи геометрического программирования (в силу леммы 6.1 такое решение заведомо существует, когда соответствующая ей двойственная задача имеет по крайней мере одно допустимое решение с положительными компонентами), то существует оптимальное решение* $\bar{\delta}$ *двойственной задачи, причем оптимальные значения обеих задач совпадают:* $f_0(\bar{x}) = v(\bar{\delta})$;

пусть $\bar{x}$ *и* $\bar{\delta}$ — *допустимые решения задачи геометрического программирования и соответствующей двойственной задачи. Для того чтобы эти решения были оптимальными, необходимо и достаточно, чтобы они были связаны равенствами*

$$c_i \bar{x}_1^{a_{i1}} \bar{x}_2^{a_{i2}} \ldots \bar{x}_n^{a_{in}} = \begin{cases} \bar{\delta}_i v(\bar{\delta}) & \text{при } i \in J_0, \\ \bar{\delta}_i / \bar{\lambda}_k & \text{при } i \in J_k,\ k \in K, \end{cases} \tag{6.21}$$

где $\bar{\lambda}_k = \sum_{i \in J_k} \bar{\delta}_i$, *а* K — *подмножество множества номеров* $\{1, 2, \ldots, m\}$, *для которых* $\bar{\lambda}_k > 0$.

□ Пусть $\bar{x}$ — точка минимума в задаче геометрического программирования. От задачи геометрического программирования с помощью замены $x_j = e^{z_j}$, $j = 1, 2, \ldots, n$, переходим к преобразованной задаче:

$$h_0(z) \to \min;\ h_k(z) \leqslant 1,\ k = 1,\ 2, \ldots, p.$$

При этом, очевидно, вектор $\bar{z}$ с компонентами

$$\bar{z}_j = \ln \bar{x}_j,\ j = 1,\ 2, \ldots, n, \tag{6.22}$$

является оптимальным решением в преобразованной задаче. Кроме того, ограничения преобразованной задачи удовлетворяют условию регулярности Слейтера: $h_k(\tilde{z}) < 1$, $k = 1, 2, \ldots, p$, где $\tilde{z}$ — вектор с компонентами $\tilde{z}_j = \ln \tilde{x}_j$, $j = 1, 2, \ldots, n$.

В § 6.1 отмечалось, что преобразованная задача является задачей выпуклого программирования. Поэтому для точки $\bar{z}$ можно применить необходимое условие оптимальности теоремы 2.9. Согласно этой теореме, найдутся такие неотрицательные числа $\bar{\mu}_1$, $\bar{\mu}_2$, ..., $\bar{\mu}_p$, что пара векторов $(\bar{x}, \bar{\mu})$, где $\bar{\mu}=(\bar{\mu}_1, \bar{\mu}_2, ..., \bar{\mu}_p)^{\text{т}}$, образует седловую точку функции Лагранжа

$$L(z, \mu)=h_0(z)+\sum_{k=1}^{p}\mu_k(h_k(z)-1),$$

т. е. неравенства

$$L(\bar{z}, \mu)\leqslant L(\bar{z}, \bar{\mu})\leqslant L(z, \bar{\mu}) \tag{6.23}$$

выполняются для всех $z\in\mathbf{R}^n$ и всех векторов μ с неотрицательными компонентами.

Правая часть неравенства (6.29) означает, что функция $L(z, \bar{\mu})$ достигает безусловного минимума в точке $\bar{z}$. Эта функция дифференцируема, поэтому, согласно теореме 2.4, справедливо равенство $\nabla_z L(z, \bar{\mu})\big|_{z=\bar{z}}=0_n$, которое в покомпонентной форме имеет вид

$$\sum_{i\in J_0}a_{is}c_i e^{\sum\limits_{j=1}^{n}a_{ij}\bar{z}_j}+\sum_{k=1}^{p}\bar{\mu}_k\left(\sum_{i\in J_k}a_{is}c_i e^{\sum\limits_{j=1}^{n}a_{ij}\bar{z}_j}\right)=0$$

для любого $s=1, 2, ..., n$. Если это равенство разделить на $h_0(\bar{z})$ и ввести m-мерный вектор $\bar{\delta}$ с неотрицательными компонентами вида

$$\bar{\delta}_i=\begin{cases}\dfrac{c_i e^{\sum\limits_{j=1}^{n}a_{ij}\bar{z}_j}}{h_0(\bar{z})} \text{ при } i\in J_0,\\[2ex] \dfrac{\bar{\mu}_k c_i e^{\sum\limits_{j=1}^{n}a_{ij}\bar{z}_j}}{h_0(\bar{z})} \text{ при } i\in J_k,\ k=1, 2, ..., p,\end{cases} \tag{6.24}$$

то придем к равенству

$$\sum_{i=1}^{m}a_{is}\bar{\delta}_i=0 \text{ для любого } s=1, 2, ..., n.$$

Из (6.24) следует, что равенство $\bar{\delta}_1+\bar{\delta}_2+...+\bar{\delta}_{m_0}=1$ также выполняется, поскольку

$$\sum_{i\in J_0}c_i e^{\sum\limits_{j=1}^{n}a_{ij}\bar{z}_j}=h_0(\bar{z}).$$

Следовательно, введенный вектор $\bar{\delta}$ удовлетворяет всем ограничениям двойственной задачи.

Если будет установлено равенство

$$f_0(\bar{x})=v(\bar{\delta}), \tag{6.25}$$

то, согласно замечанию к основной лемме геометрического программирования, $\bar{\delta}$ — оптимальное решение двойственной задачи и тем самым первое утверждение теоремы будет доказано полностью. Проверим справедливость равенства (6.25). Суммируя по $i\in J_k$ второе равенство в (6.24), получаем

$$\bar{\lambda}_k=\sum_{i\in J_k}\bar{\delta}_i=\frac{\bar{\mu}_k h_k(\bar{z})}{h_0(\bar{z})},\ k=1,\ 2,\dots,\ p. \tag{6.26}$$

Теперь докажем, что

$$\bar{\mu}_k h_k(\bar{z})=\bar{\mu}_k,\ k=1,\ 2,\dots,\ p. \tag{6.27}$$

Из левой части неравенства (6.23) при $\mu=0_p$ имеем

$$L(\bar{z},\ 0_p)=h_0(\bar{z})\leqslant L(\bar{z},\ \bar{\mu}).$$

С другой стороны, поскольку $\bar{\mu}_k\geqslant 0$ и $h_k(\bar{z})\leqslant 1$, справедливо неравенство

$$\bar{\mu}_k(h_k(\bar{z})-1)\leqslant 0,\ k=1,\ 2,\dots,\ p, \tag{6.28}$$

и поэтому

$$h_0(\bar{z})\geqslant L(\bar{z},\ \bar{\mu})=h_0(\bar{z})+\sum_{k=1}^{p}\bar{\mu}_k(h_k(\bar{z})-1).$$

Таким образом, выполнено равенство $h_0(\bar{z})=L(\bar{z},\ \bar{\mu})$, а значит,

$$\sum_{k=1}^{p}\bar{\mu}_k(h_k(\bar{z})-1)=0.$$

Отсюда в силу (6.28) вытекают равенства (6.27). Теперь, учитывая (6.27), равенства (6.26) можно записать так:

$$\bar{\lambda}_k=\bar{\mu}_k/h_0(\bar{z}),\ k=1,\ 2,\dots,\ p.$$

В соответствии с этим равенства (6.24) после выполнения замены (6.22) принимают вид

$$\bar{\delta}_i=\begin{cases} c_i\bar{x}_1^{a_{i1}}\bar{x}_2^{a_{i2}}\dots\bar{x}_n^{a_{in}}/f_0(\bar{x}) & \text{при } i\in J_0,\\ c_i\bar{x}_1^{a_{i1}}\bar{x}_2^{a_{i2}}\dots\bar{x}_n^{a_{in}}\cdot\sum\limits_{i\in J_k}\bar{\delta}_i & \text{при } i\in J_k,\ k=1,\ 2,\dots,\ p.\end{cases}$$

В этом случае основная лемма геометрического программирования гарантирует выполнение равенства (6.25). Первое утверждение теоремы доказано.

Проверим справедливость второго утверждения. Необходимость. Пусть векторы $\bar{x}$ и $\bar{\delta}$ — оптимальные решения задачи геометрического программирования и двойственной задачи геометрического программирования соответственно. В силу доказанного первого утверждения для данного вектора $\bar{x}$ можно указать оптимальное решение δ^* двойственной задачи, причем будет выполняться равенство $f_0(\bar{x})=v(\delta^*)$. Очевидно, $v(\bar{\delta})=v(\delta^*)$, поэтому верно равенство $f_0(\bar{x})=v(\bar{\delta})$. Применяя вторую часть основной леммы геометрического программирования и используя последнее равенство, приходим к требуемым равенствам (6.21).

Достаточность. Суммируя первые равенства (6.21) по $i\in J_0$, получаем $f_0(\bar{x})=v(\bar{\delta})$, откуда в силу замечания к основной лемме геометрического программирования следует оптимальность решений $\bar{x}$ и $\bar{\delta}$. ■

Если оптимальное решение $\bar{\delta}$ двойственной задачи найдено, то с помощью равенства (6.21) можно определить оптимальное решение исходной задачи геометрического программирования. Для этого обычно равенства (6.21) логарифмируют, в результате чего они превращаются в систему линейных уравнений относительно логарифмов неизвестных, т. е. относительно $\ln\bar{x}_1, \ln\bar{x}_2, \ldots, \ln\bar{x}_n$.

Таким образом, как показывает теорема 6.1, найдя оптимальное значение двойственной задачи, при достаточно общих условиях можно быть уверенным в том, что тем самым найдено оптимальное значение исходной задачи геометрического программирования. Для того чтобы по оптимальному решению двойственной задачи определить оптимальное решение исходной задачи геометрического программирования, нужно решить указанную выше систему линейных уравнений.

2. Пример использования теоремы двойственности. Вернемся к задаче (6.4) из § 6.1. При $\tilde{x}_1=1$, $\tilde{x}_2=2$, $\tilde{x}_3=1$ ограничение в этой задаче выполняется как строгое неравенство, т. е. условие Слейтера теоремы 6.1 выполнено. В конце § 6.1 было установлено, что соответствующая двойственная задача имеет единственное допустимое решение $\delta=(1/10,\ 3/5,\ 3/10,\ 4/5)^{\mathrm{T}}$, компоненты которого положительны. Согласно первому утверждению теоремы 6.1, оптимальное решение этой задачи геометрического программирования, а также оптимальное решение двойственной задачи существуют. Поскольку $\bar{\delta}$ — единственное допустимое решение двойственной задачи, оно является и оптимальным. Тогда оптимальное значение двойственной задачи равно

$$v(\bar{\delta})=v_0=10^{1/10}(5/3)^{3/5}(10/3)^{3/10}.$$

Таким же является и оптимальное значение исходной задачи геометрического программирования (согласно первому утверждению теоремы 6.1). Найдем оптимальное решение задачи (6.4), которое обозначим через $\bar{x} = (\bar{x}_1, \bar{x}_2, \bar{x}_3)^T$. Для этого запишем систему равенств (6.21) применительно к данному примеру:

$$\frac{\bar{x}_1}{\bar{x}_2 \bar{x}_3} = \frac{v_0}{10}, \quad \frac{\bar{x}_2}{\bar{x}_1 \bar{x}_3} = \frac{3v_0}{5}, \quad \frac{\bar{x}_3}{\bar{x}_1 \bar{x}_2} = \frac{3v_0}{10}, \quad \frac{\bar{x}_1 \sqrt{\bar{x}_3}}{\sqrt[4]{\bar{x}_2}} = 1$$

и прологарифмируем их:

$$\begin{cases} \ln \bar{x}_1 - \ln \bar{x}_2 - \ln \bar{x}_3 = \ln \dfrac{v_0}{10}, \\ -\ln \bar{x}_1 + \ln \bar{x}_2 - \ln \bar{x}_3 = \ln \dfrac{3v_0}{5}, \\ -\ln \bar{x}_1 - \ln \bar{x}_2 + \ln \bar{x}_3 = \ln \dfrac{3v_0}{10}, \\ \ln \bar{x}_1 - \dfrac{1}{4} \ln \bar{x}_2 + \dfrac{1}{2} \ln \bar{x}_3 = 0. \end{cases}$$

Получена система линейных уравнений относительно неизвестных $\ln \bar{x}_1$, $\ln \bar{x}_2$, $\ln \bar{x}_3$. Из первых трех уравнений находим единственное решение:

$$\ln \bar{x}_1 = \ln \frac{5\sqrt{2}}{3v_0}, \quad \ln \bar{x}_2 = \ln \frac{10\sqrt{3}}{3v_0}, \quad \ln \bar{x}_3 = \ln \frac{5}{v_0} \sqrt{\frac{2}{3}}.$$

Нетрудно проверить, что при этих значениях четвертое из уравнений также удовлетворяется. Следовательно, оптимальное решение задачи (6.4) имеет вид

$$\bar{x}_1 = \frac{5\sqrt{2}}{3v_0}, \quad \bar{x}_2 = \frac{10\sqrt{3}}{3v_0}, \quad \bar{x}_3 = \frac{5}{v_0} \sqrt{\frac{2}{3}}.$$

В данном примере решение задачи оказалось несложным и свелось к решению системы линейных уравнений. Анализ примера показывает, что *в тех случаях, когда $m = n+1$ (при условии выполнения равенства* rang $A = n$), *двойственная задача всегда имеет не более одного допустимого решения, а значит, оно является оптимальным* и процесс решения двойственной задачи (тем самым и процесс решения исходной задачи геометрического программирования) сводится к решению соответствующих систем линейных уравнений.

3. Схема решения двойственной задачи геометрического программирования. Число $m = n - 1$ (при условии выполнения равенства rang $A = n$) называют *степенью трудности* задачи геометрического программирования. В случае нулевой степени трудности процесс решения двойственной задачи сводится, как указано выше, к решению системы линейных уравнений. Если степень трудности ненулевая (т. е. $m > n+1$), то совместная система ограничений двойственной задачи имеет бесчисленное множество решений и сразу записать оптимальное решение двойственной задачи невозможно.

В таком случае поступают следующим образом. Вместо двойственной задачи решают задачу вида

$$\ln v(\delta) \to \max;$$

$$A^{\mathrm{T}}\delta = 0_m, \ \delta_1 + \delta_2 + \ldots + \delta_{m_0} = 1, \ \delta \geqq 0_m, \tag{6.29}$$

которая отличается от двойственной задачи видом целевой функции. Она эквивалентна двойственной задаче в том смысле, что каждое ее оптимальное решение является оптимальным решением двойственной задачи и наоборот. С вычислительной точки зрения преимущество задачи (6.29) в сравнении с двойственной задачей раскрывает следующее утверждение.

Лемма 6.5. *Функция* $\ln v(\delta)$ *является вогнутой при* $\delta > \theta_m$.

Доказательство этой леммы можно найти в [9].

Итак, задача (6.20) является задачей вогнутого программирования, и поэтому ее решение в смысле вычислений проще, чем решение двойственной задачи, целевая функция которой не обладает свойством вогнутости.

Иногда удается, не решая двойственную задачу, получить довольно точное представление о ее оптимальном значении.

Пример [10]. Пусть нужно минимизировать позином

$$f_0(x) = \frac{40}{x_1 \sqrt{x_2} x_3} + 20x_1x_3 + 20x_1x_2x_3$$

при условии

$$\frac{1}{3x_1^2x_2^2} + \frac{4\sqrt{x_2}}{3x_3} \leqslant 1.$$

Это ограничение удовлетворяет условию Слейтера. Ранг матрицы экспонент

$$A = \begin{pmatrix} -1 & -1/2 & -1 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \\ -2 & -2 & 0 \\ 0 & 1/2 & -1 \end{pmatrix}$$

равен 3, поэтому трудность данной задачи равна $5-3-1=1$. Запишем двойственную задачу:

$$v(\delta) = \left(\frac{40}{\delta_1}\right)^{\delta_1} \left(\frac{20}{\delta_2}\right)^{\delta_2} \left(\frac{20}{\delta_3}\right)^{\delta_3} \left(\frac{1}{3\delta_4}\right)^{\delta_4} \left(\frac{4}{3\delta_5}\right)^{\delta_5} (\delta_4 + \delta_5)^{\delta_4 + \delta_5} \to \max;$$

$$-\delta_1 + \delta_2 + \delta_3 - 2\delta_4 = 0, \ -\frac{1}{2}\delta_1 + \delta_3 - 2\delta_4 + \frac{1}{2}\delta_5 = 0,$$

$$-\delta_1 + \delta_2 + \delta_3 - \delta_5 = 0, \ \delta_1 + \delta_2 + \delta_3 = 1, \ \delta_1, \delta_2, \ldots, \delta_5 \geqslant 0.$$

Используя ограничения-равенства, любые четыре двойственные переменные можно выразить через оставшуюся переменную*. Выразим двойственные переменные, например, через δ_2:

$$\delta_1=1-\delta_2,\ \delta_3=\delta_2,\ \delta_4=-\frac{1}{2}+2\delta_2,\ \delta_5=-1+4\delta_2. \qquad (6.30)$$

Эти равенства, в частности, показывают, что в данном примере двойственная задача — это задача, в которой независимой является только одна переменная δ_2, т. е. *число независимых переменных двойственной задачи равно числу степени трудности* (последнее утверждение имеет место и в общем случае, когда степень трудности больше 1).

Все двойственные переменные должны быть неотрицательными, поэтому переменная δ_2 может изменяться в пределах $1/4\leqslant\delta_2\leqslant 1/2$. При $\delta_2=3/8$ значения всех двойственных переменных положительны, значит, оптимальные значения $f_0(\bar{x})$ задачи геометрического программирования и $v(\bar{\delta})$ двойственной задачи существуют и равны: $f_0(\bar{x})=v(\bar{\delta})$. Найдем теперь верхнюю и нижнюю оценки этих значений. Для этого рассмотрим несколько допустимых (пробных) решений задачи геометрического программирования, например $x^{(1)}=(1,\,1,\,4)^{\text{т}}$, $x^{(2)}=(1,\,1,\,3)^{\text{т}}$, $x^{3}=(1,\,1,\,5/2)^{\text{т}}$, $x^{(4)}=(1,\,1,\,17/18)^{\text{т}}$, и вычислим соответствующие значения целевой функции: $f_0(x^{(1)})=170$; $f_0(x^{(2)})=133{,}3$; $f_0(x^{(3)})=116$; $f_0(x^{(4)})=103{,}82$. Очевидно, оптимальное значение исходной задачи не превосходит наименьшего из полученных значений, т. е. $f_0(\bar{x})\leqslant 103{,}82$. Далее, поскольку $1/4\leqslant\delta_2\leqslant 1/2$, возьмем несколько пробных значений для δ_2, например $\delta_2^{(1)}=1/4$, $\delta_2^{(2)}=3/10$, $\delta_2^{(3)}=7/20$. Согласно (6.29), получаем следующие три допустимых решения двойственной задачи: $\delta^{(1)}=(1/2,\,1/4,\,1/4,\,0,\,0)^{\text{т}}$; $\delta^{(2)}=(2/5,\,3/10,\,3/10,\,1/10,\,1/5)^{\text{т}}$, $\delta^{(3)}=(3/10,\,7/20,\,7/20,\,1/5,\,2/5)^{\text{т}}$. Вычисляя значения функции $v(\delta)$ в этих точках, найдем: $v(\delta^{(1)})=80$, $v(\delta^{(2)})=90{,}9$; $v(\delta^{(3)})=97{,}14$. Оптимальное значение двойственной задачи не может быть меньше наибольшего из этих значений, т. е. $v(\bar{\delta})\geqslant 97{,}14$.

Окончательно приходим к оценкам $97{,}14\leqslant v(\bar{\delta})=f_0(\bar{x})\leqslant 103{,}82$, которые позволяют сделать вывод о том, что оптимальное значение $f_0(\bar{x})$ равно 100 с погрешностью не более чем **4**%.

§ 6.4. ПРИКЛАДНЫЕ ЗАДАЧИ

1. Задачи выбора оптимальных размеров контейнера. Для перевозки строительного материала, например гравия, требуется изготовить открытый контейнер. Объем гравия, подлежащего транспортировке, равен V, а стоимость одной перевозки составляет C. Кроме того, заданы a, b и c — соответственно стоимость единицы площади днища, боковой и торцовой поверхности. Контейнер большого

* В общем случае, если степень трудности задачи геометрического программирования равна $l\,(l\geqslant 1)$, из ограничений-равенств двойственной задачи $m-l$ двойственных переменных можно выразить через оставшиеся l переменных.

размера позволяет уменьшить число рейсов, значит, и расходы по перевозке. При этом, однако, велика стоимость самого контейнера. Контейнер небольшого размера стоит недорого, но возрастают расходы на транспортировку. Следует определить оптимальные размеры контейнера x, y, z (длину, ширину, высоту), при которых суммарные затраты на транспортировку и изготовление контейнера были бы минимальными.

Число рейсов равно $V/(xyz)$, а стоимость всех перевозок — $CV/(xyz)$. Стоимость днища, боковой и торцовой поверхности составляет соответственно axy, $2bxz$, $2cyz$. Следовательно, суммарная стоимость (целевая функция), подлежащая минимизации, имеет вид

$$f_0(x, y, z) = CV/(xyz) + axy + 2bxz + 2cyz.$$

По условию, имеются только естественные ограничения $x>0$, $y>0$, $z>0$. Запишем матрицу экспонент для сформулированной задачи:

$$A = \begin{pmatrix} -1 & -1 & -1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}.$$

Ее ранг равен 3. Степень трудности задачи нулевая, так как $4-3-1=0$. Запишем двойственную задачу:

$$(CV/\delta_1)^{\delta_1}(a/\delta_2)^{\delta_2}(2b/\delta_3)^{\delta_3}(2c/\delta_4)^{\delta_4} \to \max;$$

$$-\delta_1+\delta_2+\delta_3=0, \qquad -\delta_1+\delta_3+\delta_4=0,$$

$$-\delta_1+\delta_2+\delta_4=0, \qquad \delta_1+\delta_2+\delta_3+\delta_4=1.$$

Система ограничений двойственной задачи имеет единственное решение $\bar{\delta}_1=2/5$, $\bar{\delta}_2=\bar{\delta}_3=\bar{\delta}_4=1/5$, являющееся и оптимальным решением этой задачи. Оптимальное значение двойственной задачи совпадает с оптимальным значением исходной задачи, т. е.

$$\min f_0(x, y, z) = \left(\frac{CV}{2/5}\right)^{2/5}\left(\frac{a}{1/5}\right)^{1/5}\left(\frac{2b}{1/5}\right)^{1/5}\left(\frac{2c}{1/5}\right)^{1/5} = 5(C^2V^2abc)^{1/5}.$$

Первое равенство из (6.21) можно записать в виде

$$c_i\bar{x}_1^{a_{i1}}\bar{x}_2^{a_{i2}}\dots\bar{x}_n^{a_{in}}/\bar{\delta}_i = v(\bar{\delta}) \text{ при } i \in J_0.$$

Отсюда получаем

$$\frac{5CV}{2\bar{x}\bar{y}\bar{z}} = 5a\bar{x}\bar{y} = 10b\bar{x}\bar{z} = 10c\bar{y}\bar{z} = 5(C^2V^2abc)^{1/5}.$$

Следовательно, оптимальное решение исходной задачи имеет вид

$$\bar{x} = (CVc^3/(a^2b^2))^{1/5}, \quad \bar{y} = (CVb^3/(a^2c^2))^{1/5}, \quad \bar{z} = \frac{1}{2}(CVa^3/(b^2c^2))^{1/5}.$$

На основе полученного решения нетрудно подсчитать, что 40% общей стоимости составляет стоимость перевозок, а 60% приходится на стоимость контейнера.

Рассмотрим теперь другой вариант задачи о контейнерах. Пусть продукция предприятия перевозится в закрытых контейнерах, которые на предприятие не возвращаются. Ежемесячно отправляется 1000 м³ продукции в контейнерах, длина которых x, ширина y и высота z (м). Днище и торцовые поверхности изготовляются из отходов и на их изготовление не требуется затрат. Однако ежемесячно может быть использовано не более 10 м² отходов. Стоимость квадратного метра боковой поверхности составляет 0,2 руб., а квадратного метра крышки — 0,1 руб. Стоимость транспортировки одного контейнера составляет 0,5 руб. Найти оптимальные размеры контейнера, при которых суммарные затраты минимальны.

Суммарная стоимость выражается функцией

$$f_0(x,\ y,\ z)=0{,}5\ \frac{1000}{xyz}+0{,}4xz\frac{1000}{xyz}+0{,}1xy\frac{1000}{xyz}=\frac{500}{xyz}+\frac{400}{y}+\frac{100}{z},$$

а ограничение на использование материала при изготовлении днища и торцовых поверхностей имеет вид $xy+2yz\leqslant 10$. Итак, получаем следующую задачу геометрического программирования:

$$500x^{-1}y^{-1}z^{-1}+400y^{-1}+100z^{-1}\to\min;\ 0{,}1xy+0{,}2yz\leqslant 1;$$
$$x,\ y,\ z>0.$$

Степень трудности задачи равна 5—3—1 = 1. Составляем соответствующую двойственную задачу:

$$v(\delta)=(500/\delta_1)^{\delta_1}(400/\delta_2)^{\delta_2}(100/\delta_3)^{\delta_3}(0{,}1/\delta_4)^{\delta_4}(0{,}2/\delta_5)^{\delta_5}(\delta_4+\delta_5)\to\max;$$
$$-\delta_1+\delta_4=0,\ \ -\delta_1-\delta_2+\delta_4+\delta_5=0,$$
$$-\delta_1-\delta_3+\delta_5=0,\ \delta_1+\delta_2+\delta_3=1,\ \delta_1,\ \delta_2,\dots,\delta_5\geqslant 0.$$

Из ограничений-равенств находим

$$\delta_2=1/2,\ \delta_3=1/2-\delta_1,\ \delta_4=\delta_1,\ \delta_5=1/2. \tag{6.31}$$

В соответствии с этим получаем

$$v(\delta)=v(\delta_1)=(320)^{1/2}\left(\frac{50}{\delta_1^2}\right)^{\delta_1}\left(\frac{100}{1/2-\delta_1}\right)^{1/2-\delta_1}(1/2+\delta_1)^{1/2+\delta_1}.$$

Логарифмируя эту функцию, а затем дифференцируя и приравнивая производную нулю, найдем значение $\overline{\delta_1}=1/\sqrt{12}\approx 0{,}29$, доставляющее максимум функции $v(\delta_1)$. Из (6.31) находим оптимальные значения остальных двойственных переменных $\overline{\delta_2}=\overline{\delta_5}=0{,}5$, $\overline{\delta_3}\approx 0{,}21$, $\overline{\delta_4}\approx 0{,}29$. Теперь на основании равенств (6.21) можно определить оптимальное решение: $\overline{x}\approx 1{,}56$ м, $\overline{y}\approx 2{,}31$ м, $\overline{z}\approx 1{,}38$ м.

2. Проектирование гидравлического цилиндра минимального диаметра. На рис. 6.1 изображена схема гидравлического цилиндра, предназначенного для создания значительных усилий, например в подъемниках. Подъемная сила f и толщина t стенок цилиндра не должны быть меньше некоторых заданных значений F и T, т. е. $f \geqslant F$, $t \geqslant T$. Давление жидкости p и напряжение в стенках s не должны превышать заданных значений: $p \leqslant P$, $s \leqslant S$. Сила, давление и внутренний диаметр цилиндра связаны равенством $f = \pi d^2 p/4$, а напряжение в стенках определяется формулой $s = dp/(2t)$. При указанных ограничениях требуется спроектировать гидравлический цилиндр с наименьшим возможным диаметром $d+2t$.

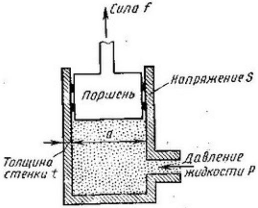

Рис. 6.1

Записываем задачу геометрического программирования:

$$d + 2t \to \min;$$

$$\frac{4F}{\pi d^2 p} \leqslant 1, \ \frac{dp}{2St} \leqslant 1, \ \frac{p}{P} \leqslant 1, \ \frac{T}{t} \leqslant 1, \tag{6.32}$$

где первые два неравенства соответствуют требованиям $f \geqslant F$ и $s \leqslant S$. Переменные d, p и t считаем положительными. Соответствующая двойственная задача имеет вид

$$v(\delta) = \left(\frac{1}{\delta_1}\right)^{\delta_1} \left(\frac{2}{\delta_2}\right)^{\delta_2} \left(\frac{4F}{\pi\delta_3}\right)^{\delta_3} \left(\frac{1}{2S\delta_4}\right)^{\delta_4} \left(\frac{P}{\delta_5}\right)^{\delta_5} \left(\frac{T}{\delta_6}\right)^{\delta_6} \delta_3^{\delta_3}\delta_4^{\delta_4}\delta_5^{\delta_5}\delta_6^{\delta_6} \to \max;$$

$$\begin{gathered} \delta_1 - 2\delta_3 + \delta_4 = 0, \ \delta_2 - \delta_4 - \delta_6 = 0, \\ -\delta_3 + \delta_4 + \delta_5 = 0, \ \delta_1 + \delta_2 = 1, \\ \delta_1, \delta_2, \ldots, \delta_6 \geqslant 0. \end{gathered} \tag{6.33}$$

Можно проверить, что, например, значения

$$\tilde{p} = P/(T+1), \ \tilde{t} = T + \sigma + 1, \ \tilde{d} = 2S(T+1)(T+\sigma)/P,$$

где $\sigma = [FP/\pi S^2 (T+1)]^{1/2}$, удовлетворяют всем ограничениям задачи (6.32) в форме строгих неравенств, а положительные значения $\delta_1^* = 0{,}3$, $\delta_2^* = 0{,}7$, $\delta_3^* = 0{,}2$, $\delta_4^* = \delta_5^* = 0{,}1$, $\delta_6^* = 0{,}6$ удовлетворяют всем ограничениям двойственной задачи (6.33). Следовательно, при любых фиксированных положительных числах F, P, S, T, согласно теореме двойственности, задачи (6.32) и (6.33) имеют оптимальные решения, причем оптимальные значения обеих задач совпадают.

Трудность исходной задачи равна 6—3—1=2, поэтому оптимальное решение двойственной задачи получить не просто. Но мы и не будем решать двойственную задачу. Найдем оптимальное решение исходной задачи, используя «простую структуру» ограничений задачи (6.32) и второе утверждение теоремы двойственности. Для этого выразим все двойственные переменные через δ_3 и δ_4:

$$\delta_1 = 2\delta_3 - \delta_4,\ \delta_2 = 1 - 2\delta_3 + \delta_4,\ \delta_5 = \delta_3 - \delta_4,\ \delta_6 = 1 - 2\delta_3. \qquad (6.34)$$

Как и ранее, оптимальные значения исходных и двойственных переменных будем обозначать через $\bar{d}$, $\bar{p}$, $\bar{t}$ и $\bar{\delta}_1, \bar{\delta}_2, \ldots, \bar{\delta}_6$.

Если $\bar{\delta}_1 = 0$, то равенство в (6.21) при $i=1$ принимает вид $\bar{d} = 0 \cdot v(\bar{\delta}) = 0$, что противоречит условиям задачи. Следовательно, $\bar{\delta}_1 \neq 0$. Аналогично можно установить справедливость неравенства $\bar{\delta}_2 \neq 0$. Выполняется и неравенство $\bar{\delta}_3 \neq 0$, так как в противном случае из первого равенства (6.34) и условия неотрицательности двойственных переменных следовало бы, что $\bar{\delta}_1 = 0$.

Рассмотрим случай, когда $\bar{\delta}_4 \neq 0$, $\bar{\delta}_5 \neq 0$, $\bar{\delta}_6 \neq 0$. Равенства (6.21) здесь принимают следующий вид:

$$\bar{d} = \bar{\delta}_1 v(\bar{\delta}),\ 2\bar{t} = \bar{\delta}_2 v(\bar{\delta}),$$

$$\frac{4F}{\pi \bar{d}^2 \bar{p}} = 1,\ \frac{\bar{d}\bar{p}}{2S\bar{t}} = 1,\ \frac{\bar{p}}{P} = 1,\ \frac{T}{\bar{t}} = 1. \qquad (6.35)$$

Отсюда, используя третье, пятое и шестое равенства, найдем

$$\bar{d} = 2(F/(\pi P))^{1/2},\ \bar{p} = P,\ \bar{t} = T. \qquad (6.36)$$

Для полученных значений переменных должно быть выполнено и четвертое из равенств (6.35), т. е. $(F/(\pi P))^{1/2}\, P/(ST) = 1$. В этом случае всегда можно подобрать такие положительные числа $\bar{\delta}_1, \bar{\delta}_2, \ldots, \bar{\delta}_6$, которые были бы допустимыми в двойственной задаче и вместе с числами $\bar{d}$, $\bar{p}$, $\bar{t}$ вида (6.36) удовлетворяли всем шести равенствам (6.35). Следовательно, согласно второму утверждению теоремы двойственности, решение вида (6.36) является оптимальным.

Если $\bar{\delta}_4 = 0$, $\bar{\delta}_5 \neq 0$, $\bar{\delta}_6 \neq 0$, то согласно (6.21) и равенству $\delta_4 = 0$ четвертое из равенств (6.35) может не выполняться. Но оно для получения значений (6.36) и не использовалось. Поэтому решение (6.36) также является оптимальным, если для него второе из неравенств (6.32) выполняется как строгое неравенство, т. е. если $(F/(\pi P))^{1/2}\, P/(ST) < 1$.

Объединяя рассмотренные случаи, приходим к следующему выводу: если $FP/(\pi S^2 T^2) \leqslant 1$, то решение вида (6.36) является оптимальным.

Из равенств (6.34) следует, что если $\bar{\delta}_4 = 0$ и верно хотя бы одно из равенств $\bar{\delta}_5 = 0$ или $\bar{\delta}_6 = 0$, то выполняются равенства $\bar{\delta}_1 = 0$ или $\delta_2 = 0$ соответственно. Этот случай можно исключить из дальнейшего рассмотрения.

Проанализируем оставшиеся возможности. Пусть $\bar{\delta}_4 \neq 0$, $\bar{\delta}_5 = 0$, $\bar{\delta}_6 \neq 0$. Из ограничений двойственной задачи (6.33) следует, что это имеет место тогда и только тогда, когда $\overline{\delta}_1<1/2$. Так как $\overline{\delta}_5=0$, то пятое из равенств (6.35) может нарушаться. Поэтому из третьего, четвертого и шестого равенств (6.35) находим

$$\overline{d}=2F/(\pi ST),\ \overline{t}=T,\ \overline{p}=\pi S^2T^2/F.$$

На основании второго утверждения теоремы двойственности это решение является оптимальным, если $\overline{p}<P$ (т. е. если $FP/(\pi S^2T^2)>1$), при одновременном выполнении требования $\overline{\delta}_1<1/2$, которое равносильно неравенству $F/(\pi ST^2)<1$.

Пусть $\overline{\delta}_4\neq 0$, $\overline{\delta}_5\neq 0$, $\overline{\delta}_6=0$. Из ограничений двойственной задачи (6.33) следует, что этот случай возможен тогда и только тогда, когда $\overline{\delta}_1>1/2$. Не принимая во внимание последнее из равенств (6.35), из третьего, четвертого и пятого равенств находим

$$\overline{d}=2(F/(\pi P))^{1/2},\ \overline{p}=P,\ \overline{t}=(FP/\pi S^2))^{1/2}.$$

Полученное решение, по теореме двойственности, является оптимальным при $\overline{t}>T$ (т. е. при $FP/(\pi S^2T^2)>1$), а также при условии $\overline{\delta}_1>1/2$, которое равносильно неравенству $P<S$.

Наконец, пусть $\overline{\delta}_4\neq 0$, $\overline{\delta}_5=\overline{\delta}_6=0$. Тогда из равенств (6.34) следует, что $\overline{\delta}_1=\overline{\delta}_2=\overline{\delta}_3=\overline{\delta}_4=1/2$. В этом случае $v(\overline{\delta})=4(F/(\pi S))^{1/2}$ и из первого, второго и третьего равенств (6.35) находим

$$\overline{d}=2(F/(\pi S))^{1/2},\ \overline{t}=(F/(\pi S))^{1/2},\ \overline{p}=S.$$

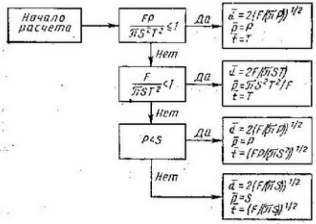

Рис. 6.2

Это решение является оптимальным, если одновременно имеют место неравенства $\overline{p}<P$ и $\overline{t}>T$, т. е. если $S<P$ и $F/(\pi ST^2)>1$.

На основании второго утверждения теоремы двойственности (необходимость) можно сделать вывод о том, что других оптимальных решений в данной задаче нет. В соответствии с полученными результатами оптимальные параметры гидравлического цилиндра можно рассчитывать по схеме, изображенной на рис. 6.2.

Другие технические приложения геометрического программирования можно найти в [10], а также в книге Д. Уайлда «Оптимальное проектирование» (М.: Мир, 1981).

Глава 7

ЭЛЕМЕНТЫ МНОГОКРИТЕРИАЛЬНОЙ ОПТИМИЗАЦИИ

Эта глава занимает особое положение. Она посвящена рассмотрению многокритериальных задач, т. е. задач оптимизации, в которых имеется не одна, а сразу несколько целевых функций (критериев). На практике многокритериальные задачи возникают, когда, удается сформулировать и формализовать в виде критериев лишь ряд отдельных требований, предъявляемых к оптимальному решению. «Объединить» эти отдельные критерии в единый, обобщенный, критерий не представляется возможным. Нередко к многокритериальным задачам приходят в случае отсутствия полных и точных сведений о решаемой задаче или тогда, когда оптимальное решение следует оценивать сразу с нескольких точек зрения.

В настоящее время теория многокритериальной оптимизации сформировалась в самостоятельный раздел теории оптимизации и находит многочисленные приложения в самых различных областях техники, экономики и математики.

Здесь введены начальные понятия и сформулированы простейшие результаты теории многокритериальной оптимизации, дающие общее представление о предмете. Более подробно с этой областью оптимизации можно ознакомиться, например, в [18].

§ 7.1. НАЧАЛЬНОЕ ПРЕДСТАВЛЕНИЕ О МНОГОКРИТЕРИАЛЬНОЙ ЗАДАЧЕ ОПТИМИЗАЦИИ

1. Общая задача принятия решений. Задача принятия решения состоит в выборе среди множества возможных решений (их называют также вариантами, планами и т. п.) такого решения, которое являлось бы в определенном смысле лучшим, или, как говорят, оптимальным.

Удобно считать, что выбор решения производит некоторое *лицо, принимающее решение,* которое преследует вполне определенные цели. В зависимости от конкретной ситуации в роли лица, принимающего решение, может выступать как отдельный человек (инженер, научный сотрудник и т. п.), так и целый коллектив (группа специалистов, занятая решением одной задачи).

Каждое возможное решение характеризуется определенной степенью достижения цели. В соответствии с этим у лица, принимающего решение, имеется свое представление о достоинствах и недостатках решений, на основании которого одно решение предпочитается другому. Оптимальное решение — это решение, которое с точки зрения лица, принимающего решение, предпочтительнее других возможных решений. Таким образом, *понятие оптимального решения связано с предпочтениями лица, принимающего решение.* Эти предпочтения на практике выражаются в различной форме, и их математическая формализация может составить сложную задачу, поскольку лицо, принимающее решение, как правило, не может ясно и четко сформулировать их.

Цель теории принятия решений и состоит в разработке методов, которые помогли бы лицу, принимающему решение, наиболее полно и точно выразить свои предпочтения в рамках соответствующей математической модели и в конечном счете обоснованно выбрать действительно оптимальное решение.

2. Многокритериальная задача оптимизации. Многокритериальные задачи оптимизации, рассмотрению которых посвящена данная глава, вместе со *множеством возможных (допустимых) решений* $X \subset \mathbf{R}^n$ включает *набор целевых функций* (называемых также *критериями*) $f_1, f_2, \ldots, f_m$, заданных на множестве X. Предполагается, что $m > 1$; при $m = 1$ задача оптимизации является однокритериальной. Набор целевых функций образует вектор-функцию, которую далее будем обозначать через $f(x) = (f_1(x), f_2(x), \ldots, f_m(x))$.

Наряду со множеством допустимых решений X удобно рассматривать множество

$$Y = f(X) = \{y \in \mathbf{R}^m \mid y = f(x), \text{ при некотором } x \in X\}, \qquad (7.1)$$

которое называют *множеством оценок*, а его элементы — *оценками*. Часто пространство $\mathbf{R}^n$, в котором содержится X, называют *пространством решений*, а пространство $\mathbf{R}^m$, в котором содержится Y, — *пространством оценок* или *критериальным пространством*.

Каждому решению $x \in X$ соответствует одна вполне определенная оценка $y = f(x) \in Y$. С другой стороны, каждой оценке $y \in Y$ отвечают те решения $x \in X$ (их может быть и более одного), для которых $f(x) = y$ (рис. 7.1). Таким образом, между множеством X и Y имеется тесная связь, и поэтому выбор решения из X в указанном смысле равносилен выбору соответствующей оценки из Y.

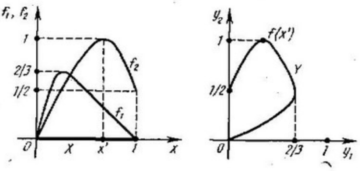

Рис. 7.1

В отличие от общей задачи принятия решений в многокритериальной задаче оптимизации заранее известна определенная информация о предпочтениях лица, принимающего решение. Это сведения о том, что желательно целевые функции максимизировать (или минимизировать). При этом может быть известна и использована и различная дополнительная информация о предпочтениях.

Для описания предпочтений лица, принимающего решение, используют математическое понятия, называемое отношением. Точному определению и изучению отношений посвящен следующий параграф.

§ 7.2. ОТНОШЕНИЯ

1. Определение отношения. С простыми примерами отношений мы уже встречались, когда для сравнения вещественных чисел использовали знаки $>$, $\geqslant$, $<$, $\leqslant$, $=$. Применение этих символов предполагает наличие пары чисел, одно из которых записывается слева от символа, а другое — справа. При этом говорят, что данные числа находятся в некотором отношении друг относительно друга (первое число больше второго, первое число больше либо равно второму и т. д.).

В определенных отношениях могут находиться не только числа, но и более сложные объекты, причем отношения между ними носят различный характер.

Дадим строгое определение отношения. Пусть A — некоторое множество. Образуем декартово произведение $A \times A$ — множество всех упорядоченных пар вида (a, b), где $a \in A$, $b \in A$. *Отношением* R, заданным на множестве A, называют подмножество множества $A \times A$, т. е. такое множество R, что $R \subset A \times A$. Если выполнено соотношение $(a, b) \in R$, то говорят, что элементы a и b *находятся в отношении* R, и при этом пишут aRb. Заметим, что декартово произведение — это множество упорядоченных пар, поэтому запись aRb и bRa означает не одно и то же (за исключением случая $a=b$). На одном и том же множестве A могут быть заданы различные отношения в зависимости от того, какие именно пары (a, b) составляют множество R. В частности, множество R может не содержать ни одной пары, содержать все возможные пары, т. е. $R = A \times A$, включать только пары из одинаковых элементов (a, a) (*отношение равенства*).

Рассмотрим примеры отношений. Пусть $A = \mathbf{R}^n$. Уже встречались отношения $\geqq$ и $>$, заданные на $\mathbf{R}^n$:

$$a \geqq b \text{ означает, что } a_i \geqslant b_i,\ i = 1, 2, \ldots, n;$$

$$a > b \text{ означает, что } a_i > b_i,\ i = 1, 2, \ldots, n,$$

где $a = (a_1, a_2, \ldots, a_n)$, $b = (b_1, b_2, \ldots, b_n)$. Далее будем использовать еще одно отношение на $\mathbf{R}^n$.

$$a \geq b \text{ означает, что } a \geqq b \text{ и } a \neq b.$$

Таким образом, $a \geq b$ имеет место тогда и только тогда, когда $a_i \geqslant b_i$, $i = 1, 2, \ldots, n$, причем хотя бы для одного номера $i \in \{1, 2 \ldots, n\}$ выполняется строгое неравенство $a_i > b_i$. При $n=1$ отношение $\geq$ совпадает с отношением $>$ для чисел.

При $m=2$ неравенство $a \geq b$ геометрически означает, что точка a находится в заштрихованной области, имеющей форму прямого угла с выколотой вершиной b, стороны которого параллельны координатным осям (рис. 7.2). А точка a, для которой выполняется неравенство $a > b$, является внутренней точкой такого угла.

Если A — это множество всех прямых на некоторой плоскости, а R составляют пары прямых, не имеющие ни одной общей точки, то R не что иное, как отношение параллельности прямых.

Если A — множество людей, то на этом множестве можно задать, например, отношения «является родственником», «старше, чем» и т. п.

На множестве A автомобильных двигателей внутреннего сгорания одинакового объема можно ввести отношение R: «более экономичен, чем». А именно: соотношение aRb (двигатель a более экономичен, чем b) верно тогда и только тогда, когда у двигателя a меньший расход горючего, чем у двигателя b.

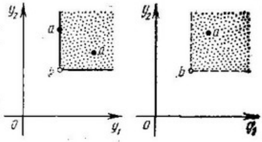

Рис. 7.2

Так как отношение — это некоторое множество, то к различным отношениям, заданным на одном и том же множестве, можно применять все известные теоретико-множественные операции. В частности, можно рассматривать объединение, пересечение и разность двух отношений. Например, отношение $\geqslant$ для чисел * представляет собой объединение отношений $>$ и $=$, отношение $=$ есть пересечение отношений $\geqslant$ и $\leqslant$, а отношение $\geq$ для векторов из $\mathbf{R}^n$ $(n>1)$ представляет собой разность отношений $\geqq$ и $=$.

2. Типы отношений. Тот или иной тип отношения получается в зависимости от того, какими свойствами обладает данное отношение. Приведем наиболее важные типы отношений.

Отношение R называют *рефлексивным*, если aRa верно для любого $a\in A$. Примерами рефлексивных отношений служат отношения равенства $=$ и неравенства $\geqq$ на $\mathbf{R}^n$.

Отношение R называют *иррефлексивным*, если aRa не может иметь места ни для какого $a\in A$. Например, отношения $>$ и $\geq$ иррефлексивны на пространстве $\mathbf{R}^n$, а отношение «старше, чем» иррефлексивно на множестве, элементами которого являются люди.

Отношение R называется *симметричным*, если для произвольной пары $a, b\in A$ из выполнения aRb всегда следует выполнение bRa; *асимметричным*, если для произвольной пары $a, b\in A$ из справедливости aRb вытекает, что bRa не выполнено, и *антисимметричным*, если для произвольной пары $a, b\in A$ выполнение двух соотношений aRb и bRa всегда влечет совпадение a и $b: a=b$. Так, на соответствующих множествах отношения равенства и параллельности симметричны, отношения $>$ и $\geq$ асимметричны, а отношение $\geqq$ антисимметрично.

Не все перечисленные типы отношений независимы. Например, *если отношение R асимметрично, то оно и иррефлексивно.* В самом деле, если оно не является иррефлексивным, то для некоторого элемента $a\in A$ выполнено соотношение aRa, откуда в силу асимметричности следует, что соотношение aRa не должно иметь места. Это противоречие говорит о том, что отношение R иррефлексивно.

---

* Для векторов объединение отношений $>$ и $=$ не совпадает с отношением $\geqq$.

Отношение R называют *транзитивным*, если из выполнения двух соотношений aRb и bRc, где $a, b, c \in A$, всегда следует соотношение aRc. Все отношения, приведенные в данном пункте в качестве примеров, являются транзитивными. Рассмотрим на множестве $\mathbf{R}^2$ отношение R, определяемое следующим образом: aRb верно тогда и только тогда, когда выполняется хотя бы одно из неравенств $a_1 > b_1$, $a_2 > b_2$. Это отношение не является транзитивным, о чем свидетельствует простой пример: справедливо $(1,1)^{\mathrm{T}} R (3,0)^{\mathrm{T}}$, $(3,0)^{\mathrm{T}} R (2,2)^{\mathrm{T}}$, однако $(1,1)^{\mathrm{T}} R (2,2)^{\mathrm{T}}$ не верно.

Взаимосвязь различных типов отношений можно продолжить. Так, *если отношение R иррефлексивно и транзитивно, то оно и асимметрично.* Предположим противное, т. е. что отношение не является асимметричным, т. е. для некоторой пары элементов a, $b \in A$, удовлетворяющей соотношению aRb, выполнено соотношение bRa. Из aRb и bRa, согласно транзитивности, следует соотношение aRa, что противоречит условию иррефлексивности. Значит, отношение R асимметрично.

Элементы a и b множества A называют *сравнимыми по отношению R*, если обязательно выполняется соотношение aRb или bRa (может быть и то и другое вместе), и *не сравнимыми по отношению R*, если не верно ни соотношение aRb, ни bRa. Например, любые два вещественных числа сравнимы по отношению $\geqslant$, но могут быть не сравнимыми по отношению $>$ (поскольку неравенство $a > a$ не верно). Если любые два элемента множества A сравнимы по отношению R, то такое отношение называют *полным*. Если же в множестве A найдется хотя бы одна пара элементов, не сравнимых от отношению R, то R называют *частичным отношением*. На множестве вещественных чисел отношение $\geqslant$ является полным, а отношение $>$ — частичным. Для множества $A = \mathbf{R}^n$ отношение $\geqq$ уже не является полным, так как, например, векторы $(1, -1)^{\mathrm{T}}$ и $(-1, 1)^{\mathrm{T}}$ не сравнимы по отношению $\geqq$. Значит, в этом случае отношение $\geqq$ является частичным.

Если отношение R рефлексивно, антисимметрично и транзитивно, то его называют *отношением порядка* (или *порядком*). Отношение порядка может быть полным или частичным в зависимости от того, сравнимы любые два элемента из множества A по этому отношению или же не сравнимы. Так, отношение $\geqslant$ представляет собой полный порядок на множестве вещественных чисел, а отношение $\geqq$ — частичный порядок на множестве $\mathbf{R}^n$ $(n > 1)$. Отношение $>$ на пространстве $\mathbf{R}^n (n \geqslant 1)$ является частичным отношением порядка.

§ 7.3. ОТНОШЕНИЯ ПРЕДПОЧТЕНИЯ И НЕРАЗЛИЧИМОСТИ. МНОЖЕСТВО ОПТИМАЛЬНЫХ РЕШЕНИЙ

1. Отношения предпочтения и неразличимости. Выбор решения из множества возможных решений X равносилен выбору оценки из множества оценок Y (см. § 7.1), поэтому в этом параграфе для удобства множество возможных решений будем обозначать буквой

Z, считая, что в качестве Z может быть взято как множество X, так и множество Y.

Прямой выбор некоторого наилучшего (наиболее предпочтительного) решения из всего множества Z, как правило, оказывается трудным. Поэтому упростим задачу и выберем лучшее из двух данных решений, что значительно проще (если, конечно, Z включает более двух решений). Далее, аналогично, выяснив предпочтение для каждой пары решений, получим информацию, на основании которой лицо, принимающее решение, произведет выбор решения и из всего множества Z.

Для определенности зафиксируем некоторое лицо, принимающее решение. Если из двух заданных решений a и b множества Z лицо, принимающее решение, выбирает решение a, то будем говорить, что *решение a более предпочтительно, чем решение b*. Все пары вида (a, b), где $a, b \in Z$, для которых решение a более предпочтительно, чем решение b, образуют некоторое множество, которое будем называть *отношением строгого предпочтения* и обозначать символом $\succ$. Указанное множество является отношением, заданным на множестве Z. В соответствии с этим запись $a \succ b$ означает, что решение a для лица, принимающего решение, более предпочтительно, чем решение b.

Если лицо, принимающее решение, ведет себя «разумно», то отношение $\succ$ должно быть иррефлексивным. В самом деле, бессмысленно решению a предпочитать то же самое решение a. Более того, естественно считать отношение $\succ$ асимметричным, так как в противном случае могут одновременно выполняться соотношения $a \succ b$ и $b \succ a$, что противоречит смыслу.

Во многих случаях целесообразно предполагать введенное отношение $\succ$ еще и транзитивным. Транзитивность отношения $\succ$ означает, что если лицо, принимающее решение, решение a считает более предпочтительным, чем решение b, а решение b — более предпочтительным, чем решение c, то из двух решений a и c будет выбрано a. Однако в некоторых ситуациях руководствуются нетранзитивным отношением предпочтения (см., например: *Фишберн П.* Теория полезности для принятия решений. — М.: Наука, 1978).

Когда сравнивается по предпочтительности пара решений a и b, возможен и такой случай, когда не будет отдано предпочтения ни одному из них. Это заведомо имеет место, если, например, $a=b$. Поэтому имеет смысл ввести следующее определение. Будем говорить, что решения a и b, где $a, b \in Z$, *неразличимы*, если не выполняется ни соотношение $a \succ b$, ни $b \succ a$. Другими словами, решения a и b неразличимы, если они не сравнимы по отношению $\succ$. Множество всех пар вида (a, b), в которых решения a и b неразличимы, называют *отношением неразличимости (отношением безразличия)* и обозначают символом $\sim$.

Не надо считать, что соотношение $a \sim b$ означает равенство $a=b$. Если, например, $Z=\mathbb{R}^n$ $(n>1)$ и в качестве отношения $\succ$ взято отношение $\geq$, то верно $(1,0)^{\mathrm{T}} \sim (0,1)^{\mathrm{T}}$, однако $(1,0)^{\mathrm{T}} \neq$

$\neq (0,1)^{\mathrm{T}}$. Соотношение $a \sim b$ может иметь место тогда, когда лицо, принимающее решение, считает, что в смысле предпочтения для него нет разницы между решениями a и b (в частности, когда $a=b$). Кроме того, неразличимость может быть и в случае, если решения a и b лицо, принимающее решение, вообще никак не может сравнить друг с другом.

Непосредственно из определения отношения $\sim$ следует, что оно симметрично, т. е. соотношение $a \sim b$ равносильно соотношению $b \sim a$. Из иррефлексивности отношения $\succ$ следует рефлексивность отношения $\sim$. В самом деле, если отношение $\sim$ не является рефлексивным, то найдется элемент $a \in Z$, для которого соотношение $a \sim a$ не выполнено. Отсюда, по определению отношения $\sim$, следует справедливость соотношения $a \succ a$, что несовместимо с условием иррефлексивности отношения $\succ$.

Отношение $\sim$ может не быть транзитивным, даже если транзитивно отношение $\succ$. Так, например, для транзитивного отношения $\geq$ на пространстве R^2 выполнены соотношения $(3, 0)^{\mathrm{T}} \sim (1, 1)^{\mathrm{T}}$, $(1, 1)^{\mathrm{T}} \sim (2, -1)^{\mathrm{T}}$, однако $(3, 0)^{\mathrm{T}} \geq (2, -1)^{\mathrm{T}}$, поэтому соотношение $(3, 0)^{\mathrm{T}} \sim (2, -1)^{\mathrm{T}}$ не имеет места.

Итак, *для произвольно выбранной пары решений $a, b \in Z$ выполняется одно и только одно из следующих трех соотношений: $a \succ b$, $b \succ a$, $a \sim b$.*

Нередко удобно рассматривать еще одно отношение «не менее предпочтительно, чем», являющееся объединением отношений $\succ$ и $\sim$. Будем называть его *отношением нестрогого предпочтения* и обозначать $\succsim$. Таким образом, соотношение $a \succsim b$ обозначает, что имеет место соотношение $a \succ b$ или же выполнено соотношение $a \sim b$.

При транзитивном отношении $\succ$ отношение неразличимости может не быть транзитивным, поэтому и отношение $\succsim$ также может не обладать свойством транзитивности. Отметим, что $\succsim$ — полное отношение, так как для любых двух решений $a, b \in Z$ верно $a \succsim b$ или $b \succsim a$ (возможно, и то и другое вместе).

2. Множество оптимальных решений. Пусть лицо, принимающее решение, при выборе решения из множества Z руководствуется некоторым отношением строгого предпочтения $\succ$, являющимся асимметричным и транзитивным. Воспользуемся отношением $\succ$ для того, чтобы выделить решения, которые могут быть «лучшими», т. е. оптимальными, решениями. Все те решения, для которых найдутся более предпочтительные, следует удалить из Z; их заведомо нельзя считать оптимальными. В результате подобного исключения в Z останутся решения (единственное решение), каждое из которых может быть признано оптимальным согласно данному отношению $\succ$.

Решение $z^{(0)} \in Z$ будем называть *оптимальным* * *по отношению* $\succ$ на множестве Z, если не существует другого решения $z \in Z$, для

* Вместо термина *оптимальное* решение в литературе часто используют термины *максимальное, минимальное, неулучшаемое, недоминируемое* решение.

которого справедливо соотношение $z \succ z^{(0)}$. Используя отношение $\succsim$, построенное описанным выше способом на основании отношений $\succ$ и $\sim$, можно привести эквивалентную формулировку: решение $z^{(0)} \in Z$ называют оптимальным по отношению $\succsim$, если для любого $z \in Z$ выполняется соотношение $z^{(0)} \succsim z$. *Множество всех оптимальных решений* множества Z будем обозначать через $\mathrm{opt}_{\succ} Z$ В зависимости от структуры Z и вида отношения $\succ$ множество $\mathrm{opt}_{\succ} Z$ может содержать единственный элемент, конечное или бесконечное множество элементов, а также не содержать ни одного элемента.

Пусть непустое множество $\mathrm{opt}_{\succ} Z$ содержит по крайней мере два элемента. Рассмотрим два произвольных оптимальных решения z' и z''. Поскольку ни $z' \succ z''$, ни $z'' \succ z'$ не может иметь места, решения z' и z'' неразличимы: $z' \sim z''$. Таким образом, *если отношение неразличимости $\sim$ совпадает с отношением равенства $=$, то множество $\mathrm{opt}_{\succ} Z$ (если оно не пусто) содержит единственный элемент.* Условие совпадения отношений $\sim$ и $=$ является довольно «жестким». Согласно этому условию, для каждой пары решений $a, b \in Z$ должно выполняться одно и только одно из следующих трех соотношений: $a \succ b$, $b \succ a$, $a = b$, т. е. решения a и b либо сравнимы по отношению $\succ$, либо равны друг другу.

Сформулируем требования, гарантирующие существование оптимальных решений.

Теорема 7.1. *Если множество Z не пусто и содержит конечное число элементов, а отношение $\succ$ асимметрично и транзитивно, то $\mathrm{opt}_{\succ} Z \neq \varnothing$.*

□ Доказательство этого утверждения носит конструктивный характер и представляет собой *алгоритм нахождения всего множества оптимальных решений.* Введем обозначение

$$Z = Z_1 = \{z^{(11)}, z^{(12)}, \dots, z^{(1n_1)}\}.$$

Если $n_1 = 1$, то $Z_1 = \{z^{(11)}\} = \mathrm{opt}_{\succ} Z_1$. Поэтому далее будем считать $n_1 > 1$.

Первый шаг алгоритма заключается в попарном сравнении решения $z^{(11)}$ с каждым из остальных решений. Если для некоторого $i \in \{2, 3, \dots, n_1\}$ выполняется соотношение $z^{(11)} \succ z^{(1i)}$, то решение $z^{(1i)}$ из множества Z_1 удаляют; оно не может быть оптимальным. В противном случае, т. е. когда $z^{(11)} \sim z^{(1i)}$ либо $z^{(1i)} \succ z^{(11)}$, решение $z^{(1i)}$ сохраняют. После выполнения всех сравнений решение $z^{(11)}$ также следует удалить из Z_1. При этом если ни для какого $i = 2, 3, \dots, n_1$ не оказалось выполненным соотношение $z^{(1i)} \succ z^{(11)}$, то решение $z^{(11)}$ является оптимальным и его нужно запомнить. Оставшееся в результате удаления множество решений обозначим через $Z_2 = \{z^{(21)}, z^{(22)}, \dots, z^{(2n_2)}\}$, $n_2 < n_1$.

Если $Z_2 = \varnothing$, то решение $z^{(11)}$ оптимальное (оно хранится в памяти), поскольку в силу асимметричности отношения $\succ$ из соотношения $z^{(11)} \succ z^{(1i)}$ следует, что соотношение $z^{(1i)} \succ z^{(11)}$, $i = 2, 3, \dots, n_1$, не может иметь места. В этом случае процедура отыскания множества $\mathrm{opt}_{\succ} Z_1$ закончена. Если же $Z_2 \neq \varnothing$, то переходят к следующему шагу алгоритма.

Второй шаг аналогичен первому и заключается в попарном сравнении решения $z^{(21)}$ с каждым из решений $z^{(22)}, \ldots, z^{(2n_2)}$. Все решения $z^{(2i)}$, для которых $z^{(21)} \succ z^{(2i)}$, из множества Z_2 исключают. Кроме того, удаляют решение $z^{(21)}$. При этом, если ни для какого $i=2, 3, \ldots, n_2$ не оказалось выполненным соотношение $z^{(2i)} \succ z^{(21)}$, то $z^{(21)} \in \mathrm{opt}_{\succ} Z_2$. Более того, $z^{(21)} \in \mathrm{opt}_{\succ} Z_1$ и решение $z^{(21)}$ следует запомнить. В самом деле, соотношение $z^{(11)} \succ z^{(21)}$ не может иметь места, так как решение $z^{(21)}$ не было удалено из Z_1 на первом шаге. Соотношение $z^{(1i)} \succ z^{(2i)}$ для $z^{(1i)} \in Z_1 \setminus Z_2$, $i \neq 1$, также не может быть выполнено, поскольку $z^{(11)} \succ z^{(1i)}$ и отношение $\succ$ транзитивно: из $z^{(11)} \succ z^{(1i)}$ и $z^{(1i)} \succ z^{(21)}$ следует, что $z^{(11)} \succ z^{(21)}$. Оставшееся после исключения множество решений обозначают через $Z_3 = \{z^{(31)}, z^{(32)}, \ldots, z^{(3n_3)}\}$, $n_3 < n_2$. Если $Z_3 \neq \varnothing$, то переходят к следующему шагу и т. д.

Алгоритм таков, что, согласно транзитивности отношения $\succ$, решение $z^{(k1)}$, оптимальное на множество Z_k, является оптимальным на Z_{k-1}, $k=2, 3, \ldots$, а значит, и на исходном множестве Z_1.

Так как множество Z_1 содержит конечное число элементов, то через конечное число шагов процедура закончится. Решения, хранящиеся в памяти, образуют искомое непустое множество $\mathrm{opt}_{\succ} Z$ ■

Оценим «трудоемкость» сформулированного алгоритма, т. е. определим наименьшее и наибольшее возможное число попарных сравнений, которое потребуется для нахождения всего множества $\mathrm{opt}_{\succ} Z$. Наименьшее число сравнений $n_1 - 1$ имеет место, если $z^{(11)} \succ z^{(1i)}$, $i=2, 3, \ldots, n_1$. В самом «длинном варианте» придется сравнивать между собой все возможные пары решений, и поэтому максимальное число сравнений равно $n_1(n_1-1)/2$.

§ 7.4. МНОГОКРИТЕРИАЛЬНЫЕ ЗАДАЧИ ОПТИМИЗАЦИИ. ПАРЕТО-ОПТИМАЛЬНЫЕ ОЦЕНКИ И РЕШЕНИЯ

1. Согласованность отношений предпочтения на множестве решений и на множестве оценок. Пусть целевая векторная функция $f(x) = (f_1(x), f_2(x), \ldots, f_m(x))$ определена на множестве допустимых решений $X \subset \mathbf{R}^n$. Данному множеству X при отображении f соответствует множество оценок Y, определяемое равенством (7.1). Будем полагать, что на множествах X и Y заданы отношения строгого предпочтения $\underset{X}{\succ}$ и $\underset{Y}{\succ}$ соответственно. Каждому решению $x \in X$ соответствует определенная оценка $y = f(x) \in Y$, и, наоборот, каждой оценке y соответствуют такие решения x, для которых $f(x) = y$, поэтому указанные отношения согласованы друг с другом: $y \underset{Y}{\succ} y'$ имеет место тогда и только тогда, когда $x \underset{X}{\succ} x'$, где $y = f(x)$ и $y' = f(x')$. Следовательно, результаты, сформулированные в терминах оценок, могут быть переформулированы применительно к решениям, и наоборот.

В данном параграфе рассмотрены «минимальные» требования к указанным отношениям предпочтения, на основании которых получена оценка сверху для искомого множества оптимальных решений (оценок). Под *оценкой сверху* будем понимать подмножество исходного множества возможных решений (оценок), которое содержит множество оптимальных решений (оценок).

2. Аксиома Парето. Многокритериальные задачи. Рассмотрим две произвольные оценки y, $y' \in \mathbf{R}^m$, которые связаны между собой неравенством $y \geq y'$ (т. е. $y \geqq y'$ и $y \neq y'$). При этом оценка y может оказаться для лица, принимающего решение, предпочтительнее, чем y'. *Многокритериальная задача максимизации* характеризуется тем, что оценка y всегда предпочтительнее оценки y', если только $y \geq y'$. Иначе говоря, в многокритериальной задаче максимизации считают выполненной следующую аксиому.

Аксиома Парето* (в терминах оценок). *Для любых двух оценок y, $y' \in Y$, удовлетворяющих неравенству $y \geq y'$, всегда выполнено соотношение $y \succ_Y y'$.*

Аксиома Парето (в терминах решений). *Для любых двух решений x, $x' \in X$, для которых верно $f(x) \geq f(x')$, всегда имеет место соотношение $x \succ_X x'$.*

В многокритериальной задаче минимизации считают, что для двух произвольных оценок y, $y' \in Y$, связанных неравенством $y \geq y'$, всегда выполняется соотношение $y' \succ_Y y$. Для определенности ограничимся рассмотрением задач максимизации (полученные результаты и выводы можно легко переформулировать применительно к задачам минимизации).

Аксиома Парето накладывает определенные требования на характер отношения предпочтения в многокритериальной задаче максимизации. А именно: для лица, принимающего решение, желательно по каждому из критериев $f_1, f_2, \ldots, f_m$ получить по возможности большее значение, т. е. максимизировать каждый из критериев. Точка максимума на множестве X одновременно всех функций $f_1, f_2, \ldots, f_m$ заведомо является оптимальным решением многокритериальной задачи максимизации. Однако на практике этот случай имеет место крайне редко, так как такой точки максимума, как правило, не существует. Поэтому при отсутствии дополнительной информации о предпочтениях $\succ_X$ и $\succ_Y$ в многокритериальной задаче удается лишь указать определенную оценку сверху для искомого множества оптимальных решений.

3. Парето-оптимальность. Согласно аксиоме Парето, в многокритериальных задачах отношение $\geq$ играет важную роль. Поэтому множество оптимальных оценок по отношению $\geq$ на мно-

---

* В. Парето (1848—1923) — итальянский экономист и социолог.

жестве Y имеет специальное название: *множество парето-оптимальных (оптимальных по Парето)* или *эффективных оценок*. Это множество обозначают через $P(Y)$. Используя символику, принятую в § 7.3, по определению можно записать: $P(Y) = \mathrm{opt}_{\geq} Y$. Далее будем использовать обозначение $P(Y)$. Таким образом, включение $y^{(0)} \in P(Y)$ имеет место тогда и только тогда, когда не существует оценки $y \in Y$, для которой было бы выполнено неравенство $y \geq y^{(0)}$.

При $m=1$ отношение $\geq$ превращается в отношение $>$ для чисел и парето-оптимальная оценка совпадает с максимальным элементом числового множества, $Y \subset \mathrm{R}$.

Если $m=2$ (т. е. имеется два критерия), то множество $P(Y)$ имеет простую геометрическую интерпретацию в критериальном пространстве (рис. 7.3). Для изображенного на этом рисунке множества Y парето-оптимальные оценки состоят из точек кривой bc (исключая точку c) и линии de. Образно говоря, $P(Y)$ является «северо-восточной» границей множества Y без тех ее частей, которые параллельны одной из координатных осей или же лежат в «глубоких провалах». Для того чтобы убедиться в этом, достаточно вспомнить, что все точки $y \in \mathrm{R}^2$, для которых верно неравенство $y \geq y'$ образуют прямой угол, стороны которого параллельны координатным осям, а вершиной является точка y' (см. рис. 7.2). Поэтому если для точки $y' \in Y$ указанный угол расположен вне множества Y, то эта точка парето-оптимальна, в противном случае она таковой не будет.

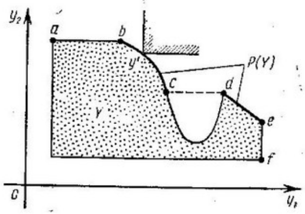

Рис. 7.3

Решение $x^{(0)} \in X$, для которого справедливо включение $y^{(0)} = f(x^{(0)}) \in P(Y)$, называют *парето-оптимальным (оптимальным по Парето)* или *эффективным решением* относительно вектор-функции f на множестве X. Множество всех таких решений обозначают через $P_f(X)$. Таким образом, включение $x^{(0)} \in P_f(X)$ имеет место тогда и только тогда, когда не существует $x \in X$ такого, что выполняется неравенство $f(x) \geq f(x^{(0)})$.

Если $m=1$, сформулированное определение парето-оптимального решения превращается в определение точки максимума числовой функции f_1. Таким образом, понятие парето-оптимальной точки можно рассматривать как обобщение понятия точки максимума функции на случай нескольких функций.

Для функций f_1 и f_2, изображенных на рис. 7.1, множество $P_f(X)$ представляет собой отрезок, который соединяет точки максимума данных функций (одна из них на рис. 7.1 обозначена буквой x'). Для функций, изображенных на рис. 7.4, имеем $X = [a, d]$ и $P_f(X) = [a, b) \cup [c, d]$.

Из аксиомы Парето вытекает следующее утверждение.

Следствие 4.1. *Справедливы включения*

$$\operatorname{opt}_{\underset{Y}{\succ}} Y \subset P(Y), \quad \operatorname{opt}_{\underset{X}{\succ}} X \subset P_f(X). \tag{7.2}$$

□ Проверим первое включение. Пусть $y^{(0)} \in \operatorname{opt}_{\succ} Y$. Предположим противное, т. е. что $y^{(0)} \notin P(Y)$. Тогда найдется такая оценка $y \in Y$, что $y \geq y^{(0)}$. Согласно аксиоме Парето, отсюда вытекает соотношение $y \underset{Y}{\succ} y^{(0)}$, а значит, $y^{(0)} \notin \operatorname{opt}_{\succ} Y$. Это противоречит начальному условию. Следовательно, выполнено включение $y^{(0)} \in P(Y)$. Аналогично можно проверить справедливость второго включения. ■

Согласно включениям (7.2), множества $P(Y)$ и $P_f(X)$ представляют собой оценки сверху для множества $\operatorname{opt}_{\succ} Y$ и $\operatorname{opt}_{\succ} X$ соответственно. Следовательно, в многокритериальной задаче окончательный выбор оптимальной оценки (оптимального решения) должен производиться в пределах множества парето-оптимальных оценок (решений). Вне множества $P(Y)$ (соответственно $P_f(X)$) оптимальных оценок (решений) быть не может.

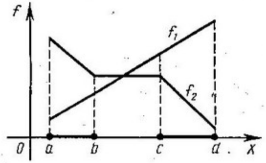

Рис. 7.4

4. Слабая аксиома Парето и слабая парето-оптимальность. Часто для того чтобы было можно рассматривать более широкий класс задач максимизации, вместо аксиомы Парето принимают ее «ослабленный» вариант.

Слабая аксиома Парето. *Для любых двух оценок $y, y' \in Y$ (двух решений $x, x' \in X$) из выполнения неравенства $y > y'$ (неравенства $f(x) > f(x')$) всегда следует, что $y \underset{Y}{\succ} y'$ ($x \underset{X}{\succ} x'$).*

По аналогии с оптимальными по Парето оценками и решениями вводят *слабо оптимальные по Парето (слабо эффективные) оценки и решения.* Множества, которые они составляют, обозначают соответственно через $S(Y)$ и $S_f(X)$. Таким образом, $S(Y) = \operatorname{opt}_{>} Y$, а включение $x^{(0)} \in S_f(X)$ выполняется тогда и только тогда, когда не существует решения $x \in X$, для которого верно неравенство $f(x) > f(x^{(0)})$.

Неравенство $y > y'$ представляет собой частный случай неравенства $y \geq y'$, поэтому из выполнения аксиомы Парето вытекает выполнение слабой аксиомы Парето, но не наоборот. На этом же основании имеют место включения

$$P(Y) \subset S(Y), \quad P_f(X) \subset S_f(X), \tag{7.3}$$

которые свидетельствуют о том, что множество слабо оптимальных по Парето оценок (решений) может быть разве что шире множества парето-оптимальных оценок (решений).

Для множества Y, изображенного на рис. 7.3, множество $S(Y)$ — это объединение кривых abc и def (включая точки a, c, d, f), т. е. это множество $P(Y)$ вместе с точкой c и теми участками северо-восточной границы множества Y, которые параллельны одной из координатных осей. Для случая, изображенного на рис. 7.4, $S_f(X)=X=[a, d]$, так как ни для какой точки $x^{(0)}\in[a, d]$ нельзя найти такую другую точку x, чтобы выполнялись оба неравенства $f_1(x)>f_1(x^{(0)})$, $f_2(x)>f_2(x^{(0)})$.

5. Условия оптимальности по Парето. Множества парето-оптимальных решений и оценок, а также слабо парето-оптимальных решений и оценок играют важную роль в теории многокритериальной оптимизации, поэтому рассмотрим их более подробно.

Приведем д о с т а т о ч н ы е у с л о в и я оптимальности по Парето и слабой оптимальности по Парето.

Говорят, что числовая функция $F(y)$, определенная на множестве $Y\subset\mathbf{R}^m$, является *возрастающей по отношению* $\geq$ *(по отношению* $>$*)*, если из выполнения неравенства $y\geq y'$ (неравенства $y>y'$) для векторов y, $y'\in Y$ всегда следует справедливость неравенства $F(y)>F(y')$. Это определение представляет собой обобщение понятия возрастающей функции одной переменной на случай функции многих переменных. Функция F, возрастающая по отношению $\geq$ на множестве Y, является возрастающей и по отношению $>$ на том же множестве.

Теорема 7.2. *Пусть функция $F(y)$ определена на множестве оценок $Y\subset\mathbf{R}^m$. Для того чтобы точка $y^{(0)}\in Y$ была оптимальной по Парето (слабо оптимальной по Парето) оценкой, достаточно, чтобы она являлась точкой максимума на множестве Y функции $F(y)$, возрастающей по отношению $\geq$ (по отношению $>$).*

□ Доказательство проведем для парето-оптимальных оценок, так как для слабо оптимальных по Парето оценок рассуждения аналогичны.

Пусть, по условию, $y^{(0)}\in Y$ и

$$F(y^{(0)})\geqslant F(y) \text{ для всех } y\in Y. \tag{7.4}$$

Предположим противное, т. е. что для некоторой оценки $y'\in Y$ верно неравенство $y'\geq y^{(0)}$. Отсюда, поскольку функция F возрастающая, получаем неравенство $F(y')>F(y^{(0)})$, которое противоречит (7.4). ■

Теорема 7.2 сформулирована в терминах оценок. Однако ее можно переформулировать и в терминах решений. Так, если функция $F(y)$ возрастает по отношению $\geq$ и $x^{(0)}$ — точка максимума функции $F[f(x)]=F[f_1(x), f_2(x), \dots, f_m(x)]$ на множестве X, то $f(x^{(0)})\in P(Y)$, а значит, и $x^{(0)}\in P_f(X)$.

Приведем примеры различных возрастающих функций многих переменных.

1) Функция $F(y_1, y_2, \dots, y_m)=\sum\limits_{i=1}^{m}\mu_i y_i$ является возрастающей по отношению $\geq$ на всем пространстве $\mathbf{R}^m$, если все числа $\mu_1, \mu_2, \dots, \mu_m$

положительны, и возрастающей по отношению $>$, когда все указанные числа не отрицательны и хотя бы одно из них положительно. Следовательно, согласно теореме 7.2, максимизация функции $\sum_{i=1}^{m} \mu_i f_i(x)$ на множестве X в первом случае приводит к парето-оптимальному решению, а во втором — к решению, слабо оптимальному по Парето. В частности, при $\mu_i=1$ и $\mu_j=0$, $j=1, 2, \ldots, m$; $j\neq i$, получаем, что всякая точка максимума функции $f_i(x)$ на множестве X является слабо оптимальным по Парето решением.

2) Рассмотрим функцию $F(y_1, y_2, \ldots, y_m)=\prod_{i=1}^{m} y_i^{\mu_i}$ при положительных фиксированных числах $\mu_1, \mu_2, \ldots, \mu_m$. Пусть $y\geq y'>0_m$. Тогда справедливо неравенство $y_i^{\mu_i}\geqslant (y_i')^{\mu_i}$ для всех $i=1, 2, \ldots, m$, причем хотя бы для одного i имеет место строгое неравенство. Следовательно,

$$\prod_{i=1}^{m} y_i^{\mu_i} > \prod_{i=1}^{m} (y_i')^{\mu_i},$$

а значит, функция $\prod_{i=1}^{m} y_i^{\mu_i}$ является возрастающей по отношению $\geq$ на множестве $\mathbf{R}^m_> =\{y\in R^m \mid y>0_m\}$. Аналогично можно проверить, что при неотрицательных фиксированных числах $\mu_1, \mu_2, \ldots, \mu_m$, среди которых есть хотя бы одно положительное, указанная функция является возрастающей на $\mathbf{R}^m_>$ по отношению $>$.

3) Рассмотрим функцию $F(y_1, y_2, \ldots, y_m)=\min_{i=1,2,\ldots,m} \mu_i y_i$ при фиксированных положительных числах $\mu_1, \mu_2, \ldots, \mu_m$. Если $y_i>y'_i$, $i=1, 2, \ldots, m$, то и $\mu_i y_i>\mu_i y'_i$, $i=1, 2, \ldots, m$. Отсюда следует неравенство

$$\min_i \mu_i y_i > \min_i \mu_i y_i',$$

свидетельствующее о том, что рассматриваемая функция является возрастающей по отношению $>$ на $\mathbf{R}^m$.

4) Пусть $y\geq y'$. Следовательно, верно неравенство $y^*-y'\geq y^*-y$. При $y^*_i\geqslant \sup_{y\in Y} y_i$, $i=1, 2, \ldots, m$, отсюда имеем

$$-\left[\sum_{i=1}^{m} (y_i^*-y_i)^2\right]^{1/2} > -\left[\sum_{i=1}^{m} (y_i^*-y_i')^2\right]^{1/2}.$$

Значит, функция $F(y_1, y_2, \ldots, y_m)=-\left[\sum_{i=1}^{m} (y_i^*-y_i)^2\right]^{1/2}$ является возрастающей по отношению $\geq$ и, по теореме 7.2, минимизация функции $\sum_{i=1}^{m}(y_i^*-y_i)^2$ на множестве Y приводит к парето-оптимальным оценкам.

Приведем необходимые условия оптимальности по Парето и слабой оптимальности по Парето. Заметим, что в силу включений (7.3) любые необходимые условия слабой оптимальности по Парето являются также и необходимыми условиями парето-оптимальности.

Теорема 7.3. *Пусть множество X выпукло и все компоненты $f_1, f_2, \dots, f_m$ вектор-функции f вогнуты на множестве X. Для того чтобы решение $x^{(0)} \in X$ было слабо оптимальным по Парето (а значит, и парето-оптимальным), необходимо, чтобы существовали числа $\mu_i \geqslant$*

$$\geqslant 0,\ i=1, 2, \dots, m, \quad \sum_{i=1}^{m}\mu_i = 1 \text{ такие, что}$$

$$\sum_{i=1}^{m}\mu_i f_i(x^{(0)}) = \max_{x\in X}\sum_{i=1}^{m}\mu_i f_i(x).$$

Доказательство теоремы можно найти в [18].

При $m=2$ утверждение теоремы 7.3 допускает простую геометрическую интерпретацию в критериальном пространстве (рис. 7.5). Для всякой слабо оптимальной по Парето оценки $y^{(0)}$ можно подобрать вектор $\mu=(\mu_1, \mu_2) \geq 0_2$, $\mu_1+\mu_2=1$, такой, что линейная функция $\langle\mu, y\rangle$ достигает максимума на множестве Y в точке $y^{(0)}$. На рис. 7.5 вектор μ приближен к точке $y^{(0)}$.

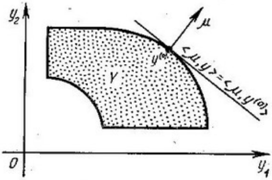

Рис. 7.5

Требование вогнутости функций $f_1, f_2, \dots, f_m$ в последней теореме является существенным. В этом убеждает следующий простой пример. Пусть $n=1$, $m=2$, $X=[0, 1]$, $f_1(x)=x$, $f_2(x)=e^{-x}$. Функция f_1 возрастает, а функция f_2 убывает, поэтому $S_f(X)=P_f(X)=[0, 1]$. Рассмотрим слабо оптимальную по Парето точку $x^{(0)}$ внутри отрезка $[0, 1]$ и рассмотрим функцию $\mu_1 x+\mu_2 e^{-x}$ при фиксированных неотрицательных числах μ_1, μ_2, одно из которых положительно. Если $\mu_2=0$, то невозможно подобрать такое число $\mu_1>0$, чтобы функция $\mu_1 x$ достигала максимального значения во внутренней точке отрезка $[0, 1]$. Пусть $\mu_2>0$. Тогда вторая производная функции $\mu_1 x+\mu_2 e^{-x}$ есть $\mu_2 e^{-x}>0$; следовательно, указанная функция строго выпукла и ни для каких чисел $\mu_1 \geqslant 0$, $\mu_2>0$ не может иметь точкой максимума внутреннюю точку $x^{(0)} \in (0, 1)$.

Теорема 7.4. *Пусть множество X выпукло, компоненты $f_1, f_2, \ldots, f_m$ вектор-функции f вогнуты на X и положительны, т. е. $f_i(x)>0$ для всех $x \in X$, $i=1, 2, \ldots, m$. Для того чтобы решение $x^{(0)} \in X$ было слабо оптимальным по Парето (и парето-оптимальным), необходимо, чтобы существовали такие числа $\mu_i \geqslant 0$, $i=1, 2, \ldots, m$, $\sum_{i=1}^{m} \mu_i = 1$, чтобы было справедливо равенство*

$$\prod_{i=1}^{m} [f_i(x^{(0)})]^{\mu_i} = \max_{x \in X} \prod_{i=1}^{m} [f_i(x)]^{\mu_i}.$$

Доказательство теоремы см., например, в [18].

В теореме 7.5 в отличие от теорем 7.3 и 7.4 отсутствуют предположения о выпуклости множества X и вогнутости компонент вектор-функции f. Поэтому ее можно применять при решении широкого класса многокритериальных задач.

Теорема 7.5. *Предположим, что $x^{(0)} \in X$ и $f(x^{(0)}) > 0_m$. Для того чтобы решение $x^{(0)}$ было слабо оптимальным по Парето (и парето-оптимальным), необходимо, чтобы нашлись такие числа $\mu_i > 0$, $i=1, 2, \ldots, m$, $\sum_{i=1}^{m} \mu_i = 1$, что имеет место равенство*

$$\min_{i=1,2,\ldots,m} \mu_i f_i(x^{(0)}) = \max_{x \in X} \min_{i=1,2,\ldots,m} \mu_i f_i(x).$$

□ Доказательство носит конструктивный характер, т. е. для каждого слабо оптимального по Парето решения следует указать соответствующий ему набор чисел $\mu_1, \mu_2, \ldots, \mu_m$. Пусть $x^{(0)} \in S_f(X)$. Это означает, что не существует вектора $x \in X$, для которого выполнялось бы неравенство $f_i(x) > f_i(x^{(0)})$, $i=1, 2, \ldots, m$. Следовательно, для любого $x \in X$ найдется такой номер $i \in \{1, 2, \ldots, m\}$, при котором справедливо неравенство $f_i(x^{(0)}) \geqslant f_i(x)$.

Возьмем в качестве $\mu_1, \mu_2, \ldots, \mu_m$ следующие числа:

$$\mu_i = \frac{1}{f_i(x^{(0)})} \Bigg/ \sum_{k=1}^{m} \frac{1}{f_k(x^{(0)})} > 0, \quad i=1, 2, \ldots, m.$$

Очевидно, что $\mu_1 + \mu_2 + \ldots + \mu_m = 1$.

Из неравенства $f_i(x^{(0)}) \geqslant f_i(x)$ получаем

$$\mu_i f_i(x^{(0)}) \geqslant \mu_i f_i(x). \tag{7.5}$$

Следовательно, для каждого вектора $x \in X$ существует номер $i \in \{1, 2, \ldots, m\}$, при котором выполняется неравенство (7.5). Число

$$\mu_i f_i(x^{(0)}) = \left[\sum_{k=1}^{m} \frac{1}{f_k(x^{(0)})} \right]^{-1}$$

одинаковое для всех номеров $i=1, 2, \ldots, m$, поэтому можно записать равенство

$$\mu_i f_i(x^{(0)}) = \min_{i=1,2,\ldots,m} \mu_i f_i(x^{(0)}).$$

Учитывая это равенство, из (7.5) получаем неравенство

$$\min_{i=1,2,\ldots,m} \mu_i f_i(x^{(0)}) \geqslant \min_{i=1,2,\ldots,m} \mu_i f_i(x) \text{ для любого } x \in X. \blacksquare$$

Теоремы 7.2—7.5 показывают, что при определенных предположениях, подбирая коэффициенты $\mu_1, \mu_2, \ldots, \mu_m$, **любое** слабо оптимальное по Парето (и парето-оптимальное) решение можно получить в результате решения соответствующей однокритериальной задачи максимизации. Аналогичный вывод справедлив и для оценок.

6. О существовании и нахождении парето-оптимальных оценок и решений. Рассмотрим функцию $\sum_{i=1}^{m} f_i(x)$, где функции $f_1, f_2, \ldots, f_m$ непрерывны на множестве X. Сумма непрерывных функций также непрерывная функция, поэтому если множество X не пусто, замкнуто и ограничено, то, по теореме Вейерштрасса, существует точка $x^{(0)} \in X$, доставляющая максимум функции $\sum_{i=1}^{m} f_i(x)$ на множестве X. Согласно теореме 7.2, верно включение $x^{(0)} \in P_f(X)$, а значит, и включение $x^{(0)} \in S_f(X)$. Таким образом, доказано следующее утверждение.

Теорема 7.6. *Пусть компоненты вектор-функции $f(x) = (f_1(x), f_2(x), \ldots, f_m(x))$ непрерывны на непустом компактном множестве X. Тогда $P_f(X) \neq \varnothing$, $S_f(X) \neq \varnothing$.*

Следствие 7.2. *Если $Y \subset \mathbf{R}^m$ — непустое компактное множество, то $P(Y) \neq \varnothing$ и $S(Y) \neq \varnothing$.*

☐ Рассмотрим линейную вектор-функцию вида $(y_1, y_2, \ldots, y_m)$. Ее компоненты непрерывны на $\mathbf{R}^m$, и, в частности, на Y. По теореме 7.6, существует точка $y^{(0)} \in Y$, являющаяся парето-оптимальной относительно этой линейной вектор-функции на множестве Y. Ясно, что $y^{(0)} \in P(Y)$. Следовательно, $P(Y) \neq \varnothing$, а значит, и $S(Y) \neq \varnothing$. ■

Из доказательства теоремы 7.6 вытекает следующий способ отыскания какого-нибудь одного парето-оптимального решения (если, конечно, предположения теоремы выполнены): найти точку, реализующую максимальное значение функции $\sum_{i=1}^{m} f_i(x)$ на множестве X.

Найти все парето-оптимальные оценки (решения) значительно сложнее. Если множество оценок Y конечно, то найти все множество $P(Y)$ можно с помощью алгоритма, описанного в доказательстве теоремы 7.1, используя вместо отношения $>$ отношение $\geq$. Если множество Y бесконечно, но $Y \subset \mathbf{R}^2$, то часто множество $P(Y)$ удается найти геометрически, предварительно изобразив на плоскости все точки множества Y (см. рис. 7.3). Аналогично можно поступать и при нахождении множества $S(Y)$. Если же количество критериев больше двух и Y — бесконечное множество, то наглядно изобразить множество Y трудно. Поэтому в таких случаях геометрически найти все элементы множества $P(Y)$ (или $S(Y)$), как правило, не удается.

§ 7.5. ВЫБОР РЕШЕНИЯ ПРИ НАЛИЧИИ ДОПОЛНИТЕЛЬНЫХ СВЕДЕНИЙ ОБ ОТНОШЕНИИ ПРЕДПОЧТЕНИЯ

1. Проблема сужения множества Парето. Как отмечалось в предыдущем параграфе, в многокритериальной задаче оптимальную оценку (оптимальное решение) следует выбирать среди элементов множества $P(Y)$ $(P_f(X))$. Если множество $P(Y)$ сравнительно «узкое» (в частности, если $P(Y)$ содержит единственный элемент), то в качестве искомой оптимальной оценки можно выбрать любую парето-оптимальную оценку. В таких случаях можно считать, что отношение строгого предпочтения $\succ$ совпадает с отношением $\geq$ и поэтому $\mathrm{opt}_{\succ} Y = P(Y)$.

Однако часто на практике множество $P(Y)$ оказывается достаточно «широким», так что для лица, принимающего решение, небезразлично, какую именно оценку выбрать из $P(Y)$. Такое положение свидетельствует о том, что отношения $\succ$ и $\geq$ хотя и связаны аксиомой Парето, однако не совпадают. Иначе говоря, разница между множествами $\mathrm{opt}_{\succ} Y$ и $P(Y)$ значительная и оценка сверху $\mathrm{opt}_{\succ} Y \subset P(Y)$ является слишком «грубой» и требует уточнения. Обоснованное уточнение оценки сверху или, как говорят, *сужение множества парето-оптимальных оценок* возможно лишь при использовании дополнительных сведений об отношении предпочтения $\succ$. В конкретных ситуациях дополнительные сведения могут быть заданы в различном виде. Иногда они вообще отсутствуют и их получают в результате анализа отношения предпочтения $\succ$, которое используется в данной задаче.

2. Выбор решения в случае существования функции ценности. Пусть $\succ$ — некоторое асимметричное и транзитивное отношение строгого предпочтения на множестве оценок $Y \subset \mathbf{R}^m$. Числовую функцию Φ переменных $y_1, y_2, \ldots, y_m$ называют *функцией ценности для отношения* $\succ$, если для произвольных векторов $y, y' \in Y$ неравенство $\Phi(y) > \Phi(y')$ имеет место тогда и только тогда, когда $y \succ y'$.

Предполагаем, что отношение $\succ$ удовлетворяет аксиоме Парето. Поэтому из неравенства $y \geq y'$ следует соотношение $y \succ y'$, а значит, $\Phi(y) > \Phi(y')$. Следовательно, *функция ценности (если она существует) является возрастающей по отношению* $\geq$.

Если $\Phi(y)$ — функция ценности для отношения $\succ$ и h — возрастающая функция одной переменной, то $h[\Phi(y)]$ также функция ценности. Таким образом, *функция ценности определяется с точно-*

стью до возрастающего преобразования h. В частности, всякая функция $\alpha\Phi(y)+a$, где $\alpha>0$ и $a\in\mathbf{R}$, — функция ценности, если таковой является функция Φ.

По определению отношения неразличимости $\sim$, отношение $y\sim\sim y'$ выполняется в том и только том случае, если не верно ни соотношение $y\succ y'$, ни соотношение $y'\succ y$. Поэтому если имеет место равенство $\Phi(y)=\Phi(y')$, то не может быть ни $\Phi(y)>\Phi(y')$, ни $\Phi(y')>\Phi(y)$, а значит, верно соотношение $y\sim y'$. Справедливо и обратное, т. е. из соотношения $y\sim y'$ всегда вытекает равенство $\Phi(y)=\Phi(y')$.

Используя функцию ценности, вопрос сравнения по предпочтительности векторных оценок y и y' можно свести к сравнению соответствующих чисел $\Phi(y)$ и $\Phi(y')$: более предпочтительной оценке соответствует бóльшее значение фукции ценности, а двум оценкам, находящимся в отношении неразличимости, отвечают равные значения функции ценности.

Если для отношения $\succ$ существует функция ценности Φ, то очевидно, что

$$\operatorname{opt}_{\succ} Y=\{y^{(0)}\in Y|\Phi(y^{(0)})=\max_{y\in Y}\Phi(y)\}$$

и *отыскание оптимальной оценки сводится к решению однокритериальной задачи максимизации функции Φ на множестве Y.*

Необходимым условием существования функции ценности является транзитивность отношения $\sim$. Это следует непосредственно из определения функции ценности и транзитивности отношения равенства $=$.

Отношение $\geq$ на $\mathbf{R}^m$ порождает нетранзитивное отношение неразличимости (см. п. 1, § 7.3), поэтому для отношения $\geq$ на пространстве $\mathbf{R}^m$ функции ценности не существует. Однако в частных случаях, например если множество Y конечно и имеет вид

$$Y=\{y^{(1)},\ y^{(2)},\dots,\ y^{(l)}\in\mathbf{R}^m|y^{(l)}\geq y^{(l-1)}\geq\dots\geqq y^{(1)}\},$$

функция ценности существует; это, например, функция вида $\Phi(y^{(i)})=i$, $i=1,\ 2,\dots,l$. Таким образом, существование функции ценности зависит и от структуры множества Y.

Вопросы существования функции ценности подробно изучаются в книге П. Фишберна «Теория полезности для принятия решений» (М.: Наука, 1976).

Если отношение $\sim$ транзитивно, то его называют *отношением безразличия* и обозначают $\approx$. Оценки y и y', для которых верно соотношение $y\approx y'$, одинаково предпочтительны (безразличны) в том смысле, что если найдется оценка y^* строго предпочтительнее одной из них, то она строго предпочтительнее и другой оценки, т. е. из соотношения $y^*\succ y$, $y\approx y'$ всегда следует соотношение $y^*\succ y'$. В этом можно легко убедиться, рассуждая «от противного». Заметим, что для нетранзитивного отношения $\sim$ подобное утверждение, вообще говоря, неверно; в подтверждение этого достаточно в качестве $\succ$ взять отношение $\geq$.

Таким образом, если функция ценности Φ существует, то все оценки y, имеющие одно и то же фиксированное значение $\Phi(y)$, являются безразличными для лица, принимающего решение. Говорят, что такие оценки составляют *класс безразличия.* Каждому значению функции ценности соответствует свой класс безразличия. При $m=2$ класс безразличия представляет собой некоторую линию уровня функции ценности (рис. 7.6). На этом рисунке каждая оценка из класса с бо́льшим номером строго предпочтительнее любой оценки с меньшим номером: $y \succ y'$.

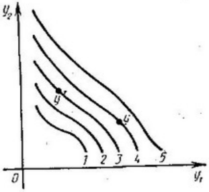

Рис. 7.6

Как отмечалось выше, в случае существования функции ценности Φ оптимальная оценка находится максимизацией этой функции на множестве Y. Ответ на вопрос о том, как построить функцию ценности, если известно, что она существует, можно найти в [13]. Укажем лишь, что функция Φ нередко допускает *аддитивное представление:*

$$\Phi(y_1, y_2, \ldots, y_m) = \sum_{i=1}^{m} v_i(y_i),$$

где возрастающие функции $v_i(y_i)$, $i=1, 2, \ldots, m$, одной переменной определяют, анализируя отношение строгого предпочтения. В частности, аддитивное представление может иметь и следующий линейный вид:

$$\Phi(y_1, y_2, \ldots, y_m) = \sum_{i=1}^{m} \mu_i y_i,$$

где $\mu_1, \mu_2, \ldots, \mu_m$ — некоторые положительные числа, подлежащие определению.

Учитывая связь между отношениями $\underset{Y}{\succ}$ и $\underset{X}{\succ}$, все сформулированные выше определения и выводы можно перенести на случай отношения $\underset{X}{\succ}$. Так, например, в случае существования функции ценности Φ множество оптимальных решений имеет вид

$$\text{opt}_{\succ} X = \{x^{(0)} \in X \mid \Phi[f(x^{(0)})] = \max_{x \in X} \Phi[f(x)]\}.$$

3. Выбор решения при строго упорядоченных по важности критериях. Часто для лица, принимающего решение, желательно получить возможно большее значение, например критерия f_1, даже за счет «потерь» по остальным критериям, т. е. критерий f_1 оказывается более важным, чем остальные. Возможен и случай, когда весь набор критериев $f_1, f_2, \ldots, f_m$ строго упорядочен по важности, так что критерий f_1 более важен, чем все остальные критерии f_2,

$f_3, \dots, f_m$, критерий f_2 более важен, чем все критерии $f_3, \dots, f_m$, и т. д. Это соответствует ситуации, когда при сравнении оценок и решений используется лексико-графическое отношение. Приведем определение этого отношения.

Пусть имеются два вектора $y, y' \in Y \subset \mathbf{R}^m$. *Лексико-графическое отношение* $\overset{\text{lex}}{\underset{Y}{\succ}}$ *определяется следующим образом: отношение* $y \overset{\text{lex}}{\underset{Y}{\succ}} y'$ *имеет место тогда и только тогда, когда выполнено одно из следующих условий:*

$$
\begin{aligned}
&1)\ y_1 > y_1',\\
&2)\ y_1 = y_1',\ y_2 > y_2', \qquad\qquad (7.6)\\
&\ldots\ldots\ldots\ldots\\
&m)\ y_i = y_i',\ \ i = 1, 2, \dots, m-1;\ y_m > y_m',
\end{aligned}
$$

где $y = (y_1, y_2, \dots, y_m)$, $y' = (y_1', y_2', \dots, y_m')$. При $m = 1$ лексико-графическое отношение совпадает с отношением $>$ на подмножестве вещественных чисел.

Если выполнено соотношение $y \overset{\text{lex}}{\succ} y'$, то говорят, что вектор *y лексико-графически больше, чем вектор y'.*

Геометрическая иллюстрация отношения $\overset{\text{lex}}{\succ}$ приведена на рис. 7.7.

Если лицо, принимающее решение, в качестве отношения строгого предпочтения использует лексико-графическое отношение, то это означает, что из пары оценок для него предпочтительнее та, первая компонента которой больше (независимо от соотношений между остальными компонентами).

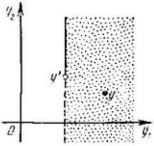

Рис. 7.7

Если первые компоненты двух оценок одинаковы, то для лица, принимающего решение, предпочтительнее оценка, имеющая бóльшую вторую компоненту; остальные компоненты данной оценки могут при этом «значительно уступать» соответствующим компонентам второй оценки и т. д. В таких случаях говорят, что компоненты $y_1, y_2, \dots, y_m$ (т. е. критерии $f_1, f_2, \dots, f_m$) *строго упорядочены по важности*: самым важным является критерий f_1, менее важным f_2, еще менее важным — f_3 и т. д.; наконец, критерий f_m имеет наименьшую важность.

В определении лексико-графического отношения важную роль играет порядок перечисления критериев. Изменение нумерации критериев приводит к другому лексико-графическому отношению.

Теорема 7.7. *Отношение* $\overset{\text{lex}}{\succ}$ *асимметрично, транзитивно и удовлетворяет аксиоме Парето.*

□ Установим транзитивность лексико-графического отношения. Пусть $y \overset{\text{lex}}{\succ} y^*$ и $y^* \overset{\text{lex}}{\succ} y'$, причем, согласно первому соотношению, имеет место условие типа (7.6) с некоторым номером $k \in \{1, 2, \dots, m\}$, а согласно второму — условие типа (7.6) с некоторым номером $l \in \{1, 2, \dots, m\}$. Положим $n = \min\{k, l\}$. Тогда для оценок y и y' выполнено условие из (7.6) с номером n. Следовательно, $y \overset{\text{lex}}{\succ} y'$ и транзитивность доказана.

Так как соотношение $y \overset{\text{lex}}{\succ} y$ не может быть выполнено ни для какого y, то лексико-графическое отношение иррефлексивно.

В п. 2 § 7.2 указывалось, что произвольное транзитивное и иррефлексивное отношение всегда асимметрично. Поэтому и отношение $\overset{\text{lex}}{\succ}$ асимметрично.

Если верно неравенство $y \geq y'$, то одно из условий в (7.6) будет выполнено. Поэтому неравенство $y \geq y'$ влечет выполнение соотношения $y \overset{\text{lex}}{\succ} y'$, т. е. лексико-графическое отношение удовлетворяет аксиоме Парето. ■

Пусть имеются векторы $y, y' \in Y$ и $y \neq y'$. Тогда справедливо либо соотношение $y \overset{\text{lex}}{\succ} y'$, либо соотношение $y' \overset{\text{lex}}{\succ} y$. Следовательно, когда не выполнено ни одно из этих двух соотношений, то $y = y'$. Поэтому для произвольной пары оценок y, y' имеет место одно и только одно из следующих трех соотношений: $y \overset{\text{lex}}{\succ} y'$, $y' \overset{\text{lex}}{\succ} y$, $y = y'$.

Лексико-графическое отношение на множестве оценок порождает отношение $\underset{X}{\overset{\text{lex}}{\succ}}$ на множестве решений X: $x \underset{X}{\overset{\text{lex}}{\succ}} x'$ тогда и только тогда, когда $y \underset{Y}{\overset{\text{lex}}{\succ}} y'$, где $y = f(x)$, $y' = f(x')$. Очевидно, отношение $\underset{X}{\overset{\text{lex}}{\succ}}$ также является асимметричным, транзитивным и удовлетворяет аксиоме Парето (в терминах решений).

Множество решений (оценок), оптимальных по отношению $\underset{X}{\overset{\text{lex}}{\succ}}$ на множестве X (соответственно Y), называют *множеством лексико-графически оптимальных* (точнее говоря, *лексико-графически максимальных*) *решений (оценок)* и обозначают через $\operatorname{opt}_{\underset{\succ}{\text{lex}}} X$ (соответственно $\operatorname{opt}_{\underset{\succ}{\text{lex}}} Y$).

Так как для любых двух векторов y, y' либо один лексико-графически больше другого, либо они равны, то *множество* $\operatorname{opt}_{\underset{\succ}{\text{lex}}} Y$, *если оно не пусто, состоит из единственного элемента.*

Для множества Y, изображенного на рис. 7.3, лексико-графически оптимальной оценкой является точка e.

Непосредственно из теорем 7.1 и 7.7 вытекает следующее следствие.

Следствие 7.3. *Если множество Y состоит из конечного числа элементов, то лексико-графически оптимальная оценка существует и единственна.*

Введем множества, определяемые рекуррентным способом:

$$X_1 = \{x^{(1)} \in X \mid f_1(x^{(1)}) = \max_{x \in X} f_1(x)\},$$

$$X_2 = \{x^{(2)} \in X_1 \mid f_2(x^{(2)}) = \max_{x \in X_1} f_2(x)\},$$

$$\dots\dots\dots\dots\dots\dots\dots$$

$$X_m = \{x^{(m)} \in X_{m-1} \mid f_m(x^{(m)}) = \max_{x \in X_{m-1}} f_m(x)\}.$$

X_1 — это множество всех точек максимума первой компоненты вектор-функции f, т. е. функции f_1, на множестве X. X_2 — это множество всех точек максимума f_2 — второй компоненты вектор-функции f — на множестве X_1 и т. д.

По определению, каждое последующее множество является подмножеством предыдущего, т. е. имеют место включения

$$X_m \subset X_{m-1} \subset \dots \subset X_2 \subset X_1 \subset X. \tag{7.7}$$

Теорема 7.8. *Для того чтобы решение $x^{(0)} \in X$ было лексико-графически-оптимальным, необходимо и достаточно, чтобы выполнялось включение $x^{(0)} \in X_m$.*

□ Необходимость. Пусть выполнено $x^{(0)} \in \mathrm{opt}_{\underset{\succ}{\mathrm{lex}}} X$. Предположим противное: $x^{(0)} \notin X_m$. Последнее означает, что либо выполнено $x^{(0)} \in X_{m-1}$ и $f_m(x^{(0)}) < \max_{X_{m-1}} f_m(x)$, либо $x^{(0)} \notin X_{m-1}$. В первом случае найдется такое решение $x' \in X_{m-1}$, что $f_i(x') = f_i(x^{(0)})$, $i=1, 2, \dots, m-1$, $f_m(x') > f_m(x^{(0)})$. Следовательно, верно соотношение $x' \overset{\mathrm{lex}}{\succ} x^{(0)}$, что противоречит лексико-графической оптимальности решения $x^{(0)}$. Поэтому выполняется $x^{(0)} \notin X_{m-1}$. Отсюда следует, что либо имеет место $x^{(0)} \in X_{m-2}$ и $f_{m-1}(x^{(0)}) < \max_{X_{m-2}} f_{m-1}(x)$, либо справедливо $x^{(0)} \notin X_{m-2}$. Рассуждая, как и выше, придем к соотношению $x^{(0)} \notin X_{m-2}$ и т. д. Окончательно получаем противоречие: $x^{(0)} \notin X$.

Достаточность. Рассмотрим решение $x^{(0)} \in X_m$ и допустим противное: существует решение $x \in X$, для которого справедливо соотношение $x \overset{\mathrm{lex}}{\succ} x^{(0)}$. Отсюда следует соотношение $y \overset{\mathrm{lex}}{\succ} y^{(0)}$, где $y = f(x)$, $y^{(0)} = f(x^{(0)})$. Пусть согласно соотношению $y \overset{\mathrm{lex}}{\succ} y^{(0)}$ выполняется k-е условие из (7.6), $k \in \{1, 2, \dots, m-1\}$. Тогда $x^{(0)} \in X_k$, а значит, в силу (7.7) имеем $x^{(0)} \in X_m$. ■

Следствие 7.4. *Если все функции $f_1, f_2, \dots, f_m$ непрерывны на непустом компактном множестве $X \subset \mathbb{R}^n$, то справедливо неравенство $\mathrm{opt}_{\underset{\succ}{\mathrm{lex}}} X \neq \varnothing$.*

□ По теореме Вейерштрасса (см. § 2.2), множество X_1 не пусто и компактно. Тогда таким же свойством обладает и множество X_2 и т. д. до X_{m-1}. Множество X_m не пусто в силу того, что X_{m-1} не пусто и компактно, а функция f_m непрерывна. В таком случае, по теореме 7.8, множество лексико-графически-оптимальных решений также не пусто. ■

Теорема 7.8 дает следующий поэтапный метод нахождения лексико-графически-оптимального решения. Сначала находят множество точек максимума функции f_1 на множестве X, т. е. множество X_1. Далее на этом множестве максимизируют функцию f_2 и определяют множество X_2 и т. д. до множества X_{m-1}. Наконец, максимизируя f_m на множестве X_{m-1}, находят лексико-графически-оптимальное решение.

С точки зрения вычислений описанный метод сложен, поскольку на каждом k-м этапе (кроме последнего) нужно целиком строить множество X_k. Удобнее для нахождения лексико-графически-оптимального решения решать такую последовательность задач:

1) найти $f^*_1=\max f_1(x)$ при условии $x\in X$;

2) найти $f^*_2=\max f_2(x)$ при условиях $x\in X$, $f_1(x)=f^*_1$;

. . . .

$m-1$) найти $f^*_{m-1}=\max f_{m-1}(x)$ при условиях $x\in X$, $f_i(x)=f^*_i$, $i=1, 2, \dots, m-2$;

m) найти точку максимума функции $f_m(x)$ при условиях $x\in X$, $f_i(x)=f^*_i$, $i=1, 2, \dots, m-1$.

4. Оценка сверху для множества оптимальных решений в условиях отношения предпочтения, инвариантного относительно перенумерации критериев. Лексико-графическое отношение не является инвариантным относительно перенумерации критериев. Однако на практике встречаются задачи, в которых для лица, принимающего решение, несущественно, в каком порядке перечисляются компоненты $y_1, y_2, \dots, y_m$ оценки $y\in\mathbf{R}^m$, а важны только числовые значения этих компонент. В таком случае если одна из оценок предпочтительнее другой, то и каждая оценка, полученная из первой перестановкой компонент, является более предпочтительной, чем оценка, образованная из второй оценки произвольной перестановкой компонент. Такие дополнительные сведения можно использовать при построении оценки сверху для искомого множества оптимальных оценок и тем самым сузить множество парето-оптимальных оценок.

Пусть имеется вектор $y=(y_1, y_2, \dots, y_m)$. Множество векторов, которое состоит из y и всех векторов, получающихся из y перестановкой его компонент, обозначим через $\Pi(y)$. Очевидно, множество $\Pi(y)$ содержит не более чем $m!$ элементов. В частном случае при $y_1=y_2=\dots=y_m$ выполняется равенство $\Pi(y)=\{y\}$. Элементы множества $\Pi(y)$ будем обозначать через $\pi(y)$, $\pi'(y)$ и т. п.

Далее считаем, что отношение строгого предпочтения $\succ$ определено на множестве $\bigcup\limits_{y\in Y}\Pi(y)\subset\mathbf{R}^m$. Кроме того, как и раньше, предполагаем, что это отношение асимметрично, транзитивно и удовлетворяет аксиоме Парето.

Отношение $\succ$ называют *инвариантным относительно перенумерации критериев*, если для произвольных векторов y, $y' \in Y$ из соотношения $y \succ y'$ всегда следует соотношение $\pi(y) \succ \pi'(y')$ для любых векторов $\pi(y) \in \Pi(y)$, $\pi'(y') \in \Pi(y')$.

Отношение $\geq$ на пространстве $\mathbf{R}^m$ не является инвариантным относительно перенумерации критериев, так как, например, из неравенства $(2,3)^{\mathrm{T}} \geq (1,3)^{\mathrm{T}}$ не следует справедливость неравенства $(2,3)^{\mathrm{T}} \geq (3,1)^{\mathrm{T}}$. Если для отношения $\succ$ существует функция ценности вида $F(y_1, y_2, \dots, y_m) = \sum_{i=1}^{m} v(y_i)$, то, очевидно, такое отношение инвариантно относительно перенумерации критериев.

Далее *будем считать, что об отношении строгого предпочтения $\succ$ известно, что оно является инвариантным относительно перенумерации критериев.* Оценку сверху для множества $\operatorname{opt}_{\succ} Y$ получим с помощью симметрического отношения.

Симметрическое отношение $\overset{S}{\succ}$ на множестве $\bigcup_{y \in Y} \Pi(y)$ определяется следующим образом: соотношение $y \overset{S}{\succ} y'$ выполняется тогда и только тогда, когда для некоторого $\pi(y') \in \Pi(y')$ верно неравенство $y \geq \pi(y')$.

Геометрическая иллюстрация симметрического отношения при $m=2$ приведена на рис. 7.8. Точка y, для которой верно соотношение $y \overset{S}{\succ} y'$, принадлежит объединению двух заштрихованных углов, расположенных симметрично относительно биссектрисы координатного угла.

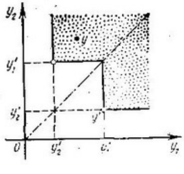

Рис. 7.8

Симметрическое отношение транзитивно и асимметрично. В самом деле, пусть согласно соотношениям $y \overset{S}{\succ} y^*$, $y^* \overset{S}{\succ} y'$ выполнены неравенства $y \geq \pi^*(y^*)$, $y^* \geq \pi'(y')$ для некоторых векторов $\pi^*(y^*) \in \Pi(y^*)$, $\pi'(y') \in \Pi(y')$. Переставим компоненты вектора y^* так, чтобы получился вектор $\pi^*(y^*)$. Аналогичная перестановка компонент вектора $\pi'(y')$ приведет к некоторому вектору $\pi''(y')$. В результате имеем неравенства $y \geq \pi^*(y^*) \geq \pi''(y')$, т. е. верно соотношение $y \overset{S}{\succ} y'$. Транзитивность установлена. Симметрическое отношение иррефлексивно. Действительно, выполнение соотношения $y \overset{S}{\succ} y'$ означает, что $y \geq \pi(y)$, откуда следует, что сумма компонент вектора y больше суммы компонент вектора $\pi(y)$, а это не может иметь места. Асимметричность является следствием транзитивности и иррефлексивности.

Симметрическое отношение инвариантно относительно перенумерации критериев и удовлетворяет аксиоме Парето, т. е. из неравенства $y \geq y'$ следует соотношение $y \overset{s}{\succ} y'$.

Теорема 7.9. *Справедливы соотношения*

$$\operatorname{opt}_{\succ} Y \subset \operatorname{opt}_{\overset{s}{\succ}} Y = Y \cap P\Big(\bigcup_{y \in Y} \Pi(y)\Big), \tag{7.8}$$

где $\operatorname{opt}_{\overset{s}{\succ}} Y$ — множество оптимальных оценок по отношению $\overset{s}{\succ}$ на множестве Y, а $P\Big(\bigcup_{y \in Y} \Pi(y)\Big)$ — множество парето-оптимальных оценок на множестве $\bigcup_{y \in Y} \Pi(y)$.

□ Проверим включение из соотношения (7.8). Пусть имеется вектор $y^{(0)} \in \operatorname{opt}_{\succ} Y$. Предположим противное: для некоторого вектора $y \in Y$ верно соотношение $y \overset{s}{\succ} y^{(0)}$. Тогда $y \geq \pi(y^{(0)})$ для некоторого вектора $\pi(y^{(0)}) \in \Pi(y^{(0)})$. Отсюда, согласно аксиоме Парето, следует соотношение $y \succ \pi(y^{(0)})$ и, используя инвариантность отношения $\succ$ относительно перенумерации критериев, получаем $y \succ y^{(0)}$, что противоречит начальному допущению.

Теперь докажем справедливость включения $\operatorname{opt}_{\overset{s}{\succ}} Y \subset [Y \cap P\Big(\bigcup_{y \in Y_l} \Pi(y)\Big)$. Пусть имеется вектор $y^{(0)} \in \operatorname{opt}_{\overset{s}{\succ}} Y$. Предполагая противное, получаем $y^{(0)} \in Y$ и $y^{(0)} \notin P\Big(\bigcup_{y \in Y} \Pi(y)\Big)$. Следовательно, существует вектор $\pi(y) \in \Pi(y)$ для которого $y \in Y$ такой, что выполнено $\pi(y) \geq y^{(0)}$. Надлежащим образом, переставляя компоненты векторов, входящих в это неравенство, получаем неравенство $y \geq \pi^{\circ}(y^{(0)})$, где вектор $\pi^{\circ}(y^{(0)}) \in \Pi(y^{(0)})$. Последнее означает, что выполнено соотношение $y \overset{s}{\succ} y^{(0)}$. Это противоречит предположению о том, что $y^{(0)} \in \operatorname{opt}_{\succ} Y$, и тем самым справедливость включения доказана. Справедливость обратного включения устанавливается аналогично. ■

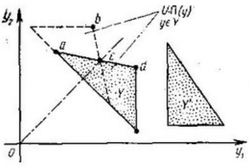

Рис. 7.9

При $m=2$ оценку сверху для множества $\operatorname{opt}_{\succ} Y$, т. е. множество $Y \cap P\Big(\bigcup_{y \in Y} \Pi(y)\Big)$, часто можно получить геометрически. Так, на рис. 7.9 изображены множества Y и $\bigcup_{Y \in y} \Pi(y)$

Здесь $P\Big(\bigcup_{y \in Y} \Pi(y)\Big)$ — это объединение отрезков bc и cd. Но множеству Y принадлежат только точки отрезка cd; они и дают искомую оценку сверху. Множество $P(Y)$ — это весь отрезок acd, так что использование информации об инвариантности отношения строгого предпочтения относительно перенумерации критериев его существенно сужает. Однако можно привести пример множества Y', для которого $P(Y')=\text{opt}_{\underset{\succ}{S}} Y'$ (рис. 7.9). В этом случае использование информации об инвариантности отношения строгого предпочтения к сужению множества парето-оптимальных оценок не приводит.

§ 7.6. ОЦЕНКА СВЕРХУ ДЛЯ МНОЖЕСТВА ОПТИМАЛЬНЫХ РЕШЕНИЙ В УСЛОВИЯХ ОТНОШЕНИЯ ПРЕДПОЧТЕНИЯ, ИНВАРИАНТНОГО ОТНОСИТЕЛЬНО ПОЛОЖИТЕЛЬНОГО ЛИНЕЙНОГО ПРЕОБРАЗОВАНИЯ

1. Отношение предпочтения, инвариантное относительно положительного линейного преобразования. В данном параграфе рассматриваются только отношения предпочтения, связанные со множеством оценок, поэтому условимся для краткости вместо $\underset{Y}{\succ}$ писать $\succ$. Заметим, что полученные результаты можно переформулировать в терминах решений.

Бинарное отношение строгого предпочтения $\succ$, заданное на пространстве $\mathbf{R}^m$, называют *инвариантным относительно положительного линейного преобразования*, если для произвольных векторов $y,\ y'\in\mathbf{R}_m$ из выполнения соотношения $y\succ y'$ следует справедливость соотношения $ky+c\succ ky'+c$ при всех $k>0$, $c\in\mathbf{R}^m$.

Инвариантность отношения $\succ$ относительно положительного линейного преобразования означает, что если оценка y предпочтительнее оценки y', то оценка вида ky предпочтительнее оценки вида ky' при всех $k>0$ и, кроме того, оценка вида $y+c$ предпочтительнее оценки вида $y'+c$ при каждом векторе $c\in\mathbf{R}^m$. Свойством инвариантности обладает, например, отношение $\geq$. Если для отношения строгого предпочтения существует линейная функция ценности $\Phi(y)=\sum_{i=1}^{m}\mu_i y_i$, где $\mu_1,\ \mu_2,\ldots,\ \mu_m$ — фиксированные числа, то отношение $\succ$ будет инвариантно относительно положительного линейного преобразования. Этим же свойством обладает лексико-графическое отношение. Однако симметрическое отношение $\overset{S}{\succ}$ не является инвариантным относительно положительного линейного преобразования. Например, имеет место соотношение $(1,3)^{\text{т}}\overset{S}{\succ}(2,1)^{\text{т}}$, но соотношение

$$(1,3)^{\text{т}}+(-1,0)^{\text{т}}=(0,3)^{\text{т}}\overset{S}{\succ}(1,1)^{\text{т}}=(2,1)^{\text{т}}+(-1,0)^{\text{т}}$$

не верно.

2. Основные требования к отношению предпочтения. Дальнейшее изложение ограничено случаем двух критериев: $m=2$.

В этом параграфе считаем, что отношение строгого предпочтения данного лица, принимающего решение, удовлетворяет следующим требованиям.

Требование 1 (иррефлексивность). *Ни для какого вектора $y \in \mathbb{R}^2$ не может быть выполнено соотношение $y \succ y$.*

Это требование естественно для отношения строгого предпочтения.

Требование 2 (аксиома Парето). *Для любых векторов $y, y' \in \mathbb{R}^2$ из выполнения соотношения $y \geq y'$ следует выполнение соотношения $y \succ y'$.*

Это требование свидетельствует о том, что дальнейшее рассмотрение ограничено задачами многокритериальной максимизации.

Требование 3 (транзитивность). *Для любых векторов y, y', $y'' \in \mathbb{R}^2$ из выполнения соотношений $y \succ y'$, $y' \succ y''$ следует выполнение соотношения $y \succ y''$.*

Требование 4 (инвариантность относительно положительного линейного преобразования). *Для произвольных векторов $y, y' \in \mathbb{R}^2$ из выполнения соотношения $y \succ y'$ следует соотношение $ky + c \succ ky' + c$ при всех $k > 0$, $c \in \mathbb{R}^2$.*

Всем перечисленным требованиям удовлетворяют, например, отношение $\geq$, лексико-графическое отношение и отношение $\succ$, для которого существует линейная функция ценности. Отношений, удовлетворяющих всем указанным требованиям, существует бесчисленное множество.

Если отношение строгого предпочтения $\succ$ точно известно, то в качестве «наиболее предпочтительного» решения можно выбрать любое решение из множества $\operatorname{opt}_{\succ} Y$. Однако на практике подобные ситуации встречаются редко. Как правило, об отношении $\succ$ известно только, что оно обладает некоторыми свойствами или удовлетворяет некоторым требованиям. Неполная информация об отношении строгого предпочтения порождает неполные и неточные представления о множестве $\operatorname{opt}_{\succ} Y$. В этих условиях актуален вопрос построения точной оценки сверху для неизвестного множества $\operatorname{opt}_{\succ} Y$. Решению данного вопроса и посвящен § 7.6.

Нередко на практике в результате анализа ранее предпринимавшихся действий лица, принимающего решение, или непосредственным опросом удается выявить несколько пар оценок y, y', для которых не верно ни неравенство $y \geq y'$, ни неравенство $y' \geq y$ однако лицо, принимающее решение, может с уверенностью сказать, какая из оценок данной пары предпочтительнее. Как будет показано ниже, такая информация дает возможность построить оценку сверху для искомого множества $\operatorname{opt}_{\succ} Y$ и тем самым способствует обоснованному выбору оптимального решения.

Введем еще одно требование к отношению строгого предпочтения.

Требование 5 (дополнительная информация об отношении $\succ$). *Имеются пары векторов такие $u^{(i)}$, $v^{(i)} \in \mathbb{R}^2$, $i = 1, 2, \ldots, l$, что*

$$u^{(i)} \succ v^{(i)} \text{ для всех } i \in N = \{1, 2, \ldots, l\},$$

где *$N \neq \varnothing$. Причем указанные векторы таковы, что не выполняется ни неравенство $u^{(i)} \geq v^{(i)}$, ни неравенство $v^{(i)} \geqq u^{(i)}$.*

Тот факт, что не может выполняться неравенство $u^{(i)} \geq v^{(i)}$, означает содержательность дополнительной информации. В самом деле, если неравенство $u^{(i)} \geqq v^{(i)}$ имеет место, то, согласно требованию 2, выполняется соотношение $u^{(i)} \succ v^{(i)}$, аналогичное имеющемуся в требовании 5. Второе условие $v^{(i)} \not\succ u^{(i)}$ вводится для того, чтобы не возникло противоречия с требованиями 1 и 3. Действительно, если справедливо соотношение $v^{(i)} \geqq u^{(i)} \succ v^{(i)}$, то, согласно требованиям 2 и 3, верно соотношение $v^{(i)} \succ v^{(i)}$, что несовместимо с требованием 1. Условие $N \neq \varnothing$ в последнем требовании означает, что какая-то дополнительная информация в этом требовании содержится.

Замечание. В силу требования 4, если потребуется, всегда можно считать, что в требовании 5 все векторы $v^{(i)}$ нулевые. В самом деле, если выполнено соотношение $u^{(i)} \succ v^{(i)}$, то, согласно требованию 4, также выполняется соотношение $u^{(i)} - v^{(i)} \succ 0_2$. Теперь, если левую часть полученного соотношения обозначить через $u^{(i)}$, придем к требованию 5 с нулевыми векторами $v^{(i)}$.

Далее для краткости вместо слова «требование» будем писать букву Т с соответствующим индексом.

3. Мажорантное отношение и его свойства. Введем *мажорантное отношение* $\overset{M}{\succ}$, на основании которого будет получена оценка сверху для неизвестного множества $\operatorname{opt}_{\succ} Y$. Запись $y \overset{M}{\succ} y'$ для векторов $y, y' \in \mathbb{R}^2$, по определению, означает, что выполняется неравенство $y \geq y'$ или же что при некоторых $k > 0$, $i \in N$ имеет место неравенство $y \geqq k(u^{(i)} - v^{(i)}) + y'$, где $u^{(i)}$, $v^{(i)}$ — векторы из T_5. В соответствии с этим определением выполнение соотношения $y \overset{M}{\succ} 0_2$ означает, что выполняется неравенство $y \geqslant 0_2$ или же что при некоторых $k > 0$, $i \in N$ выполняется неравенство $y \geqq k(u^{(i)} - v^{(i)})$. Геометрическая интерпретация соотношения $y \overset{M}{\succ} 0_2$ при $l = 2$ и $v^{(1)} = v^{(2)} = 0_2$ приведена на рис. 7.10. В отмеченном штриховкой тупом углу (исключая начало координат) находятся все векторы y, для которых справедливо соотношение $y \overset{M}{\succ} 0_2$. Для того чтобы получить представление о векторах y^*, для которых имеет место соотношение $y^* \overset{M}{\succ} y'$, необходимо этот угол с помощью параллельного переноса сместить так, чтобы его вершиной оказалась точка y' (рис. 7.10).

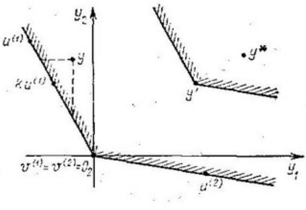

Рис. 7.10

Установим свойства мажорантного отношения.

1°. *Отношение* $\overset{M}{\succ}$ *удовлетворяет требованиям типа* T_1, T_2, T_4, T_5.

□ Предположим противное, т. е. что отношение $\overset{M}{\succ}$ не удовлетворяет T_1. Тогда для некоторого вектора y будет верно соотношение $y \overset{M}{\succ} y$. Так как неравенство $y \geq y$ невозможно, то верно неравенство $y \geqq k\ (u^{(i)}-v^{(i)})+y$ при некоторых $k>0$, $i \in N$. Отсюда получаем неравенство $v^{(i)} \geqq u^{(i)}$, которое несовместно с T_5. Следовательно, отношение $\overset{M}{\succ}$ иррефлексивно.

Выполнение аксиомы Парето для отношения $\overset{M}{\succ}$ вытекает непосредственно из определения этого отношения.

Пусть имеет место соотношение $y \overset{M}{\succ} y'$. Если, согласно этому соотношению, верно неравенство $y \geq y'$, то имеет место инвариантность относительно положительного линейного преобразования. Когда справедливо неравенство $y \geqq k\,(u^{(i)}-v^{(i)})+y'$, верно и неравенство $Ay+c \geqq Ak\ (u^{(i)}-v^{(i)})+Ay'+c$ для любых $A>0$, $c \in \mathbf{R}^2$. Следовательно, T_4 для отношения $\overset{M}{\succ}$ выполняется.

Справедливость T_5 для отношения $\overset{M}{\succ}$ вытекает из очевидных неравенств $u^{(i)} \geqq (u^{(i)}-v^{(i)})+v^{(i)}$, $i=1, 2, \ldots, l$. ■

Следующее свойство раскрывает «мажорантный характер» отношения $\overset{M}{\succ}$.

2°. *Пусть $\succ$ — произвольное отношение, удовлетворяющее T_1—T_5. Тогда для любых векторов y, $y' \in \mathbf{R}^2$ из выполнения соотношения $y \overset{M}{\succ} y'$ следует выполнение соотношения $y \succ y'$.*

□ Пусть $y \overset{M}{\succ} y'$. Выполнение этого соотношения, по определению, означает, что возможны два случая: справедливо неравенство $y \geq y'$ или неравенство $y \geqq k\,(u^{(i)}-v^{(i)})+y'$ при некоторых $k>0$, $i \in N$.

Выберем произвольное отношение $\succ$, удовлетворяющее T_1—T_5. В силу T_5 имеет место соотношение $u^{(i)} \succ v^{(i)}$, а согласно T_4 верно соотношение

$$k\,(u^{(i)}-v^{(i)})+y' \succ y'. \tag{7.9}$$

Если имеет место первый случай, т. е. когда выполняется неравенство $y \geq y'$, согласно T_2 выполнено соотношение $y \succ y'$. Если реализуется второй случай, то, учитывая соотношение (7.9), получаем соотношение $y \geqq k\ (u^{(i)}-v^{(i)})+y' \succ y'$. Если неравенство $\geqq$ выполняется как равенство, то верно соотношение $y \succ y'$. Если же неравенство $\geqq$ реализуется в виде $\geq$ то, согласно T_2 и T_3, также получаем соотношение $y \succ y'$. ■

Введем следующие множества индексов:

$$N_1=\{i \in N \mid u_1^{(i)} < v_1^{(i)},\ u_2^{(i)} > v_2^{(i)}\},$$

$$N_2=\{i \in N \mid u_1^{(i)} > v_1^{(i)},\ u_2^{(i)} < v_2^{(i)}\}.$$

На основании T_5 имеют место равенства $N_1 \cup N_2=N$, $N_1 \cap N_2=\varnothing$. При $i \in N_1$ паре векторов $u^{(i)}$, $v^{(i)}$ соответствует точка $u^{(i)}-v^{(i)}$, расположенная во II координатном углу. Если же $i \in N_2$, то точка $u^{(i)}-v^{(i)}$ находится в IV координатном углу. Возможно выполнение равенства $N_1=\varnothing$ или же равенства $N_2=\varnothing$, но одновременное выполнение этих равенств в силу того, что $N\neq\varnothing$, невозможно.

Введем положительные числа

$$K_i=\frac{u_2^{(i)}-v_2^{(i)}}{v_1^{(i)}-u_1^{(i)}}, \quad i=1, 2,\ldots, l. \tag{7.10}$$

В случае необходимости перенумеровав векторы $u^{(i)}$, $v^{(i)}$, всегда можно считать, что справедливы равенства

$$K_1=\min_{i\in N_1} K_i, \quad K_2=\max_{i\in N_2} K_i. \tag{7.11}$$

Если одно из множеств N_1 или N_2 пустое, то определено только одно число K_2 или K_1 соответственно.

Определим *сокращенное мажорантное отношение* $\overset{m}{\succ}$ следующим образом: соотношение $y\overset{m}{\succ}y'$ для векторов y, $y'\in\mathbf{R}^2$ выполняется, если справедливо неравенство $y\geqslant y'$, или при некоторых $k>0$, $i\in\{1, 2\}$ имеет место неравенство $y\geq k(u^{(i)}-v^{(i)})+y'$, где векторы $u^{(i)}$, $v^{(i)}$ из T_5 удовлетворяют равенствам (7.11). Если $N_1=\varnothing$, то в этом определении следует считать $i=2$, а при $N_2=\varnothing$ полагаем $i=1$.

Как показывает следующее свойство, отношения $\overset{M}{\succ}$ и $\overset{m}{\succ}$ совпадают друг с другом.

3°. *Соотношение $y\overset{M}{\succ}y'$ имеет место тогда и только тогда, когда верно соотношение $y\overset{m}{\succ}y'$.*

□ Достаточность следует непосредственно из определений рассматриваемых двух отношений. Необходимость. Пусть выполнено соотношение $y\overset{M}{\succ}y'$. Проверим, что в этом случае верно соотношение $y\overset{m}{\succ}y'$. На основании свойства 1° и T_4, не уменьшая общности, можно считать, что $y'=0_2$, $v^{(i)}=0_2$, $i=1, 2,\ldots, l$ (см. замечание из предыдущего пункта). В этих условиях выполнение соотношения $y\overset{M}{\succ}0_2$ означает, что верно неравенство $y\geq 0_2$ или неравенство $y\geqq ku^{(i)}$ при некоторых $k>0$, $i\in N$. Если, согласно данному соотношению, выполнено неравенство $y\geq 0_2$, то, по определению отношения $\overset{m}{\succ}$, имеем $y\overset{m}{\succ}0_2$. Пусть верно неравенство $y\geqq ku^{(i)}$ при некоторых $k>0$, $i\in N$. Для определенности предположим, что $i\in N_1$. Случай $i=1$ тривиален. Поэтому пусть $i\neq 1$. Из равенств (7.10)—(7.11) имеем неравенство $u_1^{(i)}/u_1^{(1)}\leqslant u_2^{(i)}/u_2^{(1)}$. Следовательно, найдется такое число $k'>0$, что

$$k\,\frac{u_1^{(i)}}{u_1^{(1)}}\leqslant k'\leqslant k\,\frac{u_2^{(i)}}{u_2^{(1)}}\,.$$

Из левого и правого неравенств следует соответственно неравенство $ku_1^{(i)}\geqslant k'u_1^{(1)}$ и неравенство $ku_2^{(i)}\geqslant k'u_2^{(1)}$. В свою очередь, эти неравенства влекут неравенство $ku^{(i)} \geqq k'u^{(1)}$. Окончательно получаем неравенства $y \geqq ku^{(i)} \geqq k'u^{(1)}$, означающие, что верно соотношение $y \overset{m}{\succ} 0_2$.

Случай $i \in N_2$ аналогичен. ■

Свойство 3° утверждает, что из всех l соотношений $u^{(i)} \succ v^{(i)}$, $i=1, 2, \ldots, l$, представляющих собой дополнительные сведения об отношении строгого предпочтения $\succ$, на самом деле нужно учитывать только два из этих соотношений (если же одно из множеств N_1, N_2 пусто, то одно), остальные можно из рассмотрения исключить.

4. Оценка сверху для множества оптимальных решений. Обозначим множества оптимальных оценок по отношениям $\overset{M}{\succ}$ и $\overset{m}{\succ}$ на множестве $Y \subset \mathbf{R}^2$ соответственно через $\mathrm{opt}_M Y$ и $\mathrm{opt}_m Y$. Согласно свойству 3° мажорантного отношения $\mathrm{opt}_M Y = \mathrm{opt}_m Y$. Следующее утверждение показывает, что множество оптимальных оценок по мажорантному отношению представляет собой оценку сверху для неизвестного множества $\mathrm{opt}_{\succ} Y$.

Теорема 7.10. *Для любого отношения строгого предпочтения удовлетворяющего T_1—T_5, выполняются включения*

$$\mathrm{opt}_{\succ} Y \subset \mathrm{opt}_m Y \subset P(Y).$$

□ Левое включение вытекает непосредственно из свойства 2° мажорантного отношения. Правое включение имеет место в силу свойства 1°, согласно которому мажорантное отношение удовлетворяет аксиоме Парето. ■

Как утверждает следующая теорема, оценка сверху, т. е. множество $\mathrm{opt}_m Y$, совпадает с множеством парето-оптимальных решений относительно некоторой двумерной вектор-функции на множестве Y.

Теорема 7.11. *Имеет место равенство*

$$\mathrm{opt}_m Y = P_F(Y), \tag{7.12}$$

где $P_F(Y)$ — множество парето-оптимальных решений относительно вектор-функции F вида

$$F(y) = \begin{pmatrix} \lambda_1 y_1 + (1-\lambda_1) y_2 \\ \lambda_2 y_1 + (1-\lambda_2) y_2 \end{pmatrix} \tag{7.13}$$

на множестве $Y \subset \mathbf{R}^2$. Здесь λ_1, λ_2 — числа, определяемые в случае $N_1 \neq \varnothing$, $N_2 \neq \varnothing$ равенствами

$$\lambda_i = \frac{u_2^{(1)} - v_2^{(1)}}{u_2^{(i)} - v_2^{(i)} + v_1^{(i)} - u_1^{(i)}} > 0, \quad i = 1, 2. \tag{7.14}$$

В случае выполнения равенства $N_1 = \varnothing$ число λ_2 определяется из (7.14), *а $\lambda_1 = 0$; если же справедливо равенство $N_2 = \varnothing$, то следует положить $\lambda_2 = 0$, а число λ_1 находить из равенства* (7.14).

□ Проверим, что соотношение $y \overset{m}{>} y'$ для векторов $y,\ y' \in \mathbf{R}^2$ выполняется тогда и только тогда, когда выполняется неравенство $F(y) \geq F(y')$. Предварительно отметим следующее. На основании свойства 3° мажорантного отношения и замечания п. 2, не ограничивая общности, положим $v^{(1)} = v^{(2)} = 0_2$. Далее так как неравенство $F(y) \geq F(y')$ в силу линейности вектор-функции F равносильно неравенству $F(y-y') \geqslant F(0_2)$, а соотношение $y \overset{m}{>} y'$ равносильно соотношению $y-y' \overset{m}{>} 0_2$, то можно считать, что $y' = 0_2$.

Пусть $N_1 \neq \varnothing$ и $N_2 \neq \varnothing$.

Достаточность. Проверим, что выполнение неравенства $F(y) \geq F(0_2) = 0_2$ влечет выполнение соотношения $y \overset{m}{>} 0_2$. Неравенство $F(y) \geq 0_2$ можно записать в виде

$$\begin{pmatrix} -\dfrac{u_2^{(1)}}{u_1^{(1)}}\, y_1 + y_2 \\[2ex] -\dfrac{u_2^{(2)}}{u_1^{(2)}}\, y_1 + y_2 \end{pmatrix} \geqslant \begin{pmatrix} 0 \\[2ex] 0 \end{pmatrix}. \tag{7.15}$$

Отсюда следует, что выполнение неравенства $y \leqq 0_2$ невозможно. Рассмотрим оставшиеся случаи.

а) Если $y \geq 0_2$, то, по определению отношения $\overset{m}{>}$, выполняется соотношение $y \overset{m}{>} 0_2$.

б) Пусть $y_1 < 0,\ y_2 > 0$. Из неравенства (7.15) следует неравенство $y \geqq ku^{(1)}$ при $k = y_2/u_2^{(1)} > 0$. Поэтому верно соотношение $y \overset{m}{>} 0_2$.

в) Пусть $y_1 > 0,\ y_2 < 0$. В этом случае из неравенства (7.15) вытекает неравенство $y \geqq ku^{(2)}$ при $k = y_1/u_1^{(2)} > 0$, а значит, справедливо соотношение $y \overset{m}{>} 0_2$.

Необходимость. Пусть $y \overset{m}{>} 0_2$. Здесь возможны два случая.

а) Если имеет место неравенство $y \geq 0_2$, то, очевидно, выполняется и неравенство $F(y) \geq F(0_2) = 0_2$, поскольку все числа λ_1, $(1-\lambda_1)$, λ_2, $(1-\lambda_2)$ положительны.

б) Пусть верно неравенство $y \geqq ku^{(i)}$ при некотором $k > 0$ и $i = 1$ или $i = 2$. Для определенности будем считать, что $i = 1$ (второй случай аналогичен). Из неравенства $y \geqq ku^{(1)}$ вытекают следующие покоординатные неравенства:

$$\begin{gathered} -\frac{u_2^{(1)}}{u_1^{(1)}}\, y_1 + y_2 \geqslant -\frac{u_2^{(1)}}{u_1^{(1)}}\, ku_1^{(1)} + y_2 = -ku_2^{(1)} + y_2 \geqslant 0, \\ -\frac{u_2^{(2)}}{u_1^{(2)}}\, y_1 + y_2 \geqslant -\frac{u_2^{(2)}}{u_1^{(2)}}\, ku_1^{(1)} + y_2. \end{gathered} \tag{7.16}$$

Докажем неравенство

$$-u_2^{(2)}/u_1^{(2)} < -u_2^{(1)}/u_1^{(1)}. \tag{7.17}$$

Предполагая противное, получаем, что существует такое положительное k', для которого верно неравенство $-u_1^{(2)}/u_1^{(1)} \leqslant k' \leqslant -u_2^{(2)}/u_2^{(1)}$. Отсюда следует векторное неравенство $0_2 \geqq k'u^{(1)} + u^{(2)}$, означающее справедливость соотношения $0_2 \overset{m}{\succ} u^{(2)}$. Но отношение $\overset{m}{\succ}$ удовлетворяет требованию типа T_5 (см. свойства 1° и 3° мажорантного отношения), поэтому имеет место соотношение $u^{(2)} \overset{m}{\succ} 0_2$. Таким образом, выполняется соотношение $0_2 \overset{m}{\succ} u^{(2)} \overset{m}{\succ} 0_2$, из которого, согласно свойствам 2° и 3° мажорантного отношения, следует соотношение $0_2 \succ u^{(2)} \succ 0_2$. Используя транзитивность отношения $\succ$, получаем соотношение $0_2 \succ 0_2$, противоречащее T_1.

Теперь, согласно доказанному неравенству (7.17), из второго неравенства (7.16) получаем неравенства

$$-\frac{u_2^{(2)}}{u_1^{(2)}} y_1 + y_2 > -ku_2^{(1)} + y_2 \geqslant 0,$$

где нестрогое неравенство имеет место, поскольку $y \geqq ku^{(1)}$. Отсюда и из первого неравенства (7.16) окончательно имеем

$$u_2^{(1)} y_1 - u_1^{(1)} y_2 \geqslant 0, \quad -u_2^{(2)} y_1 + u_1^{(2)} y_2 > 0.$$

Отсюда следует требуемое неравенство $F(y) \geq 0_2$.

Случаи $N_1 = \varnothing$ или $N_2 = \varnothing$ рассматриваются аналогично. ■

Используя формулу (7.14), можно проверить справедливость равенств

$$\lambda_i (u_1^{(i)} - v_1^{(i)}) + (1 - \lambda_i)(u_2^{(i)} - v_2^{(i)}) = 0, \quad i = 1, 2,$$

которые означают, что вектор $(\lambda_i, 1-\lambda_i)^{\text{т}}$ перпендикулярен вектору $u^{(i)} - v^{(i)}$, $i = 1, 2$. Значит, линии уровня первой компоненты вектор-функции F из теоремы 7.11, т. е. функции $\lambda_1 y_1 + (1-\lambda_1) y_2$, параллельны вектору $u^{(1)} - v^{(1)}$, а линии уровня, отвечающие второй компоненте вектор-функции F, параллельны вектору $u^{(2)} - v^{(2)}$. На рис. 7.11 приведен пример геометрического построения множества $\text{opt}_m Y$. Отмеченные штриховкой углы получены в результате параллельного переноса угла, образованного точками $u^{(1)}$, 0, $u^{(2)}$. На этом рисунке множество $\text{opt}_m Y$ совпадает с кривой bc, а множество $P(Y)$ составляет вся кривая $abcd$.

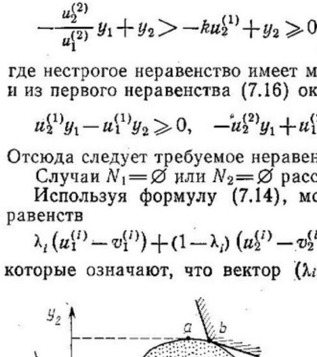

Рис. 7.11

Пример. Найдем оценку сверху для множества $\text{opt}_\succ Y$, где

$$Y = \left\{ \begin{pmatrix} 1 \\ 4{,}5 \end{pmatrix}, \begin{pmatrix} 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 3 \\ 2 \end{pmatrix}, \begin{pmatrix} 5 \\ 1{,}5 \end{pmatrix} \right\},$$

при условии, что отношение строгого предпочтения $\succ$ удовлетворяет T_1—T_4 и, кроме того, T_5 в следующей форме: $u^{(i)} \succ v^{(i)}$, $i=1, 2, 3$, причем

$$u^{(1)}=\begin{pmatrix}0\\5\end{pmatrix},\ v^{(1)}=\begin{pmatrix}2\\2\end{pmatrix},\ u^{(2)}=\begin{pmatrix}5\\0\end{pmatrix},\ v^{(2)}=\begin{pmatrix}1\\1\end{pmatrix},$$

$$u^{(3)}=\begin{pmatrix}2\\5\end{pmatrix},\ v^{(3)}=\begin{pmatrix}3\\3\end{pmatrix}.$$

Находим $N_1=\{1, 3\}$ и $N_2=\{2\}$. Далее, на основании (7.10)—(7.11) вычисляем

$$K_1=\min_{i=1,3}\frac{u_2^{(i)}-v_2^{(i)}}{v_1^{(i)}-u_1^{(i)}}=\frac{u_2^{(1)}-v_2^{(1)}}{v_1^{(1)}-u_1^{(1)}}=\frac{3}{2},$$

$$K_2=\frac{u_2^{(2)}-v_2^{(2)}}{v_1^{(2)}-u_1^{(2)}}=\frac{1}{4}.$$

Минимум при определении K_1 реализуется для $i=1$, поэтому условие $u^{(3)} \succ v^{(3)}$ может быть исключено, как не содержащее дополнительной информации по сравнению с соотношениями $u^{(i)} \succ v^{(i)}$, $i=1, 2$. Найдем оценку сверху, т. е. множество $P_F(Y)$. Согласно (7.13)—(7.14), определяем вектор-функцию

$$F(y)=\begin{pmatrix}(3/5)\,y_1+(2/5)\,y_2\\(1/5)\,y_1+(4/5)\,y_2\end{pmatrix}.$$

Теперь найдем множество парето-оптимальных решений:

$$P_F(Y)=\left\{\begin{pmatrix}1\\1{,}5\end{pmatrix},\begin{pmatrix}5\\1{,}5\end{pmatrix}\right\}.$$

Полученная оценка сверху означает, что векторы $(2,3)^{\text{т}}$ и $(3,2)^{\text{т}}$ заведомо не могут быть оптимальными, а для множества $\operatorname{opt}_{\succ} Y$ возможны следующие случаи: оно может состоять из элемента $(1,\ 1{,}5)^{\text{т}}$ или элемента $(5,\ 1{,}5)^{\text{т}}$, а также содержать оба указанных элемента. Для более определенных выводов требуется новая дополнительная информация.

§ 7.7. ПРИКЛАДНЫЕ ЗАДАЧИ

1. Анализ конечного набора проектных решений. Предположим, что имеется девять различных вариантов $x^{(1)}, x^{(2)}, \ldots, x^{(9)}$ проекта станка. Эксперты оценили их по следующим шести показателям: надежность, производительность, экономический эффект, удобство в эксплуатации, простота изготовления, внешнее оформление. Оценивание производилось по 11-балльной шкале от 0 до 10 (табл. 7.1).

Из имеющихся проектов требуется выбрать наилучший.

Заметим, что по всем шести показателям предпочтительно иметь наибольший балл. Следовательно, получаем многокритериальную задачу максимизации, в которой шесть критериев и девять возможных решений (оценок).

Выберем наилучшее решение в два этапа. Используя алгоритм теоремы 7.1, сначала найдем множество парето-оптимальных решений $\mathrm{opt}_{\geqslant} X$, где $X=\{x^{(1)}, x^{(2)}, \ldots, x^{(9)}\}$. Через $\sim$ обозначим отношение неразличимости: $a \sim b$ означает, что не верно ни $a \geq b$, ни $b \geq a$. Пусть $X_1 = X$. Сравниваем решение $x^{(1)}$ со всеми остальными решениями: $x^{(1)} \sim x^{(2)}$, $x^{(1)} \sim x^{(3)}$, $x^{(1)} \geq x^{(4)}$, $x^{(1)} \sim x^{(i)}$, $i=5, 6, \ldots, 9$. Запоминаем парето-оптимальное решение $x^{(1)}$. Удаляя из X_1 решения $x^{(1)}$ и $x^{(4)}$, приходим к множеству $X_2=\{x^{(2)}, x^{(3)}, x^{(5)}, \ldots, x^{(9)}\}$. Сравниваем теперь решение $x^{(2)}$ со всеми остальными решениями множества X_2: $x^{(3)} \geq x^{(2)}$, $x^{(2)} \sim x^{(i)}$, $i=5, 6, \ldots, 9$. Следовательно, $X_3=\{x^{(3)}, x^{(5)}, \ldots, x^{(9)}\}$. В результате сравнения решения $x^{(3)}$ с остальными решениями множества X_3 получаем $X_4=\{x^{(3)}, x^{(6)}, x^{(7)}, x^{(9)}\}$, причем решение $x^{(3)}$ запоминаем как парето-оптимальное. Аналогично находим $X_5=\{x^{(6)}, x^{(7)}, x^{(9)}\}$, $X_6=\{x^{(7)}, x^{(9)}\}$, $X_7 \neq \varnothing$, причем $x^{(7)}$ — парето-оптимальное решение, так как $x^{(7)} \geq x^{(9)}$. Таким образом, получено три парето-оптимальных решения:

Таблица 7.1

Основные показатели	x^1	x^2	x^3	x^4	x^5	x^6	x^7	x^8	x^9
Надежность	10	4	6	9	5	3	3	3	2
Производительность	2	7	8	2	1	8	5	6	5
Экономический эффект	3	3	3	3	0	0	4	3	3
Удобство в эксплуатации	2	4	5	2	5	2	2	4	1
Простота изготовления	2	3	4	1	2	4	7	4	5
Внешнее оформление	4	1	2	3	2	1	7	1	7

$$\mathrm{opt}_{\geqslant} X = \{x^{(1)}, x^{(3)}, x^{(7)}\}.$$

Остальные решения оказались заведомо непригодными.

На втором этапе из трех решений нужно отобрать наилучшее. Понятие «наилучшее решение» здесь четко не формализовано, поэтому строго обосновать окончательный выбор невозможно. Необходима дополнительная информация об отношении предпочтения. Например, если поломка станка может повлечь травму рабочего, то главным показателем следует считать надежность и принять решение $x^{(1)}$. Если же это не так, то главным является показатель производительности и наилучшим следует признать решение $x^{(3)}$, поскольку этот вариант более надежен в эксплуатации, чем $x^{(7)}$, и, кроме того, незначительно уступает ему по показателю экономического эффекта. Решение $x^{(7)}$ может оказаться наилучшим, если, например, сроки изготовления станка ограничены и определяющим фактором является простота изготовления. Насколько выбранное на втором этапе решение окажется наилучшим, в значительной степени зависит от опыта и интуиции лица, принимающего решение.

2. Лексико-графическая транспортная задача. Транспортная задача (см. § 3.4) состоит в определении такого объема перевозок продукции из одних пунктов в другие, при котором транспортные расходы минимальны. Нередко транспортная задача имеет не единственное решение. В таких случаях можно рассматривать лексико-графическую транспортную задачу. Например, если маршрут, соединяющий некоторые два пункта, перегружен транспортными средствами, то еще одним критерием, подлежащим минимизации на множестве решений транспортной задачи, может быть объем перевозок по данному маршруту.

Рассмотрим конкретную задачу. Пусть $m=n=2$.

$$f_1=2x_{11}+3x_{12}+x_{21}+2x_{22}\to\min;$$

$$f_2=x_{22}\to\min;$$

$$x_{11}+x_{12}=6,\ x_{11}+x_{21}=3,$$

$$x_{21}+x_{22}=4,\quad x_{21}+x_{22}=7,\ x_{11},\ x_{12},\ x_{21},\ x_{22}\geqslant 0.$$

Решим сначала обычную транспортную задачу с критерием f_1. Для этого отбросим, например, последнее ограничение-равенство, поскольку оно является следствием остальных ограничений, а из первого и второго равенства найдем $x_{12}-x_{21}=3$. Окончательно задача принимает следующий канонический вид:

$$f_1=2x_{11}+3x_{12}+x_{21}+2x_{22}\to\min;$$

$$x_{22}+x_{21}=4,\ x_{11}+x_{21}=3,\ x_{12}-x_{21}=3,$$

$$x_{11},\ x_{12},\ x_{21},\ x_{22}\geqslant 0.$$

Приведем соответствующую этой задаче начальную симплексную таблицу (табл. 7.2).

За один шаг можно найти оптимальное решение $x_{11}=0$, $x_{12}=6$, $x_{21}=3$, $x_{22}=1$ и оптимальное значение, равное 20. Следовательно, вторая задача, соответствующая минимизации второго критерия f на множестве решений транспортной задачи, принимает вид

$$f_2=x_{22}\to\min;$$

$$2x_{11}+3x_{12}+x_{21}+2x_{22}=20,$$

$$x_{22}+x_{21}=4,\ x_{11}+x_{21}=3,\ x_{12}-x_{21}=3,$$

$$x_{11},\ x_{12},\ x_{21},\ x_{22}\geqslant 0.$$

Таблица 7.2

		x_{11}	x_{12}	x_{21}	x_{22}
f_1	23	0	0	1	0
x_{22}	4	0	0	1	1
x_{11}	3	1	0	1*	0
x_{12}	3	0	1	−1	0

Оптимальное решение этой задачи $x_{11}=0$, $x_{12}=6$, $x_{21}=3$, $x_{22}=1$ является искомым лексико-графически-оптимальным решением.

Глава 8

ТЕОРЕТИЧЕСКИЕ ОСНОВЫ ОПТИМИЗАЦИИ В ФУНКЦИОНАЛЬНЫХ ПРОСТРАНСТВАХ

Материал этой главы вводит в круг вопросов, связанных с задачами оптимизации функционалов, и подготавливает к восприятию последующих глав. Здесь сформулированы необходимые условия оптимальности для общей задачи математического программирования, являющиеся распространением соответствующих необходимых условий, рассмотренных в гл. 2, на случай функциональных пространств. Приведены условия разрешимости задачи оптимизации, а также обсуждены численные методы решения.

Вспомогательные сведения из функционального анализа (§ 8.1) будут использованы в последующих главах, посвященных вариационному исчислению и теории оптимального управления. Более подробно ознакомиться с аппаратом функционального анализа можно, например, в [14].

§ 8.1. ВСПОМОГАТЕЛЬНЫЕ СВЕДЕНИЯ ИЗ ФУНКЦИОНАЛЬНОГО АНАЛИЗА

1. Линейное пространство. Множество L называют *линейным пространством*, если для него выполняются следующие три группы аксиом: **A**, **B** и **C**.

A. Задано правило, согласно которому каждой паре элементов x, $y \in L$ сопоставлен определенный элемент из L, называемый *суммой элементов* x и y (обозначение: $x+y$), причем:

1^0) сложение коммутативно: $x+y=y+x$;

2^0) сложение ассоциативно: $(x+y)+z=x+(y+z)$;

3^0) существует единственный элемент, который обозначают $\mathbf{0}$ и называют *нулевым элементом*, обладающим следующим свойством:

$$x+\mathbf{0}=x;$$

4^0) для каждого $x \in L$ существует единственный элемент из L, называемый *противоположным* (обозначение: $-x$), который обладает следующим свойством: $x+(-x)=\mathbf{0}$.

B. Задано правило, согласно которому каждой паре $x \in L$, $\lambda \in \mathbf{R}$ сопоставлен определенный элемент из L, называемый *произведением числа* λ *на элемент* x и обозначаемый λx, причем:

1^0) $\lambda(\mu x)=(\lambda\mu)x$;

2^0) $1 \cdot x=x$.

C. Указанные операции сложения и умножения на число связаны следующими равенствами:

1^0) $\lambda(x+y)=\lambda x+\lambda y$;

2^0) $(\lambda+\mu)x=\lambda x+\mu x$.

Заметим, что равенства в приведенных аксиомах должны иметь место для всех x, y, $z \in L$ и всех λ, $\mu \in \mathbf{R}$.

Понятие линейного пространства играет важнейшую роль в математике и используется почти во всех ее разделах. Приведем ряд примеров линейных пространств.

Пример 8.1. Множество всех векторов (направленных отрезков в пространстве), для которых сложение и умножение на число определяются правилами векторной алгебры, является линейным пространством.

Пример 8.2. Множество n-мерных векторов с введенными ранее операциями покомпонентного сложения и умножения на число (см. § 1.1) образует линейное пространство. В частности, линейным пространством является множество вещественных чисел.

Пример 8.3. Множество всех сходящихся числовых последовательностей вида $\{x_k\}$, для которых сложение и умножение на число производятся по правилам математического анализа ($\{x_k\}+\{y_k\}=\{x_k+y_k\}$, $\lambda\{x_k\}=\{\lambda x_k\}$), образует линейное пространство с нулевым элементом $\mathbf{0}=\{0, 0, \ldots\}$.

Далее будем в основном рассматривать линейные пространства, состоящие из того или иного класса функций, т. е. *функцио-*

нальные пространства. Напомним, что *суммой двух функций* f и g, имеющих одну и ту же область определения, является функция, обозначаемая $f+g$, действие которой определяется равенством $(f+g)(x)=f(x)+g(x)$, а *произведение числа* λ *на функцию* f — это функция λf вида $(\lambda f)x=\lambda f(x)$.

Пример 8.4. Множество непрерывных на отрезке $[a, b]\subset \mathrm{R}$ функций образует линейное пространство, обозначаемое $C[a, b]$. Нулевым элементом $\mathbf{0}$ в этом пространстве является функция, тождественно равная 0 на $[a, b]$.

Пример 8.5. Рассмотрим множество всех функций, имеющих на отрезке $[a, b]\subset \mathrm{R}$ непрерывные производные до порядка k включительно. Это множество также является линейным функциональным пространством (обозначение: $C_k[a, b]$). При этом обычно считают, что $C_0[a, b]=C[a, b]$.

Из приведенных примеров следует, что элементы различных линейных пространств могут сильно отличаться, что свидетельствует о высокой степени общности понятия линейного пространства.

В линейном пространстве, аналогично пространству $\mathbf{R}^n$, вводят линейно независимые и линейно зависимые системы элементов. А именно: систему элементов $\{x_1, x_2, \ldots, x_m\}$ линейного пространства L называют *линейно независимой,* если равенство

$$\lambda_1 x_1+\lambda_2 x_2+\ldots+\lambda_m x_m=\mathbf{0}$$

возможно только при $\lambda_1=\lambda_2 \ldots = \lambda_m=0$. В противном случае, т. е. если существуют числа $\lambda_1, \lambda_2, \ldots, \lambda_m$, среди которых хотя бы одно отлично от нуля и такие, что $\sum\limits_{k=1}^{m} \lambda_k x_k=\mathbf{0}$, система элементов $\{x_1, x_2, \ldots, x_m\}$ называется *линейно зависимой.*

Если в линейном пространстве L найдется линейно независимая система из n элементов, а любая система, содержащая более чем n элементов, является линейно зависимой, то L называют *конечномерным пространством размерности* n. В смысле этого определения размерность $\mathbf{R}^n$ действительно равна n.

Если для любого натурального n в линейном пространстве L можно указать линейно независимую систему из n элементов, то такое пространство называют *бесконечномерным.* Например, в пространстве $C[a, b]$ для любого n система элементов вида $\{1, t, t^2, \ldots, t^{n-1}\}$ линейно независима (поскольку равенство $\sum\limits_{k=0}^{n-1} \lambda_k t^k=0$ для всех $t\subset[a, b]$ возможно только при $\lambda_0=\lambda_1=\ldots=\lambda_{n-1}=0$), поэтому $C[a, b]$ — бесконечномерное пространство. Такими же являются и функциональные пространства из примера 8.5. В дальнейшем чаще всего будем иметь дело именно с бесконечномерными пространствами.

По аналогии с $\mathbf{R}^n$ в линейном пространстве вводят определения *отрезка,* соединяющего элементы $x\in L$ и $x'\in L$, и *выпуклого мно-*

жества. Эти определения формально не отличаются от соответствующих определений из § 1.2.

2. Линейное нормированное пространство. Линейное пространство E называют *нормированным,* если для каждого элемента $x \in E$ задано неотрицательное число, называемое *нормой элемента* x и обозначаемое $\|x\|$, причем выполняются следующие аксиомы:

1°) $\|x\| \geqslant 0$; причем равенство $\|x\| = 0$ имеет место тогда и только тогда, когда $x = \mathbf{0}$;

2°) $\|\lambda x\| = |\lambda| \cdot \|x\|$, $\lambda \in \mathbf{R}$;

3°) $\|x + y\| \leqslant \|x\| + \|y\|$ (неравенство треугольника).

Пример 8.6. Пространство $\mathbf{R}^n$, введенное в гл. 1, является нормированным: $\|x\| = \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2}$, где $x_1, x_2, \ldots, x_n$ — координаты вектора x. В частности, множество чисел $\mathbf{R}$ — пример нормированного пространства, в котором нормой числа является его абсолютная величина.

Пример 8.7. В пространстве $C[a, b]$ непрерывных на $[a, b]$ функций $x = x(t)$ норма определяется формулой

$$\|x\|_0 = \max_{t \in [a,b]} |x(t)|.$$

Можно проверить, что все аксиомы нормы в данном случае выполняются.

Пример 8.8. В линейном пространстве $C_1[a, b]$, полагая

$$\|x\|_1 = \max_{t \in [a,b]} |x(t)| + \max_{t \in [a,b]} |x'(t)|,$$

можно убедиться в том, что $\|x\|_1$ также удовлетворяет всем аксиомам нормы.

Одно и то же линейное пространство можно превратить в разные нормированные пространства, если норму вводить с помощью различных формул. Например, пространство $C_1[a, b]$ можно превратить в нормированное пространство еще одним способом, вводя норму вида $\|x\|_0$.

В нормированном пространстве *расстояние* между любыми двумя его элементами определяют по формуле

$$\rho(x, y) = \|x - y\|.$$

На основании аксиом нормы можно установить следующие свойства расстояния:

1°) $\rho(x, y) \geqslant 0$, причем равенство $\rho(x, y) = 0$ имеет место тогда и только тогда, когда $x = y$;

2°) $\rho(x, y) = \rho(y, x)$;

3°) $\rho(x, y) \leqslant \rho(x, z) + \rho(z, y)$.

По аналогии с $\mathbf{R}^n$ в нормированном пространстве E множество вида

$$U_\varepsilon(x_0) = \{x \in E \mid \|x - x_0\| < \varepsilon\}$$

называют *открытым шаром радиуса* ε *с центром в* x_0 или *ε-окре-*

стностью элемента $x_0 \in E$. Если $\varepsilon > 0$ мало, то множество $U_\varepsilon(x_0) \subset \subset C[a, b]$ образуют функции, значения которых мало отличаются от соответствующих значений функции $x_0 = x_0(t)$ (в одних и тех же точках $t \in [a, b]$).

Если элемент $x \in E$ принадлежит множеству $X \in E$ вместе с некоторой своей окрестностью, то его называют *внутренним элементом* множества X. *Открытое множество* — это множество, состоящее только из внутренних элементов. Примером открытого множества является множество $U_\varepsilon(x_0)$. Говорят, что множество $X \subset E$ ограничено, если оно целиком содержится в некоторой окрестности нулевого элемента.

Последовательность $\{x_k\}_{k=1}^{\infty}$ элементов нормированного пространства E называют *сходящейся* к элементу $x_0 \in E$, если $\lim\limits_{k\to\infty} \rho(x_k, x_0) = \lim\limits_{k\to\infty} \|x_k - x_0\| = 0$; при этом пишут $\lim\limits_{k\to\infty} x_k = x_0$.

Множество X нормированного пространства *замкнуто*, если предел каждой сходящейся последовательности элементов из X принадлежит множеству X. Если из любой последовательности $\{x_k\}$ элементов множества $X \subset E$ можно выделить сходящуюся подпоследовательность, предел которой принадлежит X, то множество X называют *компактом*. Это определение не повторяет дословно определение компакта из § 1.1. Однако если $E = \mathbf{R}^n$, то, согласно теореме Больцано — Вейерштрасса [16], оба определения эквивалентны. Как показывают примеры, в бесконечномерных пространствах не всякое замкнутое и ограниченное множество является компактом.

Говорят, что последовательность $\{x_k\} \subset E$ *сходится в себе* (является *фундаментальной*), если для любого числа $\varepsilon > 0$ найдется такой номер N_ε, что неравенство $\|x_n - x_m\| < \varepsilon$ имеет место для всех $n, m \geqslant N_\varepsilon$.

Если последовательность $\{x_k\}$ сходится, то она сходится в себе. Действительно, пусть $\lim x_k = x_0$. Тогда, по определению сходимости, для любого $\varepsilon > 0$ существует такой номер N_ε, что $\|x_n - x_0\| < \varepsilon/2$ для всех $x \geqslant N_\varepsilon$. Поэтому, согласно неравенству треугольника,

$$\|x_n - x_m\| \leqslant \|x_n - x_0\| + \|x_0 - x_m\| < \varepsilon/2 + \varepsilon/2 = \varepsilon$$

для всех $n, m \geqslant N_\varepsilon$, что и требовалось проверить.

Утверждение, обратное сформулированному выше, справедливо не всегда. Поэтому введем следующее важное определение. Нормированное пространство, в котором всякая фундаментальная последовательность сходится, называется *банаховым пространством*.

Согласно критерию Коши о равносильности сходимости в $\mathbf{R}^n$ сходимости в себе [16], *пространство $\mathbf{R}^n$ является банаховым.*

Пример 8.9. Рассмотрим нормированное пространство $C[a, b]$ функций, непрерывных на отрезке $[a, b]$ с нормой $\|x\|_0$, введенной выше, и выясним, что́ представляет собой сходимость в этом пространстве. Пусть имеется последовательность $\{x_k\}\subset C[a, b]$, сходящаяся к элементу x_0. Согласно определениям сходимости в $C[a, b]$ и нормы $\|x\|_0$, это означает, что

$$\max_{t\in[a,b]} |x_k(t) - x_0(t)| \underset{k\to 0}{\longrightarrow} 0,$$

т. е. для любого числа $\varepsilon>0$ найдется номер N_ε такой, что выполнено неравенство $\max\limits_{t\in[a,b]} |x_k(t) - x_0(t)| < \varepsilon$ при всех $n \geqslant N_\varepsilon$. Следовательно, справедливо неравенство $|x_k(t) - x_0(t)| < \varepsilon$ для всех $n\geqslant N_\varepsilon$ и всех $t\in[a, b]$, что означает равномерную сходимость функциональной последовательности $\{x_k\}$ к функции x_0 на $[a, b]$. Таким образом, *сходимость в пространстве $C[a, b]$ совпадает с равномерной сходимостью функциональных последовательностей.* В силу критерия Коши функциональная последовательность сходится равномерно тогда и только тогда, когда выполняются определенные условия, которые означают фундаментальность данной последовательности [16]. Следовательно, *пространство $C[a, b]$ является банаховым.*

С помощью аналогичных рассуждений можно убедиться, что введенное выше нормированное *пространство $C_1[a, b]$ также является банаховым.*

3. Функционалы и операторы. Числовая функция многих переменных — это отображение, заданное на множестве пространства $\mathbf{R}^n$, т. е. функция, аргументом которой является n-мерный вектор (точка из $\mathbf{R}^n$). Рассмотрение более общего случая, когда аргумент принимает значения из функционального пространства, приводит к понятию функционала.

Отображение вида $L\to\mathbf{R}$, где L — линейное пространство, называют *функционалом*. Согласно этому определению, функция многих переменных также является функционалом, так как $\mathbf{R}^n$ — линейное пространство. В дальнейшем термин функционал будем, как правило, использовать для отображений, заданных на бесконечномерных функциональных пространствах. Для отображений типа $\mathbf{R}^n\to\mathbf{R}$ сохраним термин *функция*.

Задать некоторый функционал f означает указать множество $X\subset L$ (возможен и случай $X=L$) и правило, согласно которому каждому элементу из X поставлено в соответствие определенное вещественное число. Множество X называют *областью определения* данного функционала, а множество чисел

$$f(X) = \{y\in\mathbf{R} \mid y = f(x) \text{ при некотором } x\in X\}$$

— *множеством значений* (или *образом* множества X).

Функционал f, для которого

$$f(\lambda_1 x_1 + \lambda_2 x_2) = \lambda_1 f(x_1) + \lambda_2 f(x_2) \text{ для всех } x_1, x_2\in L;\ \lambda_1, \lambda_2\in\mathbf{R}, \quad (8.1)$$

называют *линейным функционалом*. Если $L=\mathbf{R}^n$, то линейный функционал — это линейная функция вида $y=\langle c, x\rangle$, где $c\in\mathbf{R}^n$ — некоторый вектор.

Пример 8.10. Каждой непрерывной функции $x=x(t)$ сопоставим число вида

$$I(x)=\int_a^b y(t)\,x(t)\,\mathrm{d}t, \tag{8.2}$$

где $y=y(t)$ — некоторая фиксированная непрерывная функция на отрезке $[a, b]$. Тем самым на линейном пространстве $C[a, b]$ задан функционал. Согласно свойствам определенного интеграла,

$$\int_a^b y(t)\,[\lambda_1 x_1(t)+\lambda_2 x_2(t)]\,\mathrm{d}t=$$

$$=\lambda_1\int_a^b y(t)\,x_1(t)\,\mathrm{d}t+\lambda_2\int_a^b y(t)\,x_2(t)\,\mathrm{d}t,$$

т. е. рассматриваемый функционал линейный.

В гл. 1 были введены вектор-функции с множеством значений из пространства $\mathbf{R}^m$. Понятие вектор-функции более общее, чем понятие числовой функции, так как последнее является его частным случаем при $m=1$. Аналогично, обобщая понятие функционала, придем к следующему определению.

Отображение вида $L_1\to L_2$, где L_1 и L_2 — линейные пространства (возможен и случай $L_1=L_2$), называют *оператором*. Для оператора термины *область определения* и *образ* используют в том смысле, что и для функционала. Если для оператора f выполняется равенство (8.1), то его называют *линейным оператором*. В случае $L_2=\mathbf{R}^m$ оператор f является *векторным функционалом*, т. е. $f=(f_1, f_2, \ldots, f_m)$, где $f_1, f_2, \ldots, f_m$ — функционалы, заданные на пространстве L_1.

Сопоставим каждой числовой функции $x=x(t)$, заданной на отрезке $[a, b]$ и имеющей производные любого порядка, ее производную $x'(t)$. Это отображение является примером линейного оператора *(оператора дифференцирования)* типа $C_\infty[a, b]\to C_\infty[a, b]$, где $C_\infty[a, b]$ — линейное пространство функций, имеющих производные любого порядка *(пространство бесконечно дифференцируемых функций)*.

Пусть функционал f определен на некотором множестве X нормированного пространства и x_0 — элемент множества X. Говорят, что функционал f *непрерывен* в x_0, если для любого числа $\varepsilon>0$ можно указать такое число $\delta_\varepsilon>0$, что для всех элементов $x\in X\cap U_{\delta_\varepsilon}(x_0)$ выполняется неравенство $|f(x)-f(x_0)|<\varepsilon$. Для того чтобы из этого определения получить определение непрерывного в x_0 оператора $f: X\to Y$, следует неравенство $|f(x)-f(x_0)|<\varepsilon$ заменить на $\|f(x)-f(x_0)\|<\varepsilon$, где $\|\ldots\|$ — норма пространства Y.

В пространстве $\mathbf{R}^n$ линейная функция $y=\langle c, x\rangle$ непрерывна на $\mathbf{R}^n$. В бесконечномерных пространствах существуют линейные функционалы, не обладающие свойством непрерывности. Примером линейного непрерывного на $C[a, b]$ функционала служит функционал I, который определяется формулой (8.2).

Рассмотрим множество всех линейных непрерывных функционалов, заданных на некотором банаховом пространстве B. Будем считать, что на этом множестве заданы операции сложения и умножения функционала на число, определяемые теми же формулами, что и для функций. В этом случае все аксиомы линейного пространства выполняются, а значит рассматриваемое множество является линейным пространством линейных непрерывных функционалов, которое называют *сопряженным* с B и обозначают B^*. Нулевым элементом пространства B^* является линейный непрерывный функционал, сопоставляющий каждому элементу $x \in B$ число 0. Значение элемента c сопряженного пространства B^* на элементе $x \in B$ обозначают $c(x)$ или $\langle c, x \rangle_B$. Второе обозначение аналогично скалярному произведению векторов, и его можно объяснить тем, что $(\mathbf{R}^n)^*$ — это множество линейных функций вида $y = \langle c, x \rangle$, $c \in \mathbf{R}^n$.

§ 8.2. НЕОБХОДИМЫЕ УСЛОВИЯ ОПТИМАЛЬНОСТИ В ЗАДАЧАХ БЕЗ ОГРАНИЧЕНИЙ

1. Дифференцируемые функционалы. Пусть x_0 — внутренний элемент множества $X \in B$, где B — банахово пространство. Функционал f (заданный на X) называют *дифференцируемым в* x_0, если существует линейный непрерывный функционал $p \in B^*$ такой, что для всех элементов $h \in B$, удовлетворяющих условию $(x_0 + h) \in X$, имеет место представление

$$f(x_0 + h) = f(x_0) + \langle p, h \rangle_B + r(x_0, h),$$

где функционал $r(x_0, h)$ обладает свойством

$$\lim_{\|h\| \to 0} (|r(x_0, h)| / \|h\|) = 0.$$

Последнее равенство означает, что по каждому числу $\varepsilon > 0$ можно указать такое $\delta_\varepsilon > 0$, что для всех элементов h, $h \neq \mathbf{0}$, удовлетворяющих неравенству $\|h\| < \delta_\varepsilon$, выполняется неравенство $|r(x_0, h)| / \|h\| < \varepsilon$. Функционал p называют *градиентом* (или *производной*) функционала f, вычисленным в x_0. При этом используют обозначение $\nabla f(x_0)$ или $f'(x_0)$. В случае $B = \mathbf{R}^n$ данное определение дифференцируемого функционала совпадает с определением дифференцируемой функции из § 1.3.

Рассмотрим заданный на пространстве $C_1[a, b]$ функционал I вида

$$I(x) = \int_a^b F(t, x(t), x'(t))\,\mathrm{d}t, \tag{8.3}$$

где F — заданная на $[a, b] \times \mathbf{R}^2$ фиксированная, дважды непрерывно дифференцируемая функция трех переменных. Имеем

$$I(x+h)-I(x)=\int_a^b [F(t,\ x(t)+h(t),\ \ x'(t)+h'(t))-$$

$$-F(t,\ x(t),\ x'(t))]\,dt.$$

Используя для функции F формулу Тейлора (по второй x и третьей x' переменным), можно записать

$$I(x+h)-I(x)=p(h)+r(x;\ h),$$

где

$$p(h)=\int_a^b \left[\frac{\partial F(t,\ x(t),\ x'(t))}{\partial x}h(t)+\frac{\partial F(t,\ x(t),\ x'(t))}{\partial x'}h'(t)\right]dt, \quad (8.4)$$

$$r(x,\ h)=\frac{1}{2}\int_a^b\left[\frac{\partial^2 F}{\partial x^2}h^2(t)+2\frac{\partial^2 F}{\partial x\,\partial x'}h(t)\,h'(t)+\frac{\partial^2 F}{\partial x'^2}h'^2(t)\right]dt. \quad (8.5)$$

В правой части равенства (8.5) вторые частные производные вычислены в «средней точке» $(t,\ x(t)+\theta(t)\cdot h(t),\ x'(t)+\theta(t)\cdot h'(t))$, где $0<\theta(t)<1$ при $t\in[a,\ b]$.

В силу линейности интеграла функционал p линеен по h: $p(\lambda_1 h_1+\lambda_2 h_2)=\lambda_1 p(h_1)+\lambda_2 p(h_2)$. Убедимся в его непрерывности. Пусть C — число, которое больше, чем каждое из наибольших значений модулей производных $\partial F(t,\ x(t),\ x'(t))/\partial x$, $\partial F(t,\ x(t),\ x'(t))/\partial x'$ на отрезке $[a,\ b]$. Тогда по формуле (8.4), используя известные свойства интеграла, получаем неравенство

$$|p(h_1)-p(h_2)|\leqslant\int_a^b \max_{t\in[a,b]}\left(\left|\frac{\partial F}{\partial x}\right|\cdot|h_1(t)-h_2(t)|+\right.$$

$$\left.+\left|\frac{\partial F}{\partial x'}\right|\cdot|h_1'(t)-h_2'(t)|\right)dt\leqslant C\cdot|b-a|\cdot\|h_1-h_2\|_1,$$

которое влечет непрерывность функционала p.

Теперь, чтобы окончательно доказать справедливость равенства $p=\nabla I(x)$, остается убедиться, что верно соотношение $\lim\limits_{\|h\|_1\to 0}|r(x,\ h)|/\|h\|_1=0$. Для этого достаточно проверить неравенство $|r(x,\ h)|\leqslant\gamma\|h\|_1^2$, где γ — положительная константа. Так как $\|h\|_1\to 0$, можно ограничиться случаем $\|h_1\|<1$. Обозначим через A положительное число, которое больше каждого из чисел $|\sup\partial^2 F/\partial x^2|$, $|\sup\partial^2 F/\partial x\partial x'|$, $|\sup\partial^2 F/\partial x'^2|$, где супремум вычисляется на множестве $\bigcup\limits_{t\in[a,\ b]} U_1(t,\ x(t),\ x'(t))\subset\mathbf{R}^3$. Согласно определению нормы

$\|h\|_1$, для всех $t \in [a, b]$ имеют место неравенства $|h(t)| \leqslant \|h\|_1$, $|h'(t)| \leqslant \|h\|_1$. Учитывая это, из формулы (8.5) получаем

$$|r(x, h)| \leqslant \frac{1}{2} \int_a^b [A\|h\|_1^2 + 2A\|h\|_1^2 + A\|h\|_1^2]\, dt =$$

$$= 2A \cdot |b-a| \cdot \|h\|_1^2 = \gamma \|h\|_1^2,$$

где $\gamma = 2A \cdot |b-a|$. Таким образом, доказано следующее утверждение.

Теорема 8.1. *При сделанных выше предположениях функционал I, определяемый равенством (8.3), является дифференцируемым на каждой функции $x \in C_1[a, b]$ и его градиент задается формулой*

$$\langle \nabla I(x), h \rangle_{C_1[a,b]} = \int_a^b [F_x h(t) + F_{x'} h'(t)]\, dt, \tag{8.6}$$

где

$$F_x = \partial F(t, x(t), x'(t))/\partial x; \quad F_{x'} = \partial F(t, x(t), x'(t))/\partial x'.$$

2. Необходимое условие оптимальности. Здесь знакомое по гл. 2 необходимое условие оптимальности, состоящее в том, что градиент в точке экстремума равен нулю, распространяется на функционалы, заданные на множестве банахова пространства. В гл. 2 приведена постановка задачи оптимизации целевой функции на множестве евклидова пространства и введена соответствующая терминология (допустимое решение, оптимальное решение, глобальный и локальный минимумы, задачи безусловной и условной оптимизации и т. д.). Задача оптимизации функционала на множестве функционального пространства по форме не отличается от случая евклидова пространства и связанная с ней терминология совпадает с соответствующими определениями из гл. 2.

Пусть X — подмножество некоторого линейного пространства L и f — (целевой) функционал, заданный на X. Говорят, что элемент $x_0 \in X$ реализует (доставляет) минимум в задаче минимизации

$$f(x) \to \min_{x \in X},$$

если неравенство $f(x_0) \leqslant f(x)$ выполняется для всех $x \in X$. Если X совпадает со всем рассматриваемым линейным пространством, то получаем *задачу без ограничений (задачу безусловной оптимизации)*; в противном случае имеем *задачу с ограничениями (задачу условной оптимизации)*.

Приведенная ниже теорема дает необходимое условие экстремума в задаче безусловной оптимизации функционала на всем банаховом пространстве.

Теорема 8.2. *Пусть функционал f задан на банаховом пространстве B и дифференцируем в $x_0 \in B$. Для того чтобы элемент x_0 реа-*

лизовал безусловный минимум (или максимум) функционала f, необходимо, чтобы имело место равенство

$$\nabla f(x_0)=\mathbf{0}. \tag{8.7}$$

□ Пусть $\nabla f(x_0)\neq\mathbf{0}$. Тогда существует элемент $h_0\in B$, для которого $\langle\nabla f(x_0), h_0\rangle_B\neq 0$. Будем считать, что $\langle\nabla f(x_0), h_0\rangle_B<0$ (если имеет место противоположное неравенство, то оно легко приводится к виду $\langle\nabla f(x_0), -h_0\rangle_B<0$). На основании определения дифференцируемости, используя свойство линейности градиента, запишем равенство

$$f(x_0+\lambda h_0)-f(x_0)=\lambda\left[\langle\nabla f(x_0), h_0\rangle_B+\frac{r(x_0, \lambda h_0)}{\lambda}\right].$$

Знак выражения в квадратных скобках при λ, достаточно близком к нулю, определяется знаком первого слагаемого. Поэтому существует такое $\lambda_0>0$, что при $\lambda=\lambda_0$ правая, а значит, и левая части записанного равенства являются отрицательными. Это несовместимо с тем, что x_0 — минимальный элемент. ■

Следствие 8.1. *В условиях доказанной теоремы, для того чтобы элемент x_0 реализовал безусловный экстремум функционала f, необходимо, чтобы выполнялось равенство*

$$\langle\nabla f(x_0), h\rangle_B=0 \text{ для всех } h\in B. \tag{8.8}$$

Левую часть равенства (8.8) называют первым дифференциалом или первой вариацией функционала f в x_0. Следствие 8.1 утверждает, что необходимым условием экстремума является тождественное равенство нулю первой вариации функционала.

Условие оптимальности (8.8) допускает обобщение на определенный класс задач **условной** оптимизации. Рассмотрим множество H вида $H\subset B$, для которого сумма любых двух элементов из H, а также произведение любого элемента из H на произвольное вещественное число, принадлежат H. Зафиксируем элемент $x_0\in B$ и введем множество

$$B_H=\{x\in B\,|\,x=x_0+h \text{ при некотором } h\in H\}.$$

Очевидно, если $h_0\in H$, то $\lambda h_0\in H$ при любом вещественном λ, а значит, и $x_0+\lambda h_0\in B_H$. Учитывая этот факт и анализируя доказательство теоремы 8.2, убеждаемся, что оно сохраняет силу и в том случае, когда x_0 реализует условный минимум функционала f на множестве B_H.

Следствие 8.2. *Пусть функционал f определен на банаховом пространстве B и является дифференцируемым в $x_0\in B$. Для того чтобы элемент x_0 реализовал условный минимум (или максимум) функционала f на множестве B_H, необходимо, чтобы выполнялось равенство*

$$\langle\nabla f(x_0), h\rangle_B=0 \text{ для всех } h\in H.$$

По аналогии с конечномерным случаем функцию x, удовлетворяющую равенству $\nabla f(x)=\mathbf{0}$, будем называть ***стационарной.*** Таким образом, решение задачи оптимизации (если оно существует) следует искать среди стационарных функций. В п. 1 было показано, что записать градиент функционала сложнее, чем градиент функции. Кроме того, решение уравнения $\nabla f(x)=\mathbf{0}$ уже в конечномерном случае — сложная вычислительная задача, в бесконечномерном же варианте трудности резко возрастают. Следовательно, непосредственно решить функциональное уравнение $\nabla f(x)=\mathbf{0}$ практически невозможно. Поэтому обычно конкретизируют целевой функционал и выделяют классы оптимизационных задач, для которых, учитывая специфику данного класса, удается получить условия, равносильные равенству $\nabla f(x)=\mathbf{0}$ и такие, чтобы их было можно проверить. Один из таких классов задач составляет предмет изучения вариационного исчисления.

§ 8.3. ПОНЯТИЕ О ВАРИАЦИОННОМ ИСЧИСЛЕНИИ

1. Постановка простейшей задачи вариационного исчисления. В *простейшей задаче вариационного исчисления* среди всех плоских кривых, являющихся графиками функций вида $x=x(t)$ и проходящих через две заданные точки $A(t_1, x_1)$ и $B(t_2, x_2)$, $t_1<t_2$, требуется выделить такую кривую, на которой реализует свой минимум интегральный функционал

$$I(x)=\int_{t_1}^{t_2} F(t,\ x(t),\ x'(t))\,\mathrm{dt}, \tag{8.9}$$

где F — некоторая фиксированная функция трех переменных. Будем считать, что F дважды непрерывно дифференцируема на $\mathbb{R}^3$ и что таким же свойством на отрезке $[t_1, t_2]$ обладает искомая функция $x(t)$.

Если, например, в качестве F взять функцию вида $F=\sqrt{1+x'^2}$, то интеграл (8.9) определяет длину кривой и простейшая вариационная задача заключается в определении плоской линии кратчайшей длины, соединяющей точки A и B.

2. Уравнение Эйлера — Лагранжа. Для получения необходимого условия оптимальности в сформулированной выше задаче воспользуемся следствием 8.2 из § 8.2. Для этого обозначим через H множество всех дважды непрерывно дифференцируемых функций $h=h(t)$, заданных на отрезке $[t_1, t_2]$ и удовлетворяющих условию

$$h(t_1)=h(t_2)=0. \tag{8.10}$$

Очевидно, сумма любых двух функций из H, а также произведение каждой функции из H на произвольное число принадлежат введенному множеству H.

Пусть функция $x=x(t)$ является решением простейшей задачи вариационного исчисления. Рассматривая сумму $x+h$ при различных $h \in H$, получаем множество всех возможных дважды непрерывно дифференцируемых функций, графики которых проходят через точки $A(t_1, x_1)$ и $B(t_2, x_2)$. Обозначим это множество через B_H. Тот факт, что функция $x=x(t)$ является решением простейшей задачи вариационного исчисления, означает, что элемент x банахова пространства $C_1[t_1, t_2]$ реализует условный минимум функционала (8.9) на множестве B_H. Поэтому, согласно следствию 8.2, учитывая теорему 8.1, делаем вывод, что функция $x=x(t)$ удовлетворяет равенству

$$\int_{t_1}^{t_2} [F_x(t, x(t), x'(t))\, h(t) + F_{x'}(t, x(t), x'(t))\, h'(t)]\, dt = 0 \tag{8.11}$$

для всех функций $h=h(t)$, подчиненных условию (8.10).

Вычислим интеграл от $F_{x'} h'(t)$, применяя формулу итегрирования по частям * и учитывая равенства (8.10):

$$\int_{t_1}^{t_2} F_{x'} h'(t)\, dt = F_{x'} h(t) \Big|_{t_1}^{t_2} - \int_{t_1}^{t_2} \left(\frac{d}{dt} F_{x'} \right) h(t)\, dt =$$

$$= -\int_{t_1}^{t_2} \left(\frac{d}{dt} F_{x'} \right) h(t)\, dt.$$

Равенство (8.11) принимает вид

$$\int_{t_1}^{t_2} \left(F_x - \frac{d}{dt} F_{x'} \right) h(t)\, dt = 0 \tag{8.12}$$

для всех функций $h=h(t)$, удовлетворяющих равенствам (8.10).

Теперь, для того чтобы произвести дальнейшее упрощение и «избавиться» в (8.12) от интеграла, докажем вспомогательное утверждение, которое называют *первой основной леммой вариационного исчисления.*

Лемма 8.1 (лемма Лагранжа). *Пусть ω — непрерывная на отрезке $[t_1, t_2]$ числовая функция одной переменной t. Если равенство*

$$\int_{t_1}^{t_2} \omega(t)\, h(t)\, dt = 0 \tag{8.13}$$

имеет место для любой дважды непрерывно дифференцируемой функции $h=h(t)$, удовлетворяющей условию (8.10), то $\omega(t)=0$ для всех $t \in [t_1, t_2]$.

□ Пусть в некоторой точке отрезка $[t_1, t_2]$ значение функции ω отлично от нуля (для определенности будем считать его положитель-

---

* Аккуратное применение формулы интегрирования по частям требует двукратной дифференцируемости функций F и x. В гл. 9 эти требования к функции x будут ослаблены.

ным). В силу непрерывности функции ω существует отрезок $[t', t''] \subset \subset (t_1, t_2)$, на котором $\omega(t) > 0$. Введем функцию

$$h_0(t) = \begin{cases} (t-t')^3(t''-t)^3 & \text{для } t \in [t', t''], \\ 0 & \text{для остальных точек отрезка } [t_1, t_2]. \end{cases}$$

Эта функция дважды непрерывно дифференцируема и удовлетворяет граничным условиям $h_0(t_1) = h_0(t_2) = 0$. Следовательно, равенство (8.13) сохранится, если $h(t)$ заменить на $h_0(t)$. С другой стороны, используя определение h_0 и учитывая, что $\omega(t) > 0$ при $t \in [t', t'']$, получаем противоречие

$$\int_{t_1}^{t_2} \omega(t)\, h_0(t)\, \mathrm{d}t = \int_{t'}^{t''} \omega(t)\,(t-t')^3 (t''-t)^3\, \mathrm{d}t > 0. \ \blacksquare$$

Принимая во внимание равенство (8.12) и лемму 8.1, приходим к следующей теореме.

Теорема 8.3. *Если функция $x = x(t)$ является решением простейшей задачи вариационного исчисления, то она удовлетворяет дифференциальному уравнению*

$$F_x(t, x, x') - \frac{\mathrm{d}}{\mathrm{d}t} F_{x'}(t, x, x') = 0, \tag{8.14}$$

которое называют уравнением Эйлера — Лагранжа.

З а м е ч а н и е. Из приведенного доказательства следует, что уравнение Эйлера — Лагранжа представляет собой более простую запись равенства $\nabla f(x) = \mathbf{0}$, которую удалось получить, учитывая специфику простейшей вариационной задачи. Таким образом, *функция $x = x(t)$ является стационарной тогда и только тогда, когда она удовлетворяет уравнению Эйлера — Лагранжа.*

Воспользуемся необходимым условием (8.14) в задаче определения кратчайшей плоской линии, соединяющей точки A и B. Имеем $F = \sqrt{1+x'^2}$, и так как $F_x = 0$, $F_{x'} = x'/\sqrt{1+x'^2}$, то получаем уравнение Эйлера — Лагранжа вида

$$\frac{\mathrm{d}}{\mathrm{d}t} \frac{x'}{\sqrt{1+x'^2}} = 0,$$

откуда последовательно находим: $x'/\sqrt{1+x'^2} = C$, $x' = C_1$, $x = C_1 t + C_2$, где постоянные C_1, C_2 определяют, используя граничные условия прохождения найденной линии через точки A и B. Как и следовало ожидать, кратчайшей линией, соединяющей две точки, является прямая.

§ 8.4. НЕОБХОДИМЫЕ УСЛОВИЯ ОПТИМАЛЬНОСТИ В ЗАДАЧАХ С ОГРАНИЧЕНИЯМИ

1. Постановка общей задачи математического программирования. Пусть X — линейное пространство, на котором заданы функционалы f, g_1, g_2, ..., g_k и оператор G, принимающий значения в линейном пространстве Y. *Общая задача математического программирования* имеет следующий вид:

$$f(x) \to \min \tag{8.15}$$

при ограничениях

$$g_j(x) \leqslant 0, \quad j=1, 2, \dots, k, \tag{8.16}$$

$$G(x) = 0. \tag{8.17}$$

Если $Y = \mathbf{R}^{m-k}$, $m > k$, то оператор G является векторным функционалом $G = (g_{k+1}, \dots, g_m)$ и ограничение (8.17) в «покоординатной форме» записывается в виде равенств $g_j(x) = 0$, $j = k+1, \dots, m$. В этом случае общая задача математического программирования отличается от задачи математического программирования (2.31) с ограничениями в форме неравенств и равенств только тем, что в ней вместо функций рассматриваются функционалы. Если пространство Y бесконечномерно, то ограничение (8.17) можно трактовать как бесконечную систему равенств.

В класс общих задач математического программирования входят задачи, являющиеся предметом изучения вариационного исчисления и теории оптимального управления. Более подробно эти два направления теории оптимизации рассмотрены в последующих главах. Здесь ограничимся формулировкой соответствующих задач. При этом фигурирующие в них линейные пространства конкретно указывать не будем.

В вариационном исчислении исследуются задачи с целевыми функционалами следующих трех типов.

1) *Интегральный функционал* задается формулой

$$I(x) = \int_a^b F(t, x(t), x'(t))\,dt,$$

где $x = x(t)$ — n-мерная вектор-функция вида $x(t) = (x_1(t), x_2(t), \dots, x_n(t))$, $x'(t) = (x'_1(t), x'_2(t), \dots, x'_n(t))$, F — числовая функция $2n+1$ переменной. Функционал такого вида при $n=1$ участвует в формулировке простейшей задачи вариационного исчисления.

2) *Терминальный функционал* определяется равенством

$$\Psi(x) = \psi[x(t)|_{t=a}, x(t)|_{t=b}],$$

где ψ — числовая функция двух переменных.

3) *Смешанный функционал* представляет собой сумму интегрального и терминального функционалов:

$$\int_a^b F(t, x(t), x'(t))\,dt + \psi[x(t)|_{t=a}, x(t)|_{t=b}].$$

Если в вариационной задаче ограничения имеют вид системы функциональных равенств $g_1(x) - l_1 = 0$, $g_2(x) - l_2 = 0$, ..., $g_k(x) -$

$-l_k=0$, где $g_1, g_2, \ldots, g_k$ — функционалы и $l_1, l_2, \ldots, l_k \in \mathrm{R}$, то ее называют *изопериметрической задачей.*

В более сложных вариационных задачах встречаются *дифференциальные ограничения (дифференциальные связи)*, имеющие вид

$$D_i(t, x(t), x'(t))=0, \quad i=1, 2, \ldots, m; \quad t\in[a, b],$$

где $D_1, D_2, \ldots, D_m$ — числовые функции $2n+1$ переменной. В частности, при $m=n$ дифференциальные ограничения могут иметь следующий вид:

$$x_i'=\varphi_i(t, x), \quad i=1, 2, \ldots, n, \tag{8.18}$$

где $\varphi_1, \varphi_2, \ldots, \varphi_n$ — числовые функции $n+1$ переменной. Ограничения (8.18) представляют частный случай общего ограничения (8.17). Для того чтобы убедиться в этом, обозначим через G_i функционал, который определен на рассматриваемом линейном пространстве векторных функций $x=x(t)$ формулой

$$G_i(x)=x_i'-\varphi_i(t, x), \quad i=1, 2, \ldots, n.$$

Равенства (8.18) равносильны равенствам $G_i(x)=\mathbf{0}$, $i=1, 2, \ldots, n$, которые, вводя векторный функционал $G=(G_1, G_2, \ldots, G_n)$, можно записать в виде (8.17).

Во всех указанных выше вариационных задачах фигурируют только первые производные искомой функции. Задачи, в целевой функционал (а также ограничения, если они имеются) которых входят производные более высокого порядка, называют *задачами высшего порядка*. Эти задачи наиболее трудны для исследования.

В задаче оптимального управления имеется два типа переменных: *переменные состояния* $x=x(t)$ (n-мерная векторфункция) и *переменные управления* $u=u(t)$ (m-мерная векторфункция). Целевой функционал в общем случае может быть суммой интегрального и терминального членов

$$\int_a^b F(t, x(t), u(t))\,\mathrm{d}t+\psi[x(t)|_{t=a}, x(t)|_{t=b}],$$

а дифференциальные ограничения, которые обязательно присутствуют в задаче, обычно имеют вид

$$x_i'=\varphi_i(t, x, u), \quad i=1, 2, \ldots, n.$$

Здесь F, φ_i, $i=1, 2, \ldots, n$ — функции $n+m+1$ переменной. Кроме того, в задаче оптимального управления имеется ограничение на возможность выбора переменных управления: $u(t)\in U\subset\mathbf{R}^m$ при $t\in$ $\in[a, b]$, где U обычно замкнутое ограниченное множество евклидова пространства. Подобных ограничений в классических вариационных задачах нет.

2. Необходимые условия оптимальности. Задачи условной оптимизации в бесконечномерных пространствах более разнообразны и сложны, чем конечномерные задачи. Однако ограничения в этих задачах можно также учитывать с помощью функции (функционала) Лагранжа, являющегося аналогом функции Лагранжа из гл. 2. При этом необходимое условие оптимальности в дифференцируемом случае заключается в равенстве нулю градиента функции Лагранжа. Для того чтобы сформулировать необходимые условия оптимальности для задачи (8.15)—(8.17), введем понятия дифференцируемого и сопряженного операторов.

Пусть задан оператор $G: X \to Y$, где X, Y — некоторые банаховы пространства. Дифференцируемый оператор определяется аналогично дифференцируемому функционалу. А именно: оператор G *дифференцируем в* x_0 если существует линейный непрерывный оператор $P: X \to Y$ такой, что для всех элементов $h \in X$, удовлетворяющих включение $(x_0+h) \in X$, имеет место равенство

$$G(x_0+h) = G(x_0) + P(h) + R(x_0, h),$$

причем оператор $R(x_0, h)$ со значениями в пространстве Y обладает свойством

$$\lim_{\|h\|_X \to 0} (\|R(x_0, h)\|_Y / \|h\|_X) = 0,$$

где $\|\ldots\|_X$ и $\|\ldots\|_Y$ — нормы в соответствующих пространствах. Оператор P называют *градиентом (производной)* оператора G и обозначают $\nabla G(x_0)$. Для обозначения значения градиента в h будем использовать запись $\nabla G(x_0)(h)$.

Пусть, например, $A: X \to Y$ — линейный непрерывный оператор и $x_0 \in X$ — произвольный элемент. Тогда $\nabla A(x_0) = A$, т. е. градиент такого оператора совпадает с самим оператором. Это следует из равенства $A(x_0+h) = A(x_0) + A(h)$, выполняющегося в силу линейности оператора A.

Рассмотрим линейный непрерывный оператор Φ, действующий из одного банахова пространства в другое: $\Phi: X \to Y$. Линейный непрерывный оператор $\Phi^*: Y^* \to X^*$, где X^* и Y^* — сопряженные пространства, называют *сопряженным* по отношению к Φ, если равенство $\langle c, \Phi(x) \rangle_Y = \langle \Phi^*(c), x \rangle_X$ выполняется для всех $x \in X$, $c \in Y^*$.

Для общей задачи математического программирования *функция (функционал) Лагранжа* определяется формулой

$$L(x, \lambda_*, \lambda_0, \lambda_1, \ldots, \lambda_k) = \lambda_0 f(x) + \sum_{j=1}^{k} \lambda_j g_j(x) + \langle \lambda_*, G(x) \rangle_Y,$$

где $(\lambda_0, \lambda_1, \ldots, \lambda_k) \in \mathbf{R}^{k+1}$, $\lambda_* \in Y^*$. Числа $\lambda_0, \lambda_1, \ldots, \lambda_k$ вместе с линейным непрерывным функционалом λ_* называют *множителями Лагранжа.* В частном случае, когда G — векторный функционал $G = (g_{k+1}, \ldots, g_m)$, т. е. когда $Y = \mathbf{R}^{m-k}$, $m > k$, последнее слагаемое в выражении для функционала Лагранжа упрощается:

$$\langle \lambda_*, G(x) \rangle_Y = \sum_{j=k+1}^{m} \lambda_j g_j(x), \tag{8.19}$$

где $(\lambda_{k+1}, ..., \lambda_m) \in \mathbb{R}^{m-k}$. Таким образом, чтобы «обслужить» конечную систему функциональных ограничений, требуется конечное число числовых множителей Лагранжа. Если же пространство Y бесконечномерно, то в функционале Лагранжа присутствует «функциональный множитель» λ_*.

Следующая теорема, доказательство которой можно найти в [2], представляет собой распространение теоремы 2.6 на случай функциональных пространств.

Теорема 8.4. *Пусть X, Y — банаховы пространства. Предположим, что оператор $G: X \to Y$ и функционалы f, g_1, g_2, ..., g_k непрерывны и дифференцируемы в некоторой окрестности элемента $x_0 \in X$, удовлетворяющего ограничениям (8.16)—(8.17), причем градиент оператора $\nabla G(x_0)$ является непрерывным в x_0. Кроме того, будем считать, что образ пространства X при отображении $\nabla G(x_0)$, т. е. множество $\nabla G(x_0)(X) \subset Y$, замкнут. Для того чтобы элемент x_0 был решением задачи* (8.15)—(8.17), *необходимо, чтобы существовали одновременно не равные нулю множители Лагранжа* $\lambda_j \geqslant 0$, $j=1, 2, ..., k$; $\lambda_* \in Y^*$, такие, что*

$$\nabla_x L(x_0, \lambda_*, \lambda_1, \ldots, \lambda_k) = \lambda_0 \nabla f(x_0) + \sum_{j=1}^{k} \lambda_j \nabla g_j(x_0) + (\nabla G(x_0))^* (\lambda_*) = 0, \quad \lambda_j g_j(x_0) = 0, \; j = 1, 2, \ldots, k. \tag{8.20}$$

Если, кроме того, выполняется условие регулярности: $\nabla G(x_0)(X) = Y$ и существует такой элемент $\tilde{x} \in X$, что $\nabla G(x_0)(\tilde{x}) = 0$, $\langle \nabla g_j(x_0), \tilde{x} \rangle_X < 0$ для всех тех индексов j, $1 \leqslant j \leqslant k$, для которых $g_j(x_0) = 0$, то в равенстве (8.20) *можно положить $\lambda_0 = 1$.*

В частном случае векторного функционала $G = (g_{k+1}, ..., g_m)$, учитывая равенство (8.19), необходимое условие оптимальности (8.20) при выполнении условия регулярности можно записать в следующем виде:

$$\nabla f(x_0) + \sum_{j=1}^{m} \lambda_j \nabla g_j(x_0) = 0,$$

что совпадает с соответствующим необходимым условием теоремы 2.6.

§ 8.5. РАЗРЕШИМОСТЬ ЗАДАЧИ ОПТИМИЗАЦИИ

1. Пример задачи оптимизации, в которой оптимального решения не существует. Рассмотрим задачу минимизации интегрального функционала, определяемого формулой

$$I(x) = \int_0^{3\pi/2} (x'^2(t) - x^2(t))\, dt$$

* Здесь имеется в виду, что либо среди чисел λ_0, λ_1, ..., λ_k имеется отличное от нуля, либо все эти числа могут быть равны нулю, но $\lambda_* \neq 0$.

на множестве функций из пространства $C_1[0, 3\pi/2]$, удовлетворяющих граничным условиям $x(0)=x(3\pi/2)=0$. Это простейшая задача вариационного исчисления, для которой уравнение Эйлера — Лагранжа имеет вид $x''+x=0$. Общее решение такого дифференциального уравнения определяется формулой $x=C_1 \sin t+C_2 \cos t$. Единственная функция этого семейства, удовлетворяющая заданным граничным условиям, есть $x_0=x_0(t)\equiv 0$, $t\in[0, 3\pi/2]$ Таким образом, в данной задаче имеется только одна стационарная функция. Казалось бы, что именно она и должна быть оптимальной. Однако это не так. В данной задаче оптимальное решение вообще отсутствует, так как целевой функционал не ограничен снизу. Например, для последовательности допустимых функций вида $x_n(t)= =n \sin 2t/3$, $n=1, 2, \ldots$, имеем

$$I(x_n)=\int\limits_0^{3\pi/2}\left(\frac{4n^2}{9}\cos^2\frac{2t}{3}-n^2\sin^2\frac{2t}{3}\right)\mathrm{d}t=-\frac{5\pi n^2}{12}\xrightarrow[n\to\infty]{}-\infty.$$

Сама стационарная функция x_0 не реализует даже локального минимума функционала. Действительно, рассмотрим последовательность допустимых функций $\bar{x}_n(t)=(1/n)\sin 2t/3$, $n=1, 2, \ldots$. Для этой последовательности имеем

$$\|\bar{x}_n-x_0\|_1=\max_{t\in[0,3\pi/2]}\left|\frac{1}{n}\sin 2t/3\right|+\max_{t\in[0,3\pi/2]}\left|\frac{2}{3n}\cos\frac{2t}{3}\right|=\frac{5}{3n}\xrightarrow[n\to\infty]{}0,$$

а значит, $\bar{x}_n\to x_0$, т. е. в любой окрестности элемента x_0 найдутся элементы последовательности $\{\bar{x}_n\}$. Вычисляя значение функционала на членах этой последовательности

$$I(\bar{x}_n)=\int\limits_0^{3\pi/2}\left(\frac{4}{9n^2}\cos^2\frac{2t}{3}-\frac{1}{n^2}\sin^2\frac{2t}{3}\right)\mathrm{d}t=-\frac{5\pi}{12n^2}<0=I(x_0),$$

убеждаемся в том, что элемент x_0 не может реализовать локальный минимум.

Рассмотренный пример показывает, насколько важно быть уверенным в существовании оптимального решения. Если оптимальное решение существует, то оно находится среди стационарных элементов; только в этом случае при условии единственности стационарного элемента можно справедливо заключить, что он же является и оптимальным.

2. Обобщенная теорема Вейерштрасса. В функциональных пространствах справедлива теорема, аналогичная теореме Вейерштрасса: непрерывный функционал достигает своего минимального (максимального) значения на компактном множестве. Ниже сформулирована и доказана более общая теорема, в которой требование непрерывности функционала заменено более слабым требованием полунепрерывности, что расширяет область ее применения.

Напомним определение непрерывного функционала. Функционал f, заданный на множестве X нормированного пространства, непрерывен в $x_0\in X$, если для любого числа $\varepsilon>0$ можно указать такое число $\delta_\varepsilon > 0$, что для всех $x \in X \cap U_{\delta_\varepsilon}(x_0)$ выполняется неравенство

$$|f(x) - f(x_0)| < \varepsilon. \tag{8.21}$$

Неравенство (8.21) эквивалентно двум следующим неравенствам:

$$f(x) > f(x_0) - \varepsilon, \quad f(x) < f(x_0) + \varepsilon. \tag{8.22}$$

Определение полунепрерывного функционала можно получить из определения непрерывного функционала, заменяя неравенство (8.21) одним из неравенств (8.22). А именно: функционал f называют *полунепрерывным снизу* (*сверху*) в $x_0 \in X$, если для любого $\varepsilon > 0$ существует $\delta_\varepsilon > 0$ такое, что для всех элементов $x \in X \cap U_{\delta_\varepsilon}(x_0)$ выполняется левой (правое) неравенство (8.22).

Из определений следует, что непрерывный функционал полунепрерывен (как снизу, так и сверху), но не наоборот. При $X \subset \mathbf{R}^n$ получаем определение полунепрерывной функции. Простым примером полунепрерывной сверху функции служит функция $y = [x]$ («целая часть числа x»), график которой изображен на рис. 8.1. Из неравенств (8.22) следует, что функционал $-f$ полунепрерывен снизу, если функционал f полунепрерывен сверху.

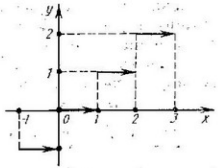

Рис. 8.1

Если функционал f полунепрерывен снизу (сверху) на каждом элементе множества X, то его называют *полунепрерывным снизу (сверху) на множестве* X.

Теорема 8.5. *Если множество X нормированного пространства компактно, а функционал f полунепрерывен снизу (сверху) на X, то задача оптимизации*

$$f(x) \to \min_{x \in X} \quad (f(x) \to \max_{x \in X})$$

имеет решение.

□ Ограничимся рассмотрением задачи минимизации и обозначим $A = \inf\limits_{x \in X} f(x)$. По определению точной нижней грани, существует минимизирующая последовательность*, т. е. последовательность элементов $\{x_n\} \subset X$ такая, что $\lim\limits_{n \to \infty} f(x_n) = A$. В силу компактности множества X найдется такая последовательность номеров $\{n_k\} \subset \{n\}$, что предел последовательности $\{x_{n_k}\}$ принадлежит X, т. е. $\lim x_{n_k} = x_0 \in X$. Используя тот факт, что функционал f полунепрерывен снизу в x_0, получаем, что для каждого числа $\varepsilon_m = 1/m$,

---

* Определение минимизирующей последовательности для задачи минимизации функции дано в § 2.2. Для задачи минимизации функционала это определение аналогично.

$(m=1, 2, \ldots)$ найдется число $\delta_m>0$ такое, что для всех $x\in X\bigcap U_{\delta_m}(x_0)$ верно неравенство $f(x)>f(x_0)-1/m$. Следовательно, так как $x_{n_k}\to x_0$, можно указать последовательность номеров $\{n_m\}\subset\{n_k\}$, обладающую следующим свойством: $f(x_{n_m})>f(x_0)-1/m$, $m=1, 2, \ldots$ Отсюда, переходя к пределу при $m\to\infty$, получаем $A=\lim f(x_{n_m})\geqslant\geqslant f(x_0)$, что вместе с неравенством $A\leqslant f(x_0)$, вытекающим из определения A, приводит к равенству $f(x_0)=A$. Таким образом, x_0 — решение задачи минимизации. ■

Пусть минимизирующая последовательность $\{x_n\}$ в задаче минимизации функционала f на множестве X (не обязательно компактном) сходится к элементу $x_0\in X$. Если функционал полунепрерывен снизу в x_0, то, повторяя рассуждения, использованные при доказательстве теоремы 8.5, придем к заключению, что элемент x_0 является решением данной задачи минимизации. Сформулируем полученный результат в виде следствия.

Следствие 8.3. *Пусть функционал f определен на множестве X нормированного пространства и последовательность элементов $\{x_n\}\subset X$ такова, что*

$$\lim_{n\to\infty} f(x_n)=\inf_X f(x),\quad \lim_{n\to\infty} x_n=x_0\in X.$$

Тогда если функционал f полунепрерывен снизу в x_0, то элемент x_0 реализует его наименьшее возможное значение на множестве X.

Это утверждение используют при обосновании численных методов решения задач оптимизации в функциональных пространствах.

§ 8.6. О ЧИСЛЕННЫХ МЕТОДАХ ОПТИМИЗАЦИИ В ФУНКЦИОНАЛЬНЫХ ПРОСТРАНСТВАХ

1. Возможности обобщения численных методов нелинейного программирования. В гл. 5 были рассмотрены численные методы приближенного решения задачи минимизации числовой функции на некотором допустимом множестве евклидова пространства $\mathbf{R}^n$ (а также на всем пространстве $\mathbf{R}^n$). Это метод покоординатного спуска, градиентные методы, метод Ньютона, метод условного градиента, метод сопряженных направлений и т. д. Среди них только в методе покоординатного спуска используется «конечномерность» пространства $\mathbf{R}^n$, и поэтому его нельзя распространить на задачи минимизации функционалов. Остальные же методы (возможно, с некоторыми непринципиальными поправками) можно обобщить на задачи в функциональных пространствах *. При этом соответствующие утверждения о характере сходимости этих методов будут аналогичны конечномерному случаю. Более подробно ознакомиться с обобщениями численных методов нелинейного программирования на задачи в функциональных пространствах можно в [9].

2. Метод Ритца. Рассмотрим кратко численный метод, который используют для оптимизации функционалов, не имеющий аналога

* Для обобщения методов второго порядка требуется понятие двукратной дифференцируемости функционала, которое здесь не определялось.

в нелинейном программировании. Его суть заключается в сведении решения бесконечномерной задачи оптимизации к решению последовательности конечномерных задач.

Предварительно введем некоторые определения. Пусть $\varphi_1, \varphi_2, \ldots, \varphi_n$ — элементы нормированного пространства X. Множество $K[\varphi_1, \ldots, \varphi_n]$, состоящее из всех возможных линейных комбинаций элементов $\varphi_1, \varphi_2, \ldots, \varphi_n$, называют *линейной оболочкой* этих элементов, т. е.

$$K[\varphi_1, \ldots, \varphi_n] = \{x \in X | x = \sum_{i=1}^{n} \lambda_i \varphi_i \text{ при некоторых } \lambda_i \in \mathbf{R},$$

$$i = 1, 2, \ldots, n\}$$

Говорят, что последовательность $\{\varphi_k\}_{k=1}^{\infty}$ элементов нормированного пространства X *полна* в X, если для каждого элемента $x_0 \in X$ и произвольного числа $\varepsilon > 0$ можно указать такое натуральное число n и элемент $x \in K[\varphi_1, \ldots, \varphi_n]$, что $\|x - x_0\| < \varepsilon$. Например, в пространстве непрерывных функций полной является последовательность многочленов $1, t, t^2, \ldots, t^n, \ldots$ [14].

Теорема 8.6. *Пусть f — непрерывный функционал на нормированном пространстве X, который достигает глобального минимума в $x_0 \in X$. Предположим, что имеется такая последовательность элементов $\{\varphi_k\}_{k=1}^{\infty} \subset X$,* что:

1) *последовательность $\{\varphi_k\}$ полна в X;*

2) *для каждого натурального n существует элемент $x_n \in K[\varphi_1, \ldots, \varphi_n]$, удовлетворяющий равенству*

$$f(x_n) = \min_{x \in K[\varphi_1, \ldots, \varphi_n]} f(x). \tag{8.23}$$

Тогда последовательность $\{x_n\}$ сходится к минимуму по функционалу, т. е. $\lim_{n \to \infty} f(x_n) = f(x_0)$.

□ Зафиксируем произвольное $\varepsilon > 0$. В силу непрерывности функционала f существует $\delta > 0$ такое, что для всех $x \in U_\delta(x_0)$ верно неравенство

$$f(x) - f(x_0) < \varepsilon. \tag{8.24}$$

Последовательность $\{\varphi_k\}$ полная, поэтому существуют натуральное n и элемент $x \in K[\varphi_1, \ldots, \varphi_n]$, для которых $\|x - x_0\| < \delta$, т. е. выполняется включение $x \in U_\delta(x_0)$. Значит, для этого элемента x верно неравенство (8.24), т. е.

$$f(x) - f(x_0) < \varepsilon, \quad x \in K[\varphi_1, \varphi_2, \ldots, \varphi_n].$$

Согласно второму условию теоремы, найдется элемент $x_n \in K[\varphi_1, \varphi_2, \ldots, \varphi_n]$, для которого выполняется равенство (8.23). Следовательно, для указанного x_n справедливо и неравенство

$$f(x_n) - f(x_0) < \varepsilon. \tag{8.25}$$

Очевидно, что с увеличением номера n множество $K[\varphi_1, ..., \varphi_n]$ становится разве что шире и поэтому числовая последовательность $\{f(x_n)\}$ не возрастает. На основании этого можно сделать вывод, что неравенство (8.25) имеет место не только для указанного номера n, но и для всех бо́льших номеров. Отсюда, учитывая, что ε произвольно, получаем требуемое. ■

Из доказанной теоремы вытекает следующий способ приближенного решения задачи минимизации непрерывного функционала. Выбирают последовательность функций $\{\varphi_k\}$, обладающую свойствами 1) и 2) теоремы 8.6, и далее последовательно решают задачи вида (8.23) при $n=1$, $n=2$ и т. д. Получающаяся в результате последовательность функций x_1^*, x_2^*, ... сходится к минимуму по функционалу. Заканчивая процесс вычислений на некотором k-м шаге, получают значение $f(x_k^*)$, приближенно равное глобальному минимуму (при этом сама функция x_k^* может «сильно отличаться» от оптимальной).

На практике последовательность $\{\varphi_k\}$ обычно строят с помощью системы многочленов $1, t, ..., t^n, ...$ или же системы тригонометрических функций $\sin t, \sin 2t, ..., \sin nt, ...$.

Так, если решается простейшая вариационная задача минимизации интегрального функционала I вида (8.3) на множестве непрерывно дифференцируемых функций, подчиненных граничным условиям $x(t_1)=x_1$, $x(t_2)=x_2$, то в качестве последовательности $\varphi_0, \varphi_1, ..., \varphi_k, ...$, как правило, выбирают одну из последовательностей

$$\varphi_0(t)=x_1+\frac{x_2-x_1}{t_2-t_1}(t-t_1),$$

$$\varphi_k(t)=(t-t_1)^k(t-t_2), \quad k=1, 2,...$$

$$\varphi_k(t)=\sin\frac{\pi k(t-t_1)}{t_2-t_1}, \quad k=1, 2,... .$$

При этом задача (8.23) состоит в безусловной минимизации функции

$$\xi(\lambda_1, \lambda_2, ..., \lambda_n)=I\left[\varphi_0(t)+\sum_{k=1}^{n}\lambda_k\varphi_k(t)\right]$$

на пространстве $\mathbf{R}^n$. Эта задача является конечномерной, и для ее решения можно использовать методы конечномерной оптимизации, изложенные в предыдущих главах.

§ 8.7. ПРИКЛАДНЫЕ ЗАДАЧИ

1. Задача определения траектории распространения света в среде с переменной плотностью. В плоской прозрачной неоднородной изотропной среде имеется точечный источник света. Свяжем с этой средой прямоугольную декартову систему координат xOy. Обозначим через $v=v(x, y)$ скорость света в точке (x, y). Выберем пару точек $A(x_1, y_1)$ и $B(x_2, y_2)$ $(x_1 < x_2$ и $y_1 > 0$; $y_2 > 0)$ и расположим источник света в точке A. Определим форму кривой, вдоль которой свет распространяется от точки A до B.

Согласно принципу Ферма, свет распространяется по такой траектории, для перемещения вдоль которой требуется наименьшее возможное время. Пусть $y = y(x)$ — функция, график которой представляет искомую кривую, проходящую через точки A и B. Обозначив через s длину пройденного светом пути, через t — время, получим для скорости света выражение

$$v = \frac{\mathrm{d}s}{\mathrm{d}t} = \frac{\sqrt{1+y'^2}\,\mathrm{d}x}{\mathrm{d}t}.$$

Отсюда, интегрируя, находим время T распространения света от точки A до B:

$$T = \int_0^T \mathrm{d}t = \int_{x_1}^{x_2} \frac{\sqrt{1+y'^2}}{v}\,\mathrm{d}x.$$

Таким образом, приходим к простейшей вариационной задаче минимизации функционала

$$T(y) = \int_{x_1}^{x_2} \frac{\sqrt{1+y'^2(x)}}{v(x, y(x))}\,\mathrm{d}x.$$

на множестве гладких функций, удовлетворяющих граничным условиям $y(x_1) = y_1$, $y(x_2) = y_2$.

Рассмотрим частный случай, когда скорость распространения света пропорциональна ординате: $v = ky$, где $k > 0$. Целевой функционал принимает вид

$$T(y) = \frac{1}{k}\int_{x_1}^{x_2} \frac{\sqrt{1+y'^2(x)}}{y(x)}\,\mathrm{d}x.$$

Воспользуемся необходимым условием оптимальности теоремы 8.3. В данном случае

$$F = \frac{\sqrt{1+y'^2}}{ky}, \quad F_y = -\frac{\sqrt{1+y'^2}}{ky^2}, \quad F_{y'} = \frac{y'}{ky\sqrt{1+y'^2}}.$$

Следовательно, уравнение Эйлера — Лагранжа, которому должна удовлетворять искомая функция, имеет вид

$$-\frac{\sqrt{1+y'^2}}{ky^2} - \frac{\mathrm{d}}{\mathrm{d}x}\,\frac{y'}{ky\sqrt{1+y'^2}} = 0. \qquad (8.26)$$

Умножим обе части равенства на y' и запишем его так:

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\sqrt{1+y'^2}}{y} - y'\,\frac{y'}{y\sqrt{1+y'^2}}\right) = 0. \qquad (8.27)$$

(Выполняя в последнем уравнении дифференцирование по x, можно непосредственно получить уравнение (8.26).) В результате интегрирования уравнения (8.27) находим

$$\frac{\sqrt{1+y'^2}}{y}-\frac{y'^2}{y\sqrt{1+y'^2}}=C_1,$$

где $C_1=\mathrm{const}$, $C_1\neq 0$, или

$$y'^2=\frac{1-C_1^2y^2}{C_1^2y^2}.$$

Общее решение этого дифференциального уравнения, т. е

$$y^2+\left(x+\frac{C_2}{C_1}\right)^2=\frac{1}{C_1^2},$$

где $C_2=\mathrm{const}$, определяет семейство окружностей с центром на оси абсцисс. Значения постоянных C_1 и C_2 находят из условия выполнения граничных условий: $y(x_1)=y_1$, $y(x_2)=y_2$. Таким образом, заданная пара точек A и B, через которые должна проходить искомая траектория, выделяет единственную окружность из найденного семейства (ее центром является точка пересечения оси абсцисс и перпендикуляра, проведенного через середину отрезка, соединяющего эти точки). Из физического смысла задачи следует, что оптимальное решение должно существовать. Таким образом, найденное стационарное решение в силу единственности является оптимальным.

2. Задача определения формы подвешенной нити. В двух точках вертикальной плоскости закреплены концы однородной абсолютно гибкой нити длины l. Требуется определить форму, которую нить примет под действием силы тяжести.

Расположим в данной плоскости прямоугольную декартову систему координат xOy так, чтобы действие силы тяжести было противоположно направлению оси ординат. Для простоты ограничимся случаем, когда точки A и B, в которых закреплена нить, расположены на одной высоте: $A(-x_0, h)$, $B(x_0, h)$. Пусть $y=y(x)$ — уравнение искомой кривой. Тогда выражение для потенциальной энергии нити имеет вид

$$W_п(y)=\int_{-x_0}^{x_0}\rho g y(x)\sqrt{1+y'^2(x)}\,\mathrm{d}x, \qquad (8.28)$$

где ρ — линейная плотность, g — ускорение свободного падения. Длина нити постоянна и равна l. Запишем это условие в виде ограничения-равенства

$$\int_{-x_0}^{x_0}\sqrt{1+y'^2(x)}\,\mathrm{d}x-l=0. \qquad (8.29)$$

Нить займет положение, в котором ее потенциальная энергия примет наименьшее возможное значение, и поэтому задача сводится к отысканию гладкой функции y, реализующей минимум функционала (8.28) при ограничении (8.29) и удовлетворяющей граничным условиям

$$y(-x_0)=y(x_0)=h. \tag{8.30}$$

Составляем функцию Лагранжа:

$$L=\int_{-x_0}^{x_0}\left[\lambda_0\rho g y\sqrt{1+y'^2}+\lambda_1\left(\sqrt{1+y'^2}-\frac{l}{2x_0}\right)\right]\mathrm{d}x. \tag{8.31}$$

На основании теорем 8.3 и 8.4 делаем вывод, что оптимальная функция должна удовлетворять уравнению Эйлера — Лагранжа, в котором

$$F=\lambda_0\rho g y\sqrt{1+y'^2}+\lambda_1\left(\sqrt{1+y'^2}-\frac{l}{2x_0}\right),$$

$$F_y=\lambda_0\rho g\sqrt{1+y'^2},$$

$$F_{y'}=(\lambda_0\rho g y+\lambda_1)y'/\sqrt{1+y'^2},$$

т. е. уравнению

$$\lambda_0\rho g\sqrt{1+y'^2}-\frac{\mathrm{d}}{\mathrm{d}x}\left[(\lambda_0\rho g y+\lambda_1)y'/\sqrt{1+y'^2})\right]=0.$$

Непосредственной проверкой (дифференцированием по x) можно убедиться в том, что полученное уравнение равносильно уравнению

$$\frac{\mathrm{d}}{\mathrm{d}x}\left[\lambda_0\rho g y\sqrt{1+y'^2}+\lambda_1\sqrt{1+y'^2}-\frac{\lambda_0\rho g y y'^2-\lambda_1 y'^2}{\sqrt{1+y'^2}}\right]=0,$$

которое в результате интегрирования и несложных преобразований принимает следующий вид:

$$\lambda_0\rho g y+\lambda_1=C_1\sqrt{1+y'^2}, \tag{8.32}$$

где $C_1=\mathrm{const}$. В случае $C_1=0$ из (8.32) получаем решение $y\equiv\equiv\mathrm{const}$, что возможно только в случае $l=2x_0$. Поэтому далее можно считать, что $C_1\neq0$ и $l>2x_0$. Если в уравнении (8.32) выполняется условие $\lambda_0=0$, то приходим к уже рассмотренному случаю $y\equiv\mathrm{const}$. Значит, можно считать, что $\lambda_0=1/(\rho g)\neq0$ и $\lambda_1=\lambda$. Тогда из (8.32) получаем уравнение

$$y+\lambda=C_1\sqrt{1+y'^2}.$$

Общее решение этого дифференциального уравнения имеет вид

$$y=C_1\operatorname{ch}\left(\frac{x}{C_1}+C_2\right)-\lambda. \tag{8.33}$$

Используя симметрию граничных условий, приходим к равенству $C_2=0$. Длина кривой, определяемой уравнением (8.33), равна

$$\int_{-x_0}^{x_0} \sqrt{1+\operatorname{sh}^2 \frac{x}{C_1}}\, dx = 2C_1 \operatorname{sh} \frac{x_0}{C_1} = l.$$

Из правой части равенства можно однозначно определить значение $C_1=C_1{}^*$, при котором оно выполняется. Граничное условие позволяет также однозначно найти значение множителя Лагранжа $\lambda^* = h - C_1{}^* \operatorname{ch} x_0/C_1{}^*$. В результате получаем единственную стационарную функцию $y=C_1^* \operatorname{ch} \frac{x}{C_1^*} - \lambda^*$, являющуюся решением исходной задачи. Найденная функция определяет кривую, называемую *цепной линией*.

Глава 9
ВАРИАЦИОННОЕ ИСЧИСЛЕНИЕ

В настоящее время в теории вариационного исчисления разработан мощный и универсальный математический аппарат, позволяющий решать экстремальные задачи в функциональных пространствах. Этот аппарат находит широкое применение при решении различного рода технических задач.

Первые четыре параграфа посвящены рассмотрению простейшей задачи вариационного исчисления, которая уже сформулирована в гл. 8. Здесь изложение автономно и основывается на понятии первой и второй вариаций целевого функционала (исторически первыми были введены именно эти понятия, а затем появились понятия дифференцируемого функционала и его градиента). Сформулированы как необходимые, так и достаточные условия оптимальности. Рассмотрены более общие вариационные задачи и соответствующие им необходимые условия оптимальности. Приведена вариационная задача Лагранжа в понтрягинской форме, формулировка которой близка к формулировке задачи оптимального управления.

Для изучения материала данной главы необходимо предварительно ознакомиться с § 8.1.

§ 9.1. ПОСТАНОВКА ПРОСТЕЙШЕЙ ЗАДАЧИ ВАРИАЦИОННОГО ИСЧИСЛЕНИЯ

1. Задача о брахистохроне*. Эта задача, о которой упоминает еще Галилей, после исследований И. Бернулли 1696 г. послужила толчком к появлению методов решения широкого класса подобных задач, составивших впоследствии основу вариационного исчисления.

* Слово «брахистохрона» образовано из двух слов, которые в переводе на русский язык означают «кратчайший» и «время».

Пусть в вертикальной плоскости имеются две точки A и B (первая выше второй). Данные точки могут быть соединены различными плоскими кривыми (в частности, и прямой линией). Предположим, что в A помещена материальная точка массы m, которая под действием силы тяжести может «скатываться» из точки A в B по различным кривым, соединяющим A и B. Геометрически задача о брахистохроне заключается в отыскании такой кривой (если она существует), по которой материальная точка достигнет B за кратчайшее время. Эту кривую называют *брахистохроной.*

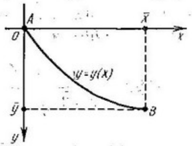

Рис. 9.1

Для того чтобы дать математическую формулировку задачи, введем на рассматриваемой плоскости прямоугольную систему координат. Поместим ее начало в точку A и направим ось ординат вертикально вниз (рис. 9.1). Обозначим координаты точки B через $(\bar{x}, \bar{y})$. По условию, материальная точка начинает двигаться из A $(0, 0)$ без начальной скорости, поэтому согласно закону сохранения энергии можно записать

$$mv^2/2 = mgy,$$

где v — скорость, y — ордината материальной точки (g — ускорение свободного падения). Отсюда находим $v=\sqrt{2gy}$. Будем считать, что $y=y(x)$ — уравнение кривой, по которой «скатывается» материальная точка. Если обозначить через s длину пройденного точкой пути, а через t — время, то можно записать

$$v=\frac{\mathrm{d}s}{\mathrm{d}t}=\frac{\sqrt{1+y'^2}\,\mathrm{d}x}{\mathrm{d}t}.$$

Следовательно,

$$\mathrm{d}t=\frac{\sqrt{1+y'^2}\,\mathrm{d}x}{v}=\frac{\sqrt{1+y'^2}\,\mathrm{d}x}{\sqrt{2gy}}.$$

Интегрируя, получаем следующее выражение:

$$T=\int_0^T \mathrm{d}t=\frac{1}{\sqrt{2g}}\int_0^{\bar{x}}\frac{\sqrt{1+y'^2}\,\mathrm{d}x}{\sqrt{y}}, \tag{9.1}$$

где T — время, в течение которого материальная точка движется вдоль кривой $y=y(x)$ $(0\leqslant x\leqslant\bar{x})$ из A в B.

Равенство (9.1) задает функционал T, определенный на множестве кривых (точнее функций) вида $y=y(x)$, подчиненных граничным условиям

$$y(0)=0,\ y(\bar{x})=\bar{y}. \tag{9.2}$$

Из физических соображений ясно, что рассматриваемые кривые не должны иметь «изломов» («углов») и поэтому функции $y=y(x)$ можно считать гладкими (непрерывно дифференцируемыми): $y\in C_1[0, \bar{x}]$.

Математическая постановка задачи о брахистохроне такова: среди всех функций вида $y=y(x)$ пространства $C_1[0, \bar{x}]$, которые удовлетворяют условиям (9.2), требуется найти функцию (если она существует), реализующую минимум функционала T, определяемого формулой,

$$T(y)=\frac{1}{\sqrt{2g}}\int\limits_0^{\bar{x}}\frac{\sqrt{1+y'^2(x)\,\mathrm{d}x}}{\sqrt{y(x)}}\,. \tag{9.3}$$

2. Постановка простейшей вариационной задачи. Опишем класс экстремальных задач, в который входит задача о брахистрохроне.

В простейшей вариационной задаче *целевой функционал*, который будем обозначать через I, имеет вид

$$I(y)=\int\limits_{x_1}^{x_2} f(x,\ y(x),\ y'(x))\,\mathrm{d}x, \tag{9.4}$$

где подынтегральную функцию f трех переменных x, y и y' считают заданной на некоторой области* $\hat{D}\subset\mathbf{R}^3$. Рассмотрим также область (проекцию множества $\hat{D}$ на плоскость xOy) вида

$D=\{(x, y)\in\mathbf{R}^2 \mid$ при некотором $y'\in\mathbf{R}$ верно $(x, y, y')\in\hat{D}\}$ и обозначим через $A(x_1, y_1)$, $B(x_2, y_2)$ фиксированные точки этого множества, предполагая, что они не лежат на одной вертикальной прямой, т. е. $x_1<x_2$ (рис. 9.2).

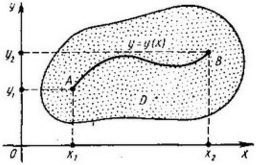

Рис. 9.2

Опишем *множество Y допустимых функций*, на котором следует минимизировать целевой функционал,

$$Y=\{y\in C_1[x_1,\ x_2] \mid y(x_1)=y_1,\ y(x_2)=y_2;$$
$$(x,\ y(x),\ y'(x))\in\hat{D} \text{ при всех } x\in[x_1,\ x_2]\}. \tag{9.5}$$

Второе требование в (9.5) геометрически означает, что кривые, определяемые допустимыми функциями, не должны выходить за пределы области $\hat{D}$. Будем предполагать, что множество Y содержит по крайней мере один элемент; в противном случае постановка задачи оптимизации теряет смысл.

Простейшая задача вариационного исчисления состоит в минимизации функционала (9.4) на множестве Y вида (9.5).

---

* *Областью* называется открытое множество, любые две точки которого можно соединить непрерывной кривой [16].

Выше не был конкретизирован класс функций, которому принадлежит функция f из (9.4). Требования, которым должна удовлетворять функция f, по крайней мере должны быть такими, чтобы существовал интеграл (9.4). Это выполняется, например, если функция f непрерывна на $\hat{D}$. На самом деле для того, чтобы можно было формулировать те или иные условия оптимальности, следует накладывать более жесткие требования непрерывной дифференцируемости или даже двукратной непрерывной дифференцируемости. Такого рода предположения в практических задачах, как правило, выполняются.

Задача о брахистохроне — это простейшая задача вариационного исчисления, в которой

$$f(x, y, y') = \sqrt{1+y'^2}/\sqrt{2gy}, \ \hat{D} = \{(x, y, y') \in \mathbf{R}^3 \mid y > -\varepsilon\},$$

$$\varepsilon > 0, \ x_1 = y_1 = 0, \ x_2 = \bar{x}, \ y_2 = \bar{y}.$$

3. Типы экстремумов в вариационных задачах. Простейшая вариационная задача, сформулированная выше, заключается в отыскании глобального минимума целевого функционала. В данной задаче можно рассматривать также и локальные минимумы.

В пространстве $C_1[x_1, x_2]$ определена норма

$$\| y \|_1 = \max_{x \in [x_1, x_2]} | y(x) | + \max_{x \in [x_1, x_2]} | y'(x) | ,$$

а значит, имеет смысл говорить об ε-окрестности элемента y_0

$$U_\varepsilon^1(y_0) = \{y \in C_1[x_1, x_2] \mid \| y - y_0 \|_1 < \varepsilon\}.$$

Такую окрестность называют *ε-окрестностью первого порядка*. При малом $\varepsilon > 0$ она состоит из функций, значения $y(x)$ которых (а также значения $y'(x)$ их первых производных) мало отличаются от значений $y_0(x)$ (соответственно $y_0'(x)$) функции y_0 в одних и тех же точках $x \in [x_1, x_2]$. Иными словами, ε-окрестность функции y_0 образуют функции, близкие к y_0 как по значениям, так и по величине производной.

В соответствии с этим элемент $y_0 \in Y$ реализует *слабый локальный минимум* функционала I на множестве $Y \subset C_1[x_1, x_2]$, если можно указать такое число $\varepsilon > 0$, что неравенство $I(y_0) \leqslant I(y)$ имеет место для всех $y \in Y \cap U_\varepsilon^1(y_0)$.

Рассмотрим теперь норму пространства непрерывных функций $C_0[x_1, x_2]$:

$$\| y \|_0 = \max_{x \in [x_1, x_2]} | y(x) | .$$

Так как $C_1[x_1, x_2] \subset C_0[x_1, x_2]$, то $\|\ldots\|_0$ можно рассматривать как еще одну норму в пространстве $C_1[x_1, x_2]$. В соответствии с этим введем понятие *ε-окрестности нулевого порядка*

$$U_\varepsilon^0(y_0) = \{y \in C_1[x_1, x_2] \mid \| y - y_0 \|_0 < \varepsilon\},$$

которую при малом $\varepsilon > 0$ составляют функции y, значения которых мало отличаются от y_0; при этом производные могут отличаться значительно.

Говорят, что элемент $y_0 \in Y$ реализует *сильный локальный минимум* функционала I на множестве $Y \subset C_1[x_1, x_2]$, если существует $\varepsilon > 0$ такое, что неравенство $I(y_0) \leqslant I(y)$ выполнено для всех $y \in Y \cap U_\varepsilon^0(y_0)$.

Приведем наглядный пример, иллюстрирующий различие слабого и сильного локальных минимумов. Предположим, что парусная лодка движется по течению широкой реки от точки A к точке B при сильном встречном ветре (рис. 9.3). Нужно как можно скорее попасть из A в B. Движение без паруса по прямой, соединяющей точки A и B, доставляет локальный минимум, так как при движении по близким гладким кривым парусом воспользоваться нельзя, а пройденный лодкой путь будет длиннее чем по прямой. Однако этот локальный минимум будет лишь слабым локальными минимумом, поскольку, двигаясь с парусом галсами (зигзагами), можно эффективно использовать силу ветра и в результате выиграть во времени.

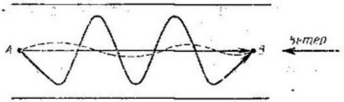

Рис. 9.3

Пример. Рассмотрим простейшую вариационную задачу, в которой

$$f(x, y, y') = y^2(1 - y'^2), \ x_1 = y_1 = y_2 = 0, \ x_2 = 1, \ \hat{D} = \mathbf{R}^3.$$

В этой задаче элемент $y_0 = y_0(x) \equiv 0$ $(x \in [0, 1])$ реализует слабый локальный минимум. Покажем это. Рассмотрим $\varepsilon \in (0, 1)$ и произвольную допустимую функцию $y \in U_\varepsilon^1(y_0)$. Поскольку $y'(x) \leqslant \max\limits_{x \in [0,1]} |y'(x)| \leqslant \|y\|_1 < \varepsilon < 1$ для всех $x \in [0, 1]$, учитывая, что $I(y_0) = 0$, получаем неравенство

$$I(y) - I(y_0) = \int\limits_0^1 y^2(x)(1 - y'^2(x))\,dx \geqslant 0.$$

В силу произвольности выбора функции y неравенство означает, что элемент y_0 реализует слабый локальный минимум. ■

Элемент y_0 не реализует сильный локальный минимум. В самом деле, рассмотрим последовательность допустимых функций $y_n = (1/\sqrt{n}) \sin 2\pi nx$, $n = 1, 2, \ldots$. Поскольку

$$\|y_n - y_0\|_0 = \max_{x \in [0,1]} |(1/\sqrt{n}) \sin 2\pi nx| = 1/\sqrt{n} \underset{n\to\infty}{\longrightarrow} 0,$$

в любой сколь угодно малой окрестности нулевого порядка элемента y_0 содержатся члены последовательности y_n, $n = 1, 2, \ldots$. Вычисляя значение функционала на элементах этой последовательности и сравнивая его с $I(y_0)$, получаем требуемое:

$$I(y_n)=\int_0^1 \frac{\sin^2 2\pi nx}{n}(1-4\pi^2 n\cos^2 2\pi nx)\,dx=\int_0^1\left[\frac{1-\cos 4\pi nx}{2n}-\right.$$

$$\left.-\frac{\pi^2}{2}(1-\cos 8\pi nx)\right]dx=\frac{1}{2n}-\frac{\pi^2}{2}<0=I(y_0),\ n=1,\ 2,\ldots$$

4. Расширение множества допустимых функций. В простейшей вариационной задаче оптимальные решения ищут среди непрерывно дифференцируемых (гладких) функций. Во многих реальных задачах предположение о гладкости решения оправдано, однако имеются и такие, в которых условие гладкости является «слишком жестким» и по смыслу задачи оптимальная функция может иметь разрывную производную.

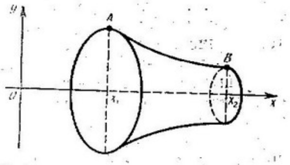

Рис. 9.4

Пусть в плоскости xOy заданы две точки $A(x_1,\ y_1)$, $B(x_2,\ y_2)$, $x_1<x_2$. Эти точки можно соединить различными кривыми, уравнения которых имеют вид $y=y(x)$. Вращая каждую такую кривую вокруг оси Ox, будем получать соответствующие поверхности вращения (рис. 9.4).

Задача о минимальной поверхности вращения заключается в отыскании такой кривой, которая при вращении образует поверхность наименьшей возможной площади. Здесь минимизируемый функционал определяется равенством

$$I(y)=2\pi\int_{x_1}^{x_2} y(x)\sqrt{1+y'^2(x)}\,dx,$$

а множество допустимых функций имеет вид, аналогичный (9.5). Однако в данной задаче допущение $y\in C_1[x_1,\ x_2]$ не представляется естественным; в число допустимых кривых можно включить и такие, которые имеют «изломы» («углы»). В результате приходим к задаче в которой допустимыми являются функции из класса кусочно-гладких.

Напомним, что числовая функция $y=y(x)$ называется *кусочно-гладкой (кусочно-непрерывно дифференцируемой)* на отрезке $[x_1,\ x_2]$, если она непрерывна на $[x_1,\ x_2]$, а ее производная непрерывна всюду на $[x_1,\ x_2]$, за исключением, быть может, некоторого конечного числа внутренних точек, где она терпит разрывы первого рода [16]. Множество кусочно-гладких функций с введенными стандартными операциями сложения и умножения на число образует линейное пространство, которое далее будем обозначать через $KC_1[x_1,\ x_2]$.

Введем еще одно определение. Пусть $y \in KC_1[x_1, x_2]$. Говорят, что точка $(x_0, y(x_0)) \in \mathbf{R}^2$, где $x_1 < x_0 < x_2$, кривой $y = y(x)$ является *угловой точкой (точкой излома)*, если односторонние производные функции $y = y(x)$ слева и справа в точке x_0 существуют, но не равны друг другу: $y'(x_0 - 0) \neq y'(x_0 + 0)$.

Простейшая расширенная вариационная задача заключается в минимизации функционала (9.4) на *расширенном допустимом множестве*

$$\hat{Y} = \{y \in KC_1[x_1, x_2] \mid y(x_1) = y_1, \ y(x_2) = y_2;$$

$$(x, y(x), y'(x)) \in \hat{D} \text{ при каждом } x \in [x_1, x_2]\}.$$

Заметим, что в предположении $y \in KC_1[x_1, x_2]$ при условии непрерывности функции f на множестве $\hat{D}$ интеграл в правой части (9.4) существует, поскольку в этом случае функция $f(x, y(x), y'(x))$ кусочно-непрерывна* на отрезке $[x_1, x_2]$ [16].

Решением расширенной вариационной задачи является функция, реализующая глобальный минимум. Для расширенной задачи вводят также понятие локального минимума. При этом используют окрестность нулевого порядка, поэтому локальный минимум в расширенной задаче — это сильный локальный минимум.

Оказывается, что при переходе от простейшей вариационной задачи к расширенной задаче оптимальное значение целевого функционала существенно не изменяется: точные нижние (точные верхние) грани обеих задач совпадают. Это следует из леммы о скруглении углов.

Лемма 9.1 (о скруглении углов). *Пусть функция f из (9.4) непрерывна на $\hat{D} = \mathbf{R}^3$. Для произвольной функции $\hat{y} \in \hat{Y}$ и любого $\varepsilon > 0$ можно указать такую функцию $y \in Y$, что верно неравенство*

$$|I(y) - I(\hat{y})| < \varepsilon \tag{9.6}$$

и, кроме того, $\|y - \hat{y}\|_0 < \varepsilon$.

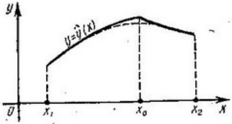

Рис. 9.5

□ Случай, когда функция $\hat{y}$ не имеет угловых точек, тривиален. Рассмотрим случай, когда существует только одна угловая точка $(x_0, y(x_0))$. Если их больше, то рассуждение проводят аналогично для каждой из них.

Доказательство заключается в «скруглении» функции $\hat{y} \in Y$ в окрестности точки x_0, т. е. в подборе такой функции $y \in Y$, которая отличается от $\hat{y}$ только в некоторой окрестности точки x_0 (рис. 9.5), причем выполняется неравенство (9.6).

---

* Числовую функцию $y = y(x)$ называют *кусочно-непрерывной* на отрезке $[x_1, x_2]$, если она непрерывна на этом отрезке, за исключением, быть может, некоторого конечного числа внутренних точек, где она терпит разрывы первого рода.

Требуемыми свойствами обладает, например, функция

$$y_\delta(x)=\hat{y}(x)+\Delta\varphi_\delta(x),\ x\in[x_1,\ x_2],$$

где $\Delta=\hat{y}'(x_0+0)-\hat{y}'(x_0-0)$ и

$$\varphi_\delta(x)=\begin{cases}\dfrac{(\delta-|x-x_0|)^2}{4\delta}, & \text{если } |x-x_0|\leqslant\delta,\\ 0, & \text{если } |x-x_0|>\delta,\end{cases}$$

причем число $\delta>0$ выбрано достаточно малым. В самом деле, функция φ_δ непрерывна на отрезке $[x_1,\ x_2]$, $\varphi_\delta(x_0\pm\delta)=0$, а производная φ_δ' только в точке x_0 имеет разрыв первого рода, причем $\varphi_\delta'(x_0\pm\delta)=\mp 1/2$. В этом легко убедиться, воспользовавшись равенством $\lim\limits_{x\to\pm 0}\mathrm{d}|x|/\mathrm{d}x=\pm 1$. Поэтому функция y_δ непрерывна и имеет непрерывную производную, так как единственный угол скруглен: $y_\delta'(x_0\pm 0)=(\hat{y}'(x_0+0)+\hat{y}'(x_0-0))/2=y_\delta'(x_0)$. Таким образом, $y_\delta\in Y$.

Определим, насколько малым должно быть число δ. Для всех $x\in[x_1,\ x_2]$ справедливы неравенства

$$|y_\delta(x)-\hat{y}(x)|\leqslant|\Delta|\delta/4,\ |y_\delta'(x)-\hat{y}'(x)|\leqslant|\Delta|/2.$$

Здесь (а также ниже в формуле для K) вместо производной $\hat{y}'(x_0)$, которая не существует, взято число $(\hat{y}'(x_0+0)+\hat{y}'(x_0-0))/2$. Рассмотрим компактное множество

$$K=\{(x,\ y,\ y')\in\mathbb{R}^3\ |\ x_1\leqslant x\leqslant x_2;\ \ |y-\hat{y}(x)|\leqslant|\Delta|/4;$$

$$|y'-\hat{y}'(x)|\leqslant|\Delta|/2 \text{ при } x\neq x_0;\ \ |y'-\hat{y}'(x_0)|\leqslant|\Delta|\}$$

и обозначим $M=\max\limits_K|f(x,\ y,\ y')|$. По теореме Вейерштрасса, этот максимум достигается. Для точки $x,\ y_\delta(x),\ y_\delta'(x))\in K$ можно записать

$$|I(y_\delta)-I(\hat{y})|\leqslant\int\limits_{x_0-\delta}^{x_0+\delta}|f(x,\ y_\delta(x),\ y_\delta'(x))-$$

$$-f(x,\ \hat{y}(x),\ \hat{y}'(x))|\,\mathrm{d}x\leqslant 4M\delta.$$

Таким образом, число δ нужно взять таким, чтобы выполнялись неравенства $\delta<1$, $4M\delta<\varepsilon$, а также включение $(x_0-\delta,\ x_0+\delta)\subset\subset[x_1,\ x_2]$.

Поскольку

$$\|y_\delta-\hat{y}\|_0=|\Delta|\max_{x\in[x_1,x_2]}|\varphi_\delta(x)|=|\Delta|\delta/4,$$

для выполнения неравенства $\|y_\delta-\hat{y}\|_0<\varepsilon$ следует взять такое число δ, для которого справедливо неравенство $|\Delta|\delta/4<\varepsilon$. ■

Следствие 9.1. *Пусть функция f непрерывна на множестве $\hat{D}=\mathbf{R}^3$. Тогда имеет место равенство*

$$\inf_{y\in\hat{Y}} I(y)=\inf_{y\in Y} I(y). \tag{9.7}$$

□ Согласно включению $KC_1[x_1, x_2]\supset C_1[x_1, x_2]$ имеет место неравенство $\inf_{\hat{Y}} I(y)\leqslant\inf_{Y} I(y)$. Для проверки справедливости равенства (9.7) предположим противное, т. е. что $\inf_{\hat{Y}} I(y)<\inf_{Y} I(y)$. По определению точной нижней грани, последнее неравенство влечет существование такой функции $\hat{y}\in\hat{Y}$ и числа $\varepsilon>0$, что $I(\hat{y})<\inf_{Y} I(y)-\varepsilon$. Согласно лемме 9.1, найдется функция $y\in Y$, для которой $I(y)<\inf_{Y} I(y)$, что невозможно. ■

Таким образом, если простейшая задача вариационного исчисления имеет решение $y_0\in C_1[x_1, x_2]$, оно же является решением соответствующей расширенной задачи (однако если расширенная задача имеет решение $y_0\in KC_1[x_1, x_2]$, то не обязательно $y_0\in C_1[x_1, x_2]$).

§ 9.2. ПЕРВАЯ ВАРИАЦИЯ. УРАВНЕНИЕ ЭЙЛЕРА — ЛАГРАНЖА

1. Определение первой вариации. Рассмотрим простейшую вариационную задачу, сформулированную в предыдущем параграфе, и сформулируем необходимое условие слабого локального минимума.

Пусть функция $y\in Y$, где множество Y имеет вид (9.5), реализует слабый локальный минимум функционала I (9.4). Это означает, что при некотором $\varepsilon>0$ неравенство $I(y)\leqslant I(\bar{y})$ выполняется для всех $\bar{y}\in Y\cap U_\varepsilon^1(y)$.

Пусть h — произвольная функция, удовлетворяющая условиям

$$h\in C_1[x_1, x_2],\ h(x_1)=h(x_2)=0. \tag{9.8}$$

Рассмотрим семейство функций вида $y+\lambda h$, зависящее от параметра $\lambda\in\mathbf{R}$. По условию, множество $\hat{D}$ открытое, поэтому при всех λ, достаточно близких к нулю, выполняется включение $(y+\lambda h)\in Y\cap U_\varepsilon^1(y)$. Значит, для всех указанных λ верно неравенство $\varphi(0)\leqslant\varphi(\lambda)$, где $\varphi(\lambda)=I(y+\lambda h)$ — функция одной переменной λ (при фиксированных y и h). Таким образом, функция φ достигает локального минимума при $\lambda=0$. Предполагая функцию f в формуле (9.4) непрерывной вместе с частными производными f_y и $f_{y'}$ на $\hat{D}$, согласно правилу Лейбница [16], получаем, что функция φ дифференцируема. Следовательно, по теореме Ферма, ее производная при $\lambda=0$ равна нулю: $\varphi'(0)=0$. Учитывая, что

$$\frac{d}{d\lambda}f(x,\ y+\lambda h,\ y'+\lambda h')=f_y h+f_{y'}h',$$

равенство $\varphi'(0)=0$ запишем в виде

$$\varphi'(0)=\left\{\frac{d}{d\lambda}\int_{x_1}^{x_2} f(x,\ y(x)+\lambda h(x),\ y'(x)+\lambda h'(x))\,dx\right\}\Bigg|_{\lambda=0}=$$

$$=\int_{x_1}^{x_2}[f_y(x,\ y(x),\ y'(x))\,h(x)+f_{y'}(x,\ y(x),\ y'(x))\,h'(x)]\,d\,x=0, \tag{9.9}$$

где дифференцирование по λ выполнено под знаком интеграла. Равенство (9.9) справедливо для любой функции h, удовлетворяющей соотношениям (9.8).

Введем следующее важное определение. Предположим, что функция f, а также ее частные производные f_y и $f_{y'}$ непрерывны на множестве D. Пусть $y\in Y$ — произвольная фиксированная функция. Функционал δI, определяемый формулой

$$\delta I(h)=\int_{x_1}^{x_2}[f_y(x,\ y(x),\ y'(x))\,h(x)+f_{y'}(x,\ y(x),\ y'(x))\,h'(x)]\,dx, \tag{9.10}$$

называют *первой вариацией функционала* I (относительно элемента y). Этот функционал задан на множестве функций, удовлетворяющих условиям (9.8).

Полученный результат в терминах первой вариации можно сформулировать следующим образом.

Лемма 9.2. *Пусть функция $y\in Y$ реализует слабый локальный минимум в простейшей вариационной задаче. При сделанных выше предположениях относительно функций f, f_y, $f_{y'}$, первая вариация целевого функционала (относительно элемента y) тождественно равна нулю:*

$$\int_{x_1}^{x_2}[f_y(x,\ y(x),\ y'(x))\,h(x)+f_{y'}(x,\ y(x),\ y'(x))\,h'(x)]\,dx=0 \tag{9.11}$$

для всех функций h, удовлетворяющих условиям (9.8).

2. Уравнение Эйлера — Лагранжа. Преобразуем необходимое условие (9.11) к более удобному виду. Введем следующие обозначения:

$$\tilde{f}_y(x)=f_y(x,\ y(x),\ y'(x)),\quad \tilde{f}_{y'}(x)=f_{y'}(x,\ y(x),\ y'(x)).$$

Интегрируя по частям первое из слагаемых (9.11) и используя равенства из (9.8), получаем

$$\int_{x_1}^{x_2}\tilde{f}_y(x)\,h(x)\,d\,x=\int_{x_1}^{x_2}h(x)\,d\left[\int_{x_1}^{x}\tilde{f}_y(t)\,dt\right]=$$

$$=-\int_{x_1}^{x_2}\left[\int_{x_1}^{x}\tilde{f}_y(t)\,d\,t\right]h'(x)\,dx.$$

В соответствии с этим равенство (9.11) можно записать в следующем виде:

$$\int_{x_1}^{x_2}\left[\tilde{f}_{y'}(x)-\int_{x_1}^{x}\tilde{f}_{y}(t)\,\mathrm{d}t\right]h'(x)\,\mathrm{d}x=0 \tag{9.12}$$

для всех h, удовлетворяющих условиям (9.8). Для получения требуемого необходимого условия следует в равенстве (9.12) «избавиться» от внешнего интеграла. Это позволит сделать следующая лемма, которую называют *второй основной леммой вариационного исчисления* *.

Лемма 9.3 (лемма Дюбуа — Реймона). *Пусть* $\omega\in C[x_1, x_2]$ *и равенство*

$$\int_{x_1}^{x_2}\omega(x)\,h'(x)\,\mathrm{d}x=0 \tag{9.13}$$

выполняется для всех h, *удовлетворяющих условиям* (9.8). *Тогда функция* ω *постоянна, т. е. для некоторого числа* C *равенство* $\omega(x)=C$ *верно при всех* $x\in[x_1, x_2]$.

□ Рассмотрим произвольное число C. Так как $h(x_1)=h(x_2)=0$, равенство (9.13) можно записать в виде

$$\int_{x_1}^{x_2}[\omega(x)-C]\,h'(x)\,\mathrm{d}x=0. \tag{9.14}$$

Введем функцию

$$h_0(x)=\int_{x_1}^{x}[\omega(t)-C_0]\,\mathrm{d}t,\quad x\in[x_1,\ x_2].$$

где $C_0=\dfrac{1}{x_2-x_1}\displaystyle\int_{x_1}^{x_2}\omega(t)\,\mathrm{d}t.$

Функция $\omega(t)-C_0$ непрерывна, поэтому $h_0\in C_1[x_1, x_2]$ [16]. Кроме того, $h_0(x_1)=h_0(x_2)=0$; следовательно, функция h_0 удовлетворяет условиям (9.8) и при $C=C_0$ должно выполняться равенство (9.14):

$$\int_{x_1}^{x_2}[\omega(x)-C_0]\,h_0'(x)\,\mathrm{d}x=\int_{x_1}^{x_2}[\omega(x)-C_0]^2\,\mathrm{d}x=0.$$

Отсюда, согласно неравенству $[\omega(x)-C_0]^2\geqslant 0$ для всех $x\in[x_1, x_2]$, получаем [16], что $\omega(x)-C=0$ для всех $x\in[x_1, x_2]$. ■

---

* Первую основную лемму вариационного исчисления см. в § 8.2.

Вернемся к равенству (9.12). Воспользовавшись леммой 9.3, получаем

$$f_{y'}(x,\ y(x),\ y'(x))=\int\limits_{x_1}^{x} f_y(x,\ y(x),\ y'(x))\,\mathrm{d}x+C,\ x\in[x_1,\ x_2], \tag{9.15}$$

где C — некоторая постоянная. Правая часть полученного равенства является непрерывно дифференцируемой функций. Значит, в левой части равенства также непрерывно дифференцируемая функция. Дифференцируя обе части равенства (9.15), имеем

$$\frac{\mathrm{d}}{\mathrm{d}x} f_{y'}(x,\ y(x),\ y'(x))=f_y(x,\ y(x),\ y'(x)),\ x\in[x_1,\ x_2].$$

Сформулируем полученный результат.

Теорема 9.1. *Предположим, что функция f непрерывна на множестве D вместе с частными производными f_y и $f_{y'}$. Если функция $y\in Y$ реализует слабый локальный минимум функционала I (9.4), то она является решением дифференциального уравнения*

$$f_y(x,\ y,\ y')-\frac{\mathrm{d}}{\mathrm{d}x} f_{y'}(x,\ y,\ y')=0, \tag{9.16}$$

называемого уравнением Эйлера — Лагранжа.

Если некоторая функция реализует сильный локальный минимум, то она же реализует и слабый локальный минимум. Поэтому всякое необходимое условие слабого локального минимума является также необходимым условием сильного локального минимума. В частности, такое условие сформулировано в теореме 9.1. В формулировке этой теоремы термин *минимум* можно заменить термином *максимум,* поскольку уравнение (9.16) при замене функции f на функцию $-f$ (т. е. функционала I на функционал $-I$) не меняется.

Тот факт, что функция $y=y(x)$ является решением уравнения Эйлера — Лагранжа, эквивалентен тождественному равенству нулю первой вариации целевого функционала (относительно y), и поэтому функции, удовлетворяющие (9.16), логично было бы назвать стационарными *. Однако, следуя исторически сложившейся в вариационном исчислении традиции, в этой главе решение уравнения Эйлера — Лагранжа (9.16) будем называть *экстремалями* функционала (9.4). Экстремаль не обязательно является слабым локальным решением простейшей вариационной задачи. В § 8.5 рассмотрен пример, где имеется единственная экстремаль, однако оптимального решения не существует.

По теореме 9.1, для определения экстремали нужно решать дифференциальное уравнение второго порядка с заданными краевыми условиями:

$$f_y-\frac{\mathrm{d}}{\mathrm{d}x} f_{y'}=0,\ y(x_1)=y_1,\ y(x_2)=y_2.$$

* См. также § 8.3 предыдушей главы.

Это так называемая *краевая задача.* Для ее решения находят общее решение уравнения Эйлера — Лагранжа, которое зависит от двух произвольных постоянных (поскольку уравнение имеет второй порядок). Затем, используя краевые условия, определяют значения этих двух постоянных.

В настоящее время не существует общего метода решения дифференциального уравнения Эйлера — Лагранжа, поскольку в общем случае оно может быть очень сложным.

3. Частные случаи уравнения Эйлера — Лагранжа. Рассмотрим некоторые частные случаи уравнения Эйлера — Лагранжа, в которых благодаря специальному виду подынтегральной функции f это уравнение можно упростить, в частности понизить его порядок.

1) *Функция f не зависит от y', т. е. $f=f(x, y)$ при $(x, y, y')\in\tilde{D}$.* В этом случае $f_{y'}=0$ и дифференциальное уравнение (9.16) превращается в функциональное уравнение (в котором отсутствуют производные):

$$f_y(x, y)=0.$$

2) *Функция f не зависит от y, т. е. $f=f(x, y')$ при $(x, y, y')\in\tilde{D}$.* Тогда $f_y=0$ и (9.16) принимает вид

$$\frac{d}{dx}f_{y'}(x, y')=0,$$

откуда, интегрируя, получаем дифференциальное уравнение первого порядка

$$f_{y'}(x, y')=C, \tag{9.17}$$

где C — некоторая постоянная.

3) *Функция f не зависит от x, т. е. $f=f(y, y')$ при $(x, y, y')\in\tilde{D}$.* Для упрощения уравнения (9.16) умножим обе его части на y', а затем прибавим и вычтем выражение $f_{y'}(y, y')y''$. Имеем

$$f_y(y, y')y'+f_{y'}(y, y')y''-f_{y'}(y, y')y''-y'\frac{d}{dx}f_{y'}(y, y')=0.$$

Полученное уравнение можно записать в виде

$$\frac{d}{dx}[f(y, y')-y'f_{y'}(y, y')]=0.$$

Таким образом, получаем дифференциальное уравнение первого порядка

$$f(y, y')-y'f_{y'}(y, y')=C, \tag{9.18}$$

где C — некоторая постоянная.

Теперь выясним, когда порядок дифференциального уравнения (9.16) можно понизить в общем случае $f=f(x, y, y')$. Для этого, предполагая функции f и y дважды непрерывно дифференцируемыми на множестве $\tilde{D}$ и отрезке $[x_1, x_2]$ соответственно, выполним в

(9.16) операцию дифференцирования по x (что возможно благодаря сделанным предположениям) и запишем уравнение Эйлера — Лагранжа более подробно:

$$f_y(x,\ y,\ y') - f_{xy'}(x,\ y,\ y') - f_{yy'}(x,\ y,\ y')\,y' - \\ - f_{y'y'}(x,\ y,\ y')\,y'' = 0. \tag{9.19}$$

Отсюда следует, что уравнение имеет первый порядок тогда и только тогда, когда равенство $f_{y'y'}(x,\ y,\ y') = 0$ верно для всех точек $(x,\ y,\ y') \in D$. Интегрируя последнее равенство, находим

$$f(x,\ y,\ y') = P(x,\ y) + Q(x,\ y)\,y',$$

где P и Q — некоторые функции двух переменных. Для этой функции f уравнение (9.19) принимает вид

$$P_y(x,\ y) - Q_x(x,\ y) = 0. \tag{9.20}$$

Здесь возможны два случая.

а) Равенство (9.20) является тождеством, т. е. имеет место для всех точек $(x,\ y) \in D$. Тогда [16] существует такая функция $u = u(x,\ y)$, что

$$\mathrm{d}u(x,\ y) = P(x,\ y)\,\mathrm{d}x + Q(x,\ y)\,\mathrm{d}y,\ (x,\ y) \in D.$$

При этом область D предполагается односвязной (она является таковой в случае, например, если D — выпуклая область). Следовательно,

$$I(y) = \int_{x_1}^{x_2} [P(x,\ y(x)) + Q(x,\ y(x))\,y'(x)]\,\mathrm{d}x =$$

$$= \int_{x_1}^{x_2} \mathrm{d}u(x,\ y) = u(x_2,\ y_2) - u(x_1,\ y_1)$$

для любой допустимой функции y. Это означает, что задача оптимизации бессодержательна: $I(y) = \text{const}$ для всех y.

б) Равенство (9.20) представляет собой функциональное уравнение относительно y. Если это уравнение имеет дважды непрерывно дифференцируемое решение и если для найденного решения выполнены краевые условия $y(x_i) = y_i$, $i = 1,\ 2$, то оно является экстремалью. На практике рассчитывать на такое «стечение обстоятельств» не приходится.

Таким образом, в *общем случае $f = f(x,\ y,\ y')$ порядок дифференциального уравнения Эйлера — Лагранжа, как правило, не понижается.*

4. Решение задачи о брахистохроне. В задаче о брахистохроне

$$f(x,\ y,\ y') = \frac{\sqrt{1 + y'^2}}{\sqrt{2gy}},\ x_1 = y_1 = 0,\ x_2 = \overline{x},\ y_2 = \overline{y}.$$

Здесь $f=f(y, y')$, поэтому можно воспользоваться уравнением (9.18), которое в данном случае имеет вид

$$\frac{\sqrt{1+y'^2}}{\sqrt{2gy}} - y'\frac{y'}{\sqrt{2gy}\sqrt{1+y'^2}} = \frac{1}{\sqrt{2gy}\sqrt{1+y'^2}} = C.$$

Отсюда следует, что $2gy(1+y'^2)=1/C^2$. Вводя новую константу $k_1 = 1/(2\,C^2 g) > 0$, получаем

$$y = \frac{k_1}{1+y'^2}. \tag{9.21}$$

Далее придется дифференцировать уравнение (9.21), поэтому, прежде чем приступить к его решению, обоснуем возможность такого действия. Для этого убедимся в существовании второй производной y'', т. е. проверим, что первая производная y' является дифференцируемой на интервале (x_1, x_2) функцией. Предположим, что решение задачи о брахистохроне $y\in C_1[x_1, x_2]$ существует. Тогда оно удовлетворяет уравнению (9.21), а значит, $y'^2(x)=k_1/y(x)-1$, $x\in[x_1, x_2]$. Поскольку y — непрерывно дифференцируемая функция, согласно последнему равенству, функция y'^2 также непрерывно дифференцируема. Далее. суперпозиция двух непрерывно дифференцируемых функций $z=y'=\sqrt{w}$, $w>0$, $w=y'^2(x)$ является непрерывно дифференцируемой функцией. Очевидно, что $w=y'^2(x)\geqslant 0$ для всех x. Остается показать, что $w\neq 0$, т. е. что

$$y'(x)\neq 0 \text{ для всех } x\in(x_1, x_2). \tag{9.22}$$

Справедливость соотношения (9.22) проверим в три этапа.

1) *Экстремаль не может быть постоянной ни на одном интервале* $(a, b)\subset(x_1, x_2)$, $a<b$. В самом деле, если $y=c$, c — const, для всех $x\in(a, b)$, то, подставляя $y=c$ в левую часть уравнения Эйлера — Лагранжа (9.16), получаем, что оно не удовлетворяется:

$$\left[f_y - \frac{d}{dx}f_{y'}\right]\Big|_{y=c} = \left[-\frac{\sqrt{1+y'^2}}{2(\sqrt{2gy})^3} - \frac{d}{dx}\left(\frac{y'}{\sqrt{2gy}\sqrt{1+y'^2}}\right)\right]\Bigg|_{y=c} =$$

$$= -\frac{1}{2(\sqrt{2gc})^3} \neq 0, \quad x\in(a, b).$$

2) *Производная экстремали может обращаться в нуль не более чем в одной точке.* Если это не так, то можно выбрать такие точки $x', x''\in[x_1, x_2]$, $x'<x''$, что $y'(x')=y'(x'')=0$, причем $y'(x)\neq 0$ при всех $x\in(x', x'')$. Тогда в силу (9.20) верно равенство $y(x')=y(x'')$ и, по теореме Ролля, производная y' обращается в нуль в некоторой внутренней точке интервала (x', x''), что противоречит выбору x' и x''.

3) *Производная экстремали может обратиться в нуль только при* $x=x_2$. Поскольку в выбранной системе координат ось Oy направлена вниз (рис. 9.1), производная на $[x_1, x_2]$ не отрицательна. Значит, функция y не убывает и, согласно равенству (9.21), в точке x, в которой $y'(x_*)=0$, она принимает наибольшее возможное значение. Следовательно, x_* не может быть внутренней точкой и поэтому $x_*=x_2$.

Решение y задачи о брахистохроне является экстремалью, а значит, для этого решения неравенство (9.22) выполняется. Тем самым двукратная непрерывная дифференцируемость оптимального решения доказана.

Используя двукратную непрерывную дифференцируемость решения задачи о брахистохроне, с помощью аналогичных рассуждений можно доказать трехкратную, а затем четырехкратную непрерывную дифференцируемость и т. д. Это означает, что *решением задачи и брахистохроне является бесконечно дифференцируемая функция.*

Решение уравнения (9.21) будем искать в параметрическом виде. Воспользовавшись подстановкой

$$y' = \operatorname{ctg}\frac{t}{2},\ t \in (0,\ 2\pi), \tag{9.23}$$

запишем уравнение (9.21) в виде

$$y = \frac{k_1}{1 + \operatorname{ctg}^2\frac{t}{2}}$$

или

$$y = \frac{k_1}{2}(1 - \cos t). \tag{9.24}$$

Для того чтобы найти функцию $x = x(t)$, продифференцируем (9.24) по x и воспользуемся подстановкой (9.23):

$$\operatorname{ctg}\frac{t}{2} = \frac{k_1}{2}(\sin t)\frac{dt}{dx}.$$

Отсюда имеем

$$dx = k_1 \sin^2\frac{t}{2}\, dt = \frac{k_1}{2}(1 - \cos t)\, dt,$$

а значит,

$$x = \frac{k_1}{2}(t - \sin t) + k_2.$$

Итак, общее решение уравнения Эйлера — Лагранжа (9.21) в параметрической форме имеет вид

$$x = \frac{k_1}{2}(t - \sin t) + k_2,$$

$$y = \frac{k_1}{2}(1 - \cos t),\ k_1 > 0,\ t \in (0,\ 2\pi)$$

и задает на плоскости xOy *семейство циклоид*, зависящих от двух произвольных постоянных k_1 и k_2. Значения этих постоянных определяют из краевых условий $y(0) = 0$, $y(\bar{x}) = \bar{y}$, что приводит к решению следующей системы уравнений:

$$0 = \frac{k_1}{2}(t_1 - \sin t_1) + k_2,$$

$$0 = \frac{k_1}{2}(1 - \cos t_1),$$

$$\bar{x} = \frac{k_1}{2}(t_2 - \sin t_2) + k_2,$$

$$\bar{y} = \frac{k_1}{2}(1 - \cos t_2).$$

Можно доказать, что эта система уравнений относительно k_1, k_2, t_1, t_2 (при $\bar{x}>0$, $\bar{y}>0$) всегда имеет единственное решение. Таким образом, для указанных краевых условий существует единственная экстремаль, соединяющая точки (0, 0) и $(\bar{x}, \bar{y})$. Только на основании выполнения необходимого условия оптимальности делать вывод о том, что экстремаль является решением задачи о брахистохроне, нельзя. Однако физический смысл данной задачи подсказывает, что оптимальное решение должно существовать; следовательно, им и является найденная экстремаль (циклоида).

5. Решение задачи о минимальной поверхности вращения. Эта задача сформулирована в п. 3 § 9.1 и сводится к минимизации функционала

$$I(y)=2\pi\int_{x_1}^{x_2} y(x)\sqrt{1+y'^2(x)}\,dx$$

на множестве функций $y=y(x)$ с граничными условиями $y(x_i)=y_i$, $i=1, 2$. Оптимальное решение задачи будем искать в пространстве гладких функций, хотя в первоначальной ее постановке фигурирует более широкое множество (пространство кусочно-гладких функций). На основании следствия 9.1 найденное решение будет оптимальным и в пространстве кусочно-гладких функций.

Данная задача является простейшей вариационной задачей, в которой $f_y=2\pi\sqrt{1+y'^2}$, $f_{y'}=2\pi yy'/\sqrt{1+y'^2}$. Подынтегральная функция $f=2\pi y\sqrt{1+y'^2}$ явно не зависит от x, поэтому мы не будем записывать уравнение Эйлера — Лагранжа, а сразу воспользуемся эквивалентным ему уравнением (9.18), которое в данном случае имеет вид

$$2\pi y\sqrt{1+y'^2}-\frac{2\pi yy'^2}{\sqrt{1+y'^2}}=C,$$

или

$$y=C_1\sqrt{1+y'^2}, \tag{9.25}$$

где $C_1=C/(2\pi)$. Запишем общее решение дифференциального уравнения первого порядка (9.25), не разрешенного относительно производной:

$$y=C_1\operatorname{ch}\left(\frac{x}{C_1}+C_2\right), \tag{9.26}$$

где C_2 — произвольная постоянная. Уравнение (9.26) задает *семейство цепных линий*, зависящее от параметров C_1 и C_2, которые следует определить из граничных условий, т. е. в результате решения следующей системы уравнений:

$$C_1\operatorname{ch}\left(\frac{x_1}{C_1}+C_2\right)=y_1;\ C_1\operatorname{ch}\left(\frac{x_2}{C_1}+C_2\right)=y_2. \tag{9.27}$$

Если C_1 и C_2 находятся однозначно, то, подставляя их значения в

(9.26), получаем оптимальное решение задачи. Однако система (9.27) при определенных комбинациях чисел x_1, x_2, y_1, y_2 может не иметь решения (в этом случае гладкого решения задачи о минимальной поверхности не существует) или же обладать не единственным решением. В последнем случае для нахождения оптимального решения необходимы дополнительные исследования.

§ 9.3. ВТОРАЯ ВАРИАЦИЯ. УСЛОВИЕ ЛЕЖАНДРА

1. Определение второй вариации. Пусть функция f, фигурирующая в определении целевого функционала (9.4), дважды непрерывно дифференцируема на множестве $\hat{D}$. Рассмотрим произвольную функцию $y \in Y$, где множество Y имеет вид (9.5). Функционал $\delta^2 I$, заданный на множестве функций (9.8) и определяемый формулой

$$\delta^2 I(h) = \int_{x_1}^{x_2} [p(x) h^2(x) + 2q(x) h(x) h'(x) + r(x) h'^2(x)] \, dx, \qquad (9.28)$$

где

$$\begin{aligned} p(x) &= f_{yy}(x, y(x), y'(x)), \\ q(x) &= f_{yy'}(x, y(x), y'(x)), \\ r(x) &= f_{y'y'}(x, y(x), y'(x)), \end{aligned} \qquad (9.29)$$

называют *второй вариацией функционала I* (относительно элемента y).

При определении первой вариации подынтегральное выражение в (9.10) было линейным относительно h, h'. Под интегралом в формуле (9.28) находится квадратичная относительно h, h' функция.

Введем ряд определений, которые потребуются в дальнейшем. Функционал $\delta^2 I$ *неотрицательно (неположительно) определен*, если для всех функций h вида (9.8) верно неравенство

$$\delta^2 I(h) \geqslant 0 (\delta^2 I(h) \leqslant 0).$$

Если же для всех функций h вида (9.8), кроме $h(x) \equiv 0$ на $[x_1, x_2]$, выполняется строгое неравенство

$$\delta^2 I(h) > 0 (\delta^2 I(h) < 0),$$

то функционал $\delta^2 I$ положительно (отрицательно) определен. В случае когда функционал $\delta^2 I$ принимает на множестве (9.8) как положительные, так и отрицательные значения, его называют *знакопеременным*.

Используя введенные понятия, сформулируем следующее необходимое условие оптимальности.

Лемма 9.4. *Пусть функция f дважды непрерывно дифференцируема на множестве $\hat{D}$. Если функция $y \in Y$ реализует слабый локальный минимум в простейшей вариационной задаче, то вторая* *вариация $\delta^2 I$ (относительно y) является неотрицательно-определенным функционалом.*

□ Рассмотрим произвольную функцию h вида (9.8). Повторяя рассуждения, аналогичные приведенным в п. 1 § 9.2, и учитывая, что функция y реализует слабый локальный минимум, приходим к утверждению, что функция вида $\varphi(\lambda)=I(y+\lambda h)$ достигает локального минимума при $\lambda=0$. Согласно сделанным предположениям относительно f, функция φ дважды дифференцируема. В соответствии с результатами п. 3 § 2.5 вторая производная функции φ в точке $\lambda=0$ должна быть неотрицательна: $\varphi''(0)\geqslant 0$. Записывая это неравенство более подробно и учитывая при этом равенство

$$\frac{d}{d\lambda^2} f(x,\ y+\lambda h,\ y'+\lambda h')=f_{yy}h^2+2f_{yy'}hh'+f_{y'y'}h'^2,$$

получаем

$$\varphi''(0)=\left\{\frac{d^2}{d\lambda^2}\int\limits_{x_1}^{x_2} f(x,\ y(x)+\lambda h(x),\ y'(x)+\lambda h'(x)\,dx\right\}\Bigg|_{\lambda=0}=$$

$$=\int\limits_{x_1}^{x_2}[f_{yy}(x,\ y(x),\ y'(x))\,h^2(x)+2f_{yy'}(x,\ y(x),\ y'(x))\,h(x)\times$$

$$\times h'(x)+f_{y'y'}(x,\ y(x),\ y'(x))\,h'^2(x)d]\,x=\delta^2 I(h)\geqslant 0.$$

Как и в случае первой вариации, теперь следует вывести условия, эквивалентные неравенству $\delta^2 I(h)\geqslant 0$ и удобные для практического применения. Однако сделать это гораздо труднее, чем получить уравнение Эйлера — Лагранжа. ■

2. Условие Лежандра. Предполагая функцию f дважды непрерывно дифференцируемой на множестве $\hat{D}$, говорят, что функция $y\in Y$ удовлетворяет *условию Лежандра (сильному условию Лежандра)*, если неравенство

$$f_{y'y'}(x,\ y(x),\ y'(x))\geqslant 0\ \ (f_{y'y'}(x,\ y(x),\ y'(x))>0)$$

имеет место для всех $x\in[x_1,\ x_2]$.

Теорема 9.2. *Предположим, что функция f дважды непрерывно дифференцируема на множестве $\hat{D}$. Если функция $y\in Y$ реализует слабый локальный минимум в простейшей вариационной задаче, то она удовлетворяет условию Лежандра.*

□ Пусть функция $y\in Y$ реализует слабый локальный минимум функционала I. Предположим, что теорема не верна, т. е. найдется такая точка $\bar{x}\in[x_1,\ x_2]$, в которой $r(\bar{x})=f_{y'y'}(\bar{x},\ y(\bar{x}),\ y'(\bar{x}))<0$. Функция r непрерывная, поэтому существуют такие положительные числа γ и τ_0 и точка $x_0\in(x_1,\ x_2)$, что

$$r(x)<-\gamma \text{ для всех } x\in U_{\tau_0}(x_0)\subset[x_1,\ x_2].$$

Введем семейство кусочно-гладких функций

$$h_\tau(x)=\begin{cases}\sqrt{\tau}-\dfrac{|x-x_0|}{\sqrt{\tau}}, & \text{если } x\in U_\tau(x_0),\\ 0, & \text{если } x\notin U_\tau(x_0),\end{cases}$$

зависящее от параметра $\tau\in(0,\ \tau_0)$ (рис. 9.6). Неравенства $|h_\tau(x)|\leqslant\sqrt{\tau}$; $|h'_x(x)|\leqslant 1/\sqrt{\tau}$; $|h_\tau(x)h'_\tau(x)|\leqslant 1$ выполняются для всех $x\in U_\tau(x_0)$, $x\neq x_0$. Используя эти неравенства, а также вводя обозначения

$$\alpha=\sup_{x\in U_\tau(x_0)} p(x),\quad \beta=\sup_{x\in U_\tau(x_0)} q(x),$$

получаем *

$$\delta^2 I(h_\tau)=\int\limits_{x_0-\tau}^{x_0+\tau}\left[p(x)h_\tau^2(x)+2q(x)h_\tau(x)h'_\tau(x)+r(x)h'^2_\tau(x)\right]\mathrm{d}x\leqslant$$

$$\leqslant 2\tau^2\alpha+4\tau\beta-2\gamma.$$

Отсюда видно, что, выбирая число τ достаточно малым, можно добиться выполнения неравенства $\delta^2 I(h_\tau)<0$. Теперь, если к функционалу $\delta^2 I$ применить лемму 9.1 о скруглении углов, то получим, что существует такая функция h вида (9.8) **, что $\delta^2 I(h)<0$. Но это не имеет места, поскольку, согласно лемме 9.4, для всех функций h вида (9.8) верно неравенство $\delta^2 I(h)\geqslant\geqslant 0$. Полученное противоречие завершает доказательство. ■

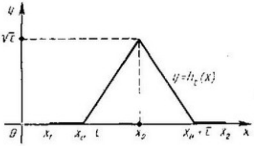

Рис. 9.6

При доказательстве теоремы 9.2 было установлено, что неотрицательность второй вариации (относительно y) влечет выполнение условия Лежандра для функции y. Обратное утверждение, вообще говоря, не имеет места. Даже из выполнения сильного условия Лежандра не обязательно следует неотрицательность второй вариации. Рассмотрим следующий пример.

Пример. Пусть в простейшей вариационной задаче

$$f=y'^2-y^2,\ x_1=y_1=y_2=0,\ x_2=2\pi.$$

---

* Вторая вариация ранее была определена на множестве гладких функций h вида (9.8). Здесь функционал $\delta^2 I$ вычисляется на кусочно-гладкой функции h_τ. Это допустимо, поскольку для кусочно-гладкой функции h и непрерывных функций p, q, r интеграл (9.28) существует (как интеграл от кусочно-непрерывной функции).

** Здесь расширенное допустимое множество составляют кусочно-гладкие функции, которые на концах отрезка $[x_1,\ x_2]$ обращаются в нуль. Значит, функция h, о существовании которой говорилось выше, также обращается в нуль на концах этого отрезка.

Здесь $p(x)=-2$, $q(x)=0$, $r(x)=2$ для всех $x\in(0, 2\pi)$. Функции p, q и r не зависят от x, поэтому вторая вариация $\delta^2 I$ одинакова относительно любой функции y:

$$\delta^2 I(h)=2\int_0^{2\pi}(-h^2(x)+h'^2(x))\,dx.$$

В частности, такой вид имеет вторая вариация относительно любой экстремали (которая здесь задается формулой $y=C\sin x$, где C — произвольная константа). Вычисляя эту вторую вариацию, например, для функции $h_0(x)=x^2-2\pi x$, получаем, что она отрицательна:

$$\delta^2 I(h_0)=2\int_0^{2\pi}[-(x^2-2\pi x)^2+(2x-2\pi)^2]\,dx=\frac{8\pi^3}{3}-\frac{32\pi^5}{5}<0,$$

хотя условие Лежандра (даже в сильной форме) выполняется: $f_{y'y'}=r(x)=2>0$.

Между задачей безусловной минимизации функции φ одной переменной и простейшей вариационной задачей имеется аналогия. А именно: необходимому условию локального минимума $\varphi'(x)=0$ соответствует необходимое условие слабого локального минимума — уравнение Эйлера — Лагранжа $f_y-\frac{d}{dx}f_{y'}=0$. Необходимое условие второго порядка имеет вид неравенства $\varphi''(x)\geqslant 0$, и ему можно сопоставить условие Лежандра $f_{y'y'}\geqslant 0$. Однако ошибочно предполагать, что экстремаль, для которой имеет место сильное условие Лежандра $f_{y'y'}>0$, должна реализовать слабый локальный минимум, так как в стационарной точке x, такой что $\varphi''(x)>0$, функция φ достигает локального минимума.

Дальнейшее исследование необходимых (достаточных) условий слабого локального минимума приводит к условию Якоби, которое связано с экстремалями функционала второй вариации.

§ 9.4. УСЛОВИЕ ЯКОБИ

1. Уравнение Якоби. Предположим, что подынтегральная функция f из (9.4) трижды непрерывно дифференцируемая на множестве D. Рассмотрим экстремаль вида $y\in C_2[x_1, x_2]$ функционала (9.4) и допустим, что для нее выполняется сильное условие Лежандра:

$$r(x)=f_{y'y'}(x, y(x), y'(x))>0 \text{ для всех } x\in[x_1, x_2].$$

Тогда дифференциальное уравнение Эйлера — Лагранжа, записанное применительно к функционалу второй вариации $\delta^2 I$ (относительно y), называют *дифференциальным уравнением Якоби*, связанным с исходным функционалом I относительно y.

Запишем в обозначениях (9.29) уравнение Эйлера — Лагранжа для второй вариации (9.28):

$$[ph^2+2qhh'+rh'^2]_h-\frac{d}{dx}[ph^2+2qhh'+rh'^2]_{h'}=$$

$$=2ph+2qh'-\frac{d}{dx}[2qh+2rh']=0;$$

дифференцируя по x и упрощая, имеем

$$[p(x)-q'(x)]h-r'(x)h'+r(x)h''=0. \tag{9.30}$$

Согласно сильному условию Лежандра, верно неравенство $r(x)\neq 0$ для всех $x\in[x_1, x_2]$, поэтому уравнение Якоби (9.30) является обыкновенным линейным однородным дифференциальным уравнением второго порядка, которое можно записать в виде

$$h''+a_1(x)h'+a_2(x)h=0,$$

где $a_1(x)=r'(x)/r(x)$, $a_2(x)=(q'(x)-p(x))/r(x)$.

При сделанных предположениях функции a_1 и a_2 непрерывны, а значит, согласно соответствующей теореме* из курса дифференциальных уравнений, задача Коши для уравнения Якоби всегда имеет, и притом единственное, решение. В частности, существует единственное решение $h=\Delta(x)$, удовлетворяющее начальным условиям

$$\Delta(x_1)=0,\quad \Delta'(x_1)=1. \tag{9.31}$$

В терминах этого решения и формулируется условие Якоби.

2. Условие Якоби. Пусть выполняются предположения, сделанные в начале п. 1. Говорят, что функция y удовлетворяет *условию Якоби,* если решение Δ уравнения Якоби (9.30), подчиненное начальным условиям (9.31), не обращается в нуль внутри отрезка $[x_1, x_2]$, т. е. $\Delta(x)\neq 0$ для всех $x\in(x_1, x_2)$. Если, кроме того, верно неравенство $\Delta(x_2)\neq 0$, то говорят, что функция y удовлетворяет *сильному условию Якоби.*

Уравнение Якоби (9.30) записано для некоторой фиксированной экстремали y функционала I, поэтому решение Δ, а также условие Якоби связаны именно с этой экстремалью. Было бы разумным вместо Δ писать, например, $\Delta(y)$, подчеркивая тем самым зависимость решения уравнения Якоби от экстремали y. Однако для краткости записи мы этого делать не будем.

В приведенном выше определении условия Якоби отмечается, что некоторое специальным образом выбранное решение Δ уравнения Якоби не обращается в нуль всюду на (x_1, x_2). На самом деле *из выполнения условия Якоби следует,*

---

* В теореме, о которой здесь идет речь, все функции рассматриваются на открытом интервале, тогда как в данной задаче фигурирует отрезок $[x_1, x_2]$. Поэтому в качестве указанного интервала следует взять любой интервал, содержащий этот отрезок. Такой интервал существует, поскольку множество D открытое.

что любое отличное от тождественно равного нулю решение уравнения Якоби, которое равно нулю в точке x_1, нигде больше на интервале (x_1, x_2) в нуль не обращается. Рассмотрим произвольное ненулевое решение h уравнения Якоби (9.30), удовлетворяющее условию $h(x_1)=0$ и решение Δ вида (9.31). Определитель Вронского для этих двух решений при $x=x_1$ равен нулю:

$$\begin{vmatrix} h(x_1) & \Delta(x_1) \\ h'(x_1) & \Delta'(x_1) \end{vmatrix} = \begin{vmatrix} 0 & 0 \\ h'(x_1) & \Delta'(x_1) \end{vmatrix} = 0,$$

а значит, согласно соответствующей теореме из курса дифференциальных уравнений, система функций h, Δ является линейно зависимой на отрезке $[x_1, x_2]$. Отсюда следует, что множество точек, в которых одна из этих функций обращается в нуль, совпадает с подобным множеством для другой функции. Тем самым справедливость сформулированного выше утверждения доказана.

В соответствии с этим при определении условия Якоби в (9.31) вместо равенства $\Delta'(x_1)=1$ можно записать равенство $\Delta'(x_1)=c$, где $c\neq 0$. Значение $c=1$ в (9.31) выбрано из соображений простоты.

Точки x', $x''\in[x_1, x_2]$, $x_1'\neq x''$, называют *парой сопряженных точек* относительно ненулевого решения h уравнения Якоби (9.30), если $h(x')=h(x'')=0$. Таким образом, выполнение условия Якоби означает, что ненулевое решение Δ, удовлетворяющее начальным условиям (9.31), не имеет сопряженных точек вида x_1, x'', где $x\in(x_1, x_2)$. Оказывается, *из выполнения условия Якоби следует, что любое ненулевое решение уравнения Якоби не может иметь ни одной пары сопряженных точек вида x', x''*, где x', $x''\in[x_1, x_2]$.

Используя условие Якоби, можно сформулировать как необходимые, так и достаточные условия слабого локального экстремума.

Теорема 9.3. *Предположим, что функция f трижды непрерывно дифференцируема на множестве $\bar{D}$, а экстремаль y функционала* (9.4) *дважды непрерывно дифференцируема на отрезке $[x_1, x_2]$. Будем считать, что экстремаль y удовлетворяет сильному условию Лежандра. Тогда если функция y реализует слабый локальный минимум функционала* (9.4) *на множестве Y* (9.5), *то она удовлетворяет условию Якоби; обратно: если функция y удовлетворяет сильному условию Якоби, то она реализует слабый локальный минимум функционала* (9.4) *на множестве Y.*

□ Доказательство достаточности (см., например, [15]. Для того чтобы доказать необходимость, предположим, что функция y реализует слабый локальный минимум, однако условие Якоби для нее не выполняется, т. е. существует такая точка $x_0\in(x_1, x_2)$, что $\Delta(x_0)=0$. Если при этом оказывается $\Delta'(x_0)=0$, то в силу единственности решения задачи Коши для уравнения Якоби с начальными условиями $\Delta(x_0)=\Delta'(x_0)=0$ имеет место тождественное равенство $\Delta(x)\equiv 0$. Но, по определению условия Якоби, Δ — ненулевое решение, значит, $\Delta'(x_0)\neq 0$.

Введем функцию $h\in KC_1[x_1, x_2]$ с помощью формулы

$$h(x)=\begin{cases} \Delta(x) & \text{при } x\in[x_1, x_0], \\ 0 & \text{при } x\in[x_0, x_2]. \end{cases}$$

Для того чтобы установить справедливость равенства

$$\delta^2 I(h)=0, \tag{9.32}$$

преобразуем с помощью формулы интегрирования по частям второе и третье слагаемые в выражении (9.28). Получаем

$$2\int\limits_{x_1}^{x_0} q(x)\,h(x)\,h'(x)\,\mathrm{d}x=\int\limits_{x_1}^{x_0} q(x)\,\mathrm{d}h^2(x)=-\int\limits_{x_1}^{x_0} q'(x)\,h^2(x)\,\mathrm{d}x,$$

$$\int\limits_{x_1}^{x_0} r(x)\,h'^2(x)\,\mathrm{d}x=\int\limits_{x_1}^{x_0} r(x)\,h'(x)\,\mathrm{d}h(x)=-\int\limits_{x_1}^{x_0}\left[\frac{\mathrm{d}}{\mathrm{d}x}\,r(x)\,h'(x)\right]h(x)\mathrm{d}x.$$

Здесь были использованы равенства $h(x_1)=h(x_0)=0$. Далее имеем

$$\delta^2 I(h)=\int\limits_{x_1}^{x_0}[p(x)\,h^2(x)+2q(x)\,h(x)\,h'(x)+r(x)\,h'^2(x)]\,\mathrm{d}x=$$

$$=\int\limits_{x_1}^{x_0}\left[(p(x)-q'(x)\Delta(x))-\frac{\mathrm{d}}{\mathrm{d}x}(r(x)\,\Delta'(x)\right]\Delta(x)\,\mathrm{d}x.$$

В квадратных скобках находится левая часть уравнения Якоби (9.30) при $h(x)=\Delta(x)$, $x_1\leqslant x\leqslant x_0$. Согласно условию, Δ — решение уравнения Якоби, поэтому равенство (9.32) действительно выполняется.

Рассмотрим семейство функций из $KC_1[x_1,\ x_2]$, определяемых формулой

$$h_\varepsilon(x)=\begin{cases} h(x)\ \text{при}\ x\in[x_1,\ x_0-\varepsilon], \\ \dfrac{(-x+x_0+c)\,h(x_0-\varepsilon)}{\xi+c}\ \text{при}\ x\in[x_0-\varepsilon,\ x_2], \end{cases}$$

зависящее от параметра ε, $\varepsilon\in(0,\ x_0-x_1)$, где $c=x_2-x_0$ (рис. 9.7). Теперь с помощью равенства $\psi(\varepsilon)=\delta^2 I(h_\varepsilon)$, $\varepsilon\in(0,\ x_0-x_1)$), определим числовую функцию ψ и докажем, что $\psi'(+0)<0$.

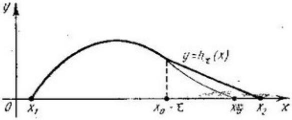

Рис. 9.7

Введем следующие обозначения:

$$\Omega=p(x)\,h^2(x)+2q(x)\times$$
$$\times h(x)\,h'(x)+r(x)\,h'^2(x),$$

$$\Omega_\varepsilon=p(x)\,h_\varepsilon^2(x)+2q(x)\,h_\varepsilon(x)\,h_\varepsilon'(x)+r(x)\,h_\varepsilon'^2(x).$$

Учитывая равенства $\psi(0)=\delta^2 I(h)=0$, составим разность:

$$\psi(\varepsilon)-\psi(0)=\psi(\varepsilon)=\int\limits_{x_1}^{x_2}\Omega_\varepsilon\mathrm{d}x=\int\limits_{x_1}^{x_0-\varepsilon}\Omega\mathrm{d}x+\int\limits_{x_0-\varepsilon}^{x_0}\Omega\mathrm{d}x-\int\limits_{x_0-\varepsilon}^{x_0}\Omega\mathrm{d}x+$$

$$+\int\limits_{x_0-\varepsilon}^{x_2}\Omega_\varepsilon\,\mathrm{d}x=\delta^2 I(h)-\int\limits_{x_0-\varepsilon}^{x_0}\Omega\,\mathrm{d}x+\int\limits_{x_0-\varepsilon}^{x_2}\Omega_\xi\,\mathrm{d}x=\int\limits_{x_0-\varepsilon}^{x_2}\Omega_\xi\,\mathrm{d}x-\int\limits_{x_0-\varepsilon}^{x_0}\Omega\,\mathrm{d}x.$$

Отсюда, используя первую теорему о среднем для интеграла [16], получаем

$$\frac{\psi(\varepsilon)-\psi(0)}{\varepsilon}=[\tilde{p}h_\varepsilon^2(\xi_1)+2\tilde{q}h_\xi(\xi_2)+\tilde{r}h_\varepsilon^2(\xi_3)]\,\frac{\varepsilon+c}{\varepsilon}-$$

$$-[\tilde{\tilde{p}}h^2(\xi_4)+2\tilde{\tilde{q}}h(\xi_5)\,h'(\xi_5)+\tilde{\tilde{r}}h'^2(\xi_6), \tag{9.33}$$

где $\tilde{p}=p(\xi_1)$, $\tilde{q}=q(\xi_2)$, $\tilde{r}=r(\xi_3)$; $\xi_1,\xi_2,\xi_3\in(x_0-\varepsilon,\,x_2)$, $\tilde{\tilde{p}}=p(\xi_4)$, $\tilde{\tilde{q}}=q(\xi_5)$, $\tilde{\tilde{r}}=r(\xi_6)$; $\xi_4,\xi_5,\xi_6\in(x_0-\varepsilon,\,x_0)$. Выбор точек ξ_1, ξ_2, ..., ξ_6 зависит от ε; это следует учитывать ниже при переходе к пределу при $\varepsilon\to+0$.

Установим поведение слагаемых в правой части равенства (9.33) при $\varepsilon\to+0$. Поскольку $\xi_1\in(x_0-\xi,\,x_2)$, по определению функции h_ε, верно равенство

$$h_\varepsilon(\xi_1)=\frac{(-\xi_1+x_0+c)\,h(x_0-\varepsilon)}{\varepsilon+c}.$$

Далее, по теореме Лагранжа найдется такая точка $\eta\in(x_0-\varepsilon,\,x_0)$, что

$$h(x_0-\varepsilon)=h(x_0-\varepsilon)-h(x_0)=-h'(\eta)\,\varepsilon. \tag{9.34}$$

Следовательно,

$$\lim_{\varepsilon\to+0}\tilde{p}\,h_\varepsilon^2(\xi_1)\,\frac{\varepsilon+c}{\varepsilon}=\lim_{\varepsilon\to+0}\frac{\tilde{p}(-\zeta_1+x_0+c)^2\,h'^2(\eta)}{\varepsilon+c}\,\varepsilon=0, \tag{9.35}$$

так как дробь в правой части равенства (9.35) является ограниченной при $\varepsilon\to+0$.

Аналогично, используя формулу (9.34), можно записать равенство

$$h_\varepsilon(\xi_2)=\frac{(-\xi_2+x_0+c)\,(-h'(\eta)\,\varepsilon)}{\varepsilon+c}.$$

По определению функции h_ε имеем $h_\varepsilon'(\xi_2)=-h(x_0-\varepsilon)/(\varepsilon+c)$, а значит, согласно формуле (9.34), верно равенство

$$h_\varepsilon'(\xi_2)=-\frac{h'(\eta)\,\varepsilon}{\varepsilon+c}.$$

Имеют место равенства

$$\lim_{\varepsilon\to+0}2\tilde{q}h_\varepsilon(\xi_2)\,h_\varepsilon'(\xi_2)\,\frac{\varepsilon+c}{\varepsilon}=\lim_{\varepsilon\to+0}\frac{2\tilde{q}\,(-\xi_2+x_0+c)\,h'^2(\eta)}{\varepsilon+c}\,\varepsilon=0, \tag{9.36}$$

поскольку и здесь функция, записанная в виде дроби, ограничена.

То же верно и для третьего слагаемого в правой части (9.33):

$$\lim_{\varepsilon\to+0} \tilde{r} h_\varepsilon'^2 (\xi_3) \frac{\varepsilon + c}{\varepsilon} = \lim_{\varepsilon\to+0} \frac{\tilde{r} h'^2(\eta)}{\varepsilon + c} \varepsilon = 0. \tag{9.37}$$

Далее, так как $h(x_0)=0$, имеем

$$\lim_{\varepsilon\to+0} \tilde{\tilde{p}} h^2 (\xi_4) = \lim_{\varepsilon\to+0} 2\tilde{q} h(\xi_5) h'(\xi_5) = 0; \tag{9.38}$$

$$\lim_{\varepsilon\to+0} \tilde{\tilde{r}} h'^2 (\xi_6) = r(x_0) h'^2 (x_0). \tag{9.39}$$

На основании равенств (9.35)—(9.39) из (9.33) при $\varepsilon\to+0$ получаем $\psi'(+0) = -r(x_0) h'^2(x_0)$, откуда, используя сильное условие Лежандра и неравенство $h'(x_0) = \Delta'(x_0) \neq 0$, получаем требуемое неравенство $\psi'(+0) < 0$. Итак, можно сделать вывод, что для достаточно малого $\varepsilon_0 > 0$ выполняется неравенство

$$\delta^2 I(h_{\varepsilon_0}) = \psi(\varepsilon_0) - \psi(0) = \psi'(+0)\varepsilon_0 + o(\varepsilon_0) < 0.$$

Отсюда на основании леммы 9.1 можно сделать вывод, что существует функция h_0, удовлетворяющая условиям (9.8), для которой вторая вариация (относительно y) отрицательна: $\delta^2 I(h_0) < 0$. Это противоречит тому, что функция y реализует слабый локальный минимум. ■

Покажем, что достаточные условия слабого локального минимума теоремы 9.3 не являются достаточными для того, чтобы функция y реализовала сильный локальный минимум.

Пример. Рассмотрим простейшую вариационную задачу, в которой

$$f(x, y, y') = y'^2 + y'^3, \; x_1 = y_1 = y_2 = 0, \; x_2 = 1.$$

Для этой задачи уравнение Эйлера — Лагранжа имеет вид

$$f_y - \frac{d}{dx} f_{y'} = -\frac{d}{dx}(2y' + 3y'^2) = -2y'' - 6y'y'' = 0.$$

Общее решение полученного дифференциального уравнения таково: $y = C_1 x + C_2$, где C_1, C_2 — произвольные постоянные. Используя краевые условия, получаем $C_1 = C_2 = 0$, т. е. $y \equiv 0$ $(x \in [0, 1])$ — единственная экстремаль. Убеждаемся, что для нее сильное условие Лежандра выполняется:

$$f_{y'y'}|_{y\equiv 0} = 2 + 6y'|_{y\equiv 0} = 2 > 0.$$

Для того чтобы записать уравнение Якоби для найденной экстремали, находим: $p(x) \equiv 0$, $q(x) \equiv 0$, $r(x) = 6 + 4y'|_{y\equiv 0} \equiv 2$. Следовательно, вторая вариация (относительно $y \equiv 0$) равна

$$\delta^2 I(h) = 2 \int_0^1 h'^2 x) \, dx$$

и уравнение Якоби принимает вид

$$0-\frac{d}{dx}2h'=-2h''=0.$$

Найдем общее решение этого уравнения: $h=C_3x+C_4$, где C_3, C_4 — произвольные постоянные. Функция $h(x)=\Delta(x)=x$ является решением, удовлетворяющим начальным условиям $h(0)=0$, $h'(0)=1$. Для найденной функции $\Delta(x)$ имеем $\Delta(x)\neq 0$ для любого $x\in(0, 1)$, т. е. сильное условие Якоби выполняется. На основании теоремы 9.3 делаем вывод, что функция $y\equiv 0$ реализует слабый локальный минимум.

Убедимся в том, что сильный локальный минимум функция $y\equiv 0$ не реализует. Для этого рассмотрим последовательность $\{y_n\}_{n=2}^{\infty}$ функций из пространства $KC_1[0, 1]$ вида

$$y_n(x)=\int_0^x g_n(t)\,dt,\quad n=2, 3,\ldots,$$

где

$$g_n(t)=\begin{cases} -\sqrt{n}, & \text{если}\quad t\in[0,\ 1/n],\\ 0, & \text{если}\quad t\in[1/n,\ 1/2],\\ 2/\sqrt{n}, & \text{если}\quad t\in(1/2,\ 1].\end{cases}$$

Очевидно, $y_n(0)=0$ и, кроме того,

$$y_n(1)=\int_0^1 g_n(t)\,dt=-\int_0^{1/n}\sqrt{n}\,dt+\int_{1/2}^1\frac{2dt}{\sqrt{n}}=0,\quad n=2, 3,\ldots,$$

т. е. функции данной последовательности принадлежат расширенному допустимому множеству. Поскольку

$$\|y_n-y\|_0=\max_{x\in[0,1]}|y_n(x)|=\frac{2}{\sqrt{n}}\underset{n\to\infty}{\longrightarrow}0,$$

в любой сколь угодно малой окрестности нулевого порядка элемента $y=0$ содержатся члены рассматриваемой последовательности. Вычислим значения целевого функционала на элементах последовательности:

$$I(y_n)=\int_0^1 g_n^2(t)\,dt+\int_0^1 g_n^3(t)\,dt=\int_0^{1/n}(n-n^{3/2})\,dt+$$

$$+\int_{1/2}^1\left(\frac{4}{n}+\frac{8}{n^{3/2}}\right)dt=1-\sqrt{n}+\frac{2}{n}+\frac{4}{n^{3/2}}\underset{n\to\infty}{\longrightarrow}-\infty.$$

На основании леммы 9.1 о скруглении углов заключаем, что функция $y\equiv 0$ не реализует сильный локальный минимум.

3. **Задача определения критической нагрузки балки.** Балка AB длины l, имеющая цилиндрическую форму, расположена вертикально. Ее нижний конец A жестко закреплен, а верхний под действием нагрузки P может перемещаться по вертикальной прямой таким образом, что касательная, проведенная к осевой линии балки в точке B, все время проходит через отрезок AB (рис. 9.8). Требуется определить критическую нагрузку P^*, при которой вертикальное положение балки является положением устойчивого равновесия.

Обозначим через s длину дуги осевой линии балки, отсчитываемую от точки A, а через $\varphi=\varphi(s)$ меньший из углов, образованных отрезком AB и касательной к осевой линии, проведенной в точке, которая расположена на расстоянии s от A. Потенциальная энергия изогнутой балки равна

$$W_{\pi}(\varphi)=\int_0^l \left[\frac{1}{2}\mu\left(\frac{d\varphi}{ds}\right)^2+P\cos\varphi\right]ds, \qquad (9.40)$$

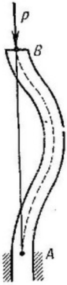

Рис. 9.8

где $\mu>0$ — константа, зависящая от коэффициента упругости и момента инерции поперечного сечения. Состояние устойчивого равновесия характеризуется минимумом потенциальной энергии, поэтому задача сводится к определению таких значений P, при которых функция $\varphi(s)\equiv 0$ реализует наименьшее возможное значение функционала (9.40) на множестве функций вида $\varphi\in C_1[0, l]$, $\varphi(0)=\varphi(l)=0$. Для решения этой задачи воспользуемся теоремой 9.3.

Запишем уравнение Эйлера — Лагранжа:

$$-P\sin\varphi+\frac{d}{ds}\left(\mu\frac{d\varphi}{ds}\right)=0.$$

Очевидно, функция $\varphi(s)\equiv 0$ удовлетворяет этому уравнению, а значит, является экстремалью. Кроме того, выполняется сильное условие Лежандра

$$f_{\varphi'\varphi'}|_{\varphi\equiv 0}=\mu>0.$$

Находим (относительно функции $\varphi(s)\equiv 0$) $p(s)=-P$, $q(s)=0$, $r(s)=\mu$ и записываем уравнение Якоби:

$$\mu h''+Ph=0.$$

Общее решение этого уравнения имеет вид

$$\Delta(s)=C_1\sin\sqrt{P/\mu}\,s+C_2\cos\sqrt{P/\mu}\,s,$$

а его решением при начальных условиях $\Delta(0)=0$, $\Delta'(0)=1$ является функция

$$\Delta(s)=\sqrt{\mu/P}\sin\sqrt{P/\mu}\,s.$$

Согласно теореме 9.3, решение $\varphi(s)\equiv 0$ является оптимальным, если $\Delta(s)\neq 0$ для всех $s\in(0,\ l)$; это равносильно условию $\sin\sqrt{P/\mu}\,l>0$. Наименьшее положительное P, при котором последнее неравенство нарушается, определяет критическую нагрузку. Следовательно, $\sqrt{P^*/\mu}\,l=\pi$, откуда находим $P^*=\pi^2\mu/l^2$.

§ 9.5. НЕКОТОРЫЕ МОДИФИКАЦИИ ПРОСТЕЙШЕЙ ВАРИАЦИОННОЙ ЗАДАЧИ

1. Простейшая задача Больца. В задаче Больца целевой функционал $B=I+\Psi$, подлежащий минимизации, является суммой интегрального и терминального функционалов:

$$B(y)=\int_{x_1}^{x_2} f(x,\ y(x),\ y'(x))\,dx+\psi[y(x_1),\ y(x_2)], \tag{9.41}$$

где $\psi[y_1,\ y_2]$ — числовая функция двух переменных, определенная на некотором открытом множестве $G\subset\mathbf{R}^2$. Терминальный функционал Ψ задан на том же множестве функций $y=y(x)$, что и интегральный функционал, и его значения вычисляют по формуле

$$\Psi(y)=\psi[y(x)|_{x=x_1},\ \ y(x)|_{x=x_2}]=\psi[y(x_1),\ y(x_2)].$$

Будем считать, что функция f из (9.41) задана и по крайней мере непрерывна на некоторой области $D\subset\mathbf{R}^3$.

Опишем множество Y допустимых функций, на котором минимизируется функционал (9.41):

$$Y=\{y\in C_1[x_1,\ x_2]\,|\,(x,\ y(x),\ y'(x))\in D\ \text{ при всех } x\in[x_1,\ x_2]\}.$$

Здесь отсутствуют краевые условия множества Y вида (9.5), фигурирующего в постановке простейшей вариационной задачи. С геометрической точки зрения это означает, что минимальное значение функционала B ищется на классе гладких кривых вида $y=y(x)$, концы которых не фиксированы и могут быть произвольно расположены на прямых $x=x_1$ и $x=x_2$ (рис. 9.9).

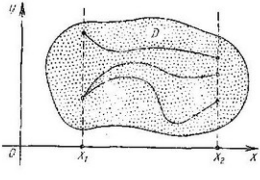

Рис. 9.9

Решением задачи Больца является допустимая функция, реализующая на указанном множестве Y наименьшее возможное значение функционала B. Введем понятие локального решения. Допустимая функция $y_0\in C_1[x_1,\ x_2]$ реализует *слабый локальный минимум* в задаче Больца, если существует такое число $\varepsilon>0$, что неравенство $B(y_0)\leqslant B(y)$ справедливо для всех функций $y\in Y\cap U^1_\varepsilon(y_0)$.

Теорема 9.4. *Предположим, что функция f определена и непрерывна вместе с частными производными f_y и $f_{y'}$ на множестве D, а* *функция ψ непрерывно дифференцируема на множестве G. Для того чтобы допустимая функция $y \in C_1[x_1, x_2]$ реализовала слабый локальный минимум в простейшей задаче Больца, необходимо, чтобы она удовлетворяла уравнению Эйлера — Лагранжа*

$$f_y(x, y, y') - \frac{d}{dx} f_{y'}(x, y, y') = 0 \tag{9.42}$$

и, кроме того, чтобы для нее были выполнены следующие условия трансверсальности:

$$f_{y'}(x_1, y(x_1), y'(x_1)) = \psi_{y_1}[y(x_1), y(x_2)], \tag{9.43}$$

$$f_{y'}(x_2, y(x_2), y'(x_2)) = -\psi_{y_2}[y(x_1), y(x_2)]. \tag{9.44}$$

□ Доказательство теоремы аналогично выводу уравнения Эйлера — Лагранжа (см. § 9.2.)

Пусть функция $y=y(x)$ реализует слабый локальный минимум в простейшей задаче Больца. Введем семейство функций вида $y+\gamma h$, где $h \in C_1[x_1, x_2]$, зависящее от параметра $\gamma \in \mathbf{R}$. При значениях γ, достаточно близких к нулю, функции данного семейства принадлежат окрестности первого порядка $U^1_\varepsilon(y)$ функции y из определения слабого локального минимума. Значит, числовая функция вида

$$\varphi(\gamma) = B(y+\gamma h) = \int_{x_1}^{x_2} f(x, y(x)+\gamma h(x), y'(x)+\gamma h'(x))\,dx +$$

$$+\psi[y(x_1+\gamma h(x_1), y(x_2)+\gamma h(x_2)]$$

достигает локального минимума при $\gamma=0$. Отсюда, используя теорему Ферма, после дифференцирования под знаком интеграла получаем

$$\varphi'(0) = \int_{x_1}^{x_2} \Big[f_y(x, y(x), y'(x))\,h(x) + f_{y'}(x, y(x), y'(x))h'(x) \Big]\,dx +$$

$$+\psi_{y_1}[y(x_1), y(x_2)]\,h(x_1) + \psi_{y_2}[y(x_1), y(x_2)]\,h(x_2) = 0 \tag{9.45}$$

для всех функций $h \in C_1[x_1, x_2]$. В частности, для всех функций h, подчиненных условиям $h(x_1)=h(x_2)=0$, из равенств (9.45) следует, что

$$\int_{x_1}^{x_2} [f_y(x, y(x), y'(x))\,h(x) + f_{y'}(x, y(x), y'(x))\,h'(x)]\,dx = 0.$$

Интегрируя первое слагаемое по частям и применяя лемму Дюбуа — Реймона, получаем, что функция y удовлетворяет уравнению Эйлера — Лагранжа (9.42).

Теперь выведем условия трансверсальности. По формуле интегрирования по частям * имеем

$$\int_{x_1}^{x_2} f_{y'}(x,\ y(x),\ y'(x))\,h'(x)\,\mathrm{d}x = \int_{x_1}^{x_2} f_{y'}(x,\ y(x),\ y'(x))\,\mathrm{d}h(x) =$$

$$= h(x)\,f_{y'}(x,\ y(x),\ y'(x))\Big|_{x_1}^{x_2} - \int_{x_1}^{x_2} h(x)\left[\frac{\mathrm{d}}{\mathrm{d}x} f_{y'}(x, y(x), y'(x))\right]\mathrm{d}x.$$

В соответствии с этим из равенства (9.45), используя уравнение Эйлера — Лагранжа (9.42), найдем

$$\varphi'(0) = \int_{x_1}^{x_2}\left[f_y(x,\ y(x),\ y'(x)) - \frac{\mathrm{d}}{\mathrm{d}x} f_{y'}(x,\ y(x),\ y'(x))\right] h(x)\,\mathrm{d}x +$$

$$+ h(x)\,f_{y'}(x,\ y(x),\ y'(x))\Big|_{x_1}^{x_2} + \psi_{y_1}[y(x_1),\ y(x_2)]\,h(x_1) +$$

$$+ \psi_{y_2}[y(x_1),\ y(x_2)]\,h(x_2) = \{f_{y'}(x_2,\ y(x_2),\ y'(x_2)) +$$

$$+ \psi_{y_2}[y(x_1),\ y(x_2)]\}\,h(x_2) + \{-f_{y'}(x_1,\ y(x_1),\ y'(x_1)) +$$

$$+ \psi_{y_1}[y(x_1),\ y(x_2)]\}\,h(x_1) = 0.$$

Подставляя последовательно в последнее равенство функции вида $h(x) = x - x_1$ и $h(x) = x - x_2$, придем к условиям трансверсальности (9.44) и (9.43) соответственно. ■

Сравнивая теоремы 9.4 и 9.1, видим, что необходимые условия для задачи Больца отличаются от соответствующих условий для простейшей вариационной задачи наличием условий трансверсальности вместо краевых условий. В записи условий трансверсальности фигурируют как значения функции в концевых точках, так и значения производных. Поэтому график оптимальной функции вблизи концевых точек (т. е. вблизи точек пересечения с прямыми $x = x_1$ и $x = x_2$) не может быть произвольным, а подчиняется определенным требованиям. Наличие условий трансверсальности в задаче Больца связано с возможностью перемещения концов допустимых линий вдоль прямых $x = x_1$, $x = x_2$.

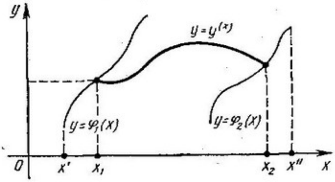

Рис. 9.10

2. Простейшая задача с подвижными концами. Пусть в плоскости xOy с помощью уравнений $y = \varphi_1(x)$ и $y = \varphi_2(x)$, $x' \leqslant x \leqslant x''$, за-

* Использование этой формулы правомерно, так как функция f_y непрерывна, а значит, согласно данному равенству $\dfrac{d}{dx} f_{y'} = f_y$, функция $f_{y'}$ непрерывно дифференцируема.

даны две кривые. Рассмотрим линии, определяемые уравнением $y=y(x)$, концы которых расположены на данных кривых (рис. 9.10). На таком классе допустимых кривых (функций) $y=y(x)$ можно рассматривать задачу минимизации некоторого функционала — *задачу с подвижными концами.*

В простейшей задаче с подвижными концами минимизируемый функционал определяется формулой

$$I(y,\ x_1,\ x_2)=\int\limits_{x_1}^{x_2} f(x,\ y(x),\ y'(x))\,\mathrm{d}x, \tag{9.46}$$

где $x'<x_1<x_2<x''$, причем x', x'' — фиксированная для данной задачи пара чисел. Будем считать, что для функции f выполнены те же предположения, что и в постановке обычной простейшей вариационной задачи.

Допустимыми здесь являются функции $y=y(x)$, удовлетворяющие следующим условиям:

$$y\in C_1[x_1,\ x_2],\quad y(x_i)=\varphi_i(x_i),\ i=1,2, \tag{9.47}$$

$$(x,\ y(x),\ y'(x))\in\widehat{D}\ \text{при всех}\ x\in[x',\ x''],$$

где φ_1, φ_2 — заданные гладкие на отрезке $[x', x'']$ функции.

Решением задачи с подвижными концами (реализующим глобальный минимум) является функция $\bar{y}$ и пара чисел $\bar{x}_1$, $\bar{x}_2$, для которых функционал (9.46) принимает наименьшее возможное значение. Тройка $\bar{y}$, $\bar{x}_1$, $\bar{x}_2$ реализует *слабый локальный минимум* в простейшей задаче с подвижными концами, если можно указать такое число $\varepsilon>0$, что для всех функций y вида (9.47) и всех пар чисел x_1, $x_2\in(x', x'')$, $x_1<x_2$, удовлетворяющих неравенствам $\|y-\bar{y}\|_1<\varepsilon$, $|x_1-\bar{x}_1|<\varepsilon$, $|x_2-\bar{x}_2|<\varepsilon$, верно неравенство $I(\bar{y}, \bar{x}_1, \bar{x}_2)\leqslant I(y, x_1, x_2)$. В этом определении числа x_1, x_2 также выбирают «локально», т. е. близкими к $\bar{x}_1$ и $\bar{x}_2$.

Теорема 9.5. *Будем считать, что функция f непрерывно дифференцируема на области $\widehat{D}$, а функции φ_1 и φ_2 непрерывно дифференцируемы на отрезке $[x', x'']$. Для того чтобы допустимая тройка $\bar{y}$, $\bar{x}_1$, $\bar{x}_2$ реализовала слабый локальный минимум в простейшей задаче с подвижными концами, необходимо, чтобы функция $\bar{y}=\bar{y}(x)$ удовлетворяла уравнению Эйлера — Лагранжа* (9.42) *при $x\in[\bar{x}_1\ \bar{x}_2]$ и, кроме того, были выполнены следующие условия трансверсальности:*

$$f(\bar{x}_i,\ \bar{y}(\bar{x}_i),\ \bar{y}'(\bar{x}_i))=f_{y'}(\bar{x}_i,\ \bar{y}(\bar{x}_i),\ \bar{y}'(\bar{x}_i))(\bar{y}'(\bar{x}_i)-\varphi_i'(\bar{x}_i)), \tag{9.48}$$

$$i=1,2.$$

□ Пусть тройка $\bar{y}$, $\bar{x}_1$, $\bar{x}_2$ реализует слабый локальный минимум. Тогда функция $\bar{y}$ реализует слабый локальный минимум в обычной простейшей вариационной задаче с фиксированным отрезком $[\bar{x}_1, \bar{x}_2]$ и краевыми условиями $y(\bar{x}_1)=\bar{y}(\bar{x}_1)$, $y(\bar{x}_2)=\bar{y}(\bar{x}_2)$. Поэтому, согласно теореме 9.1, функция $\bar{y}$ удовлетворяет уравнению Эйлера — Лагранжа (9.42) при $x \in [x_1, x_2]$.

Докажем условие трансверсальности (9.48) при $i=1$ (т. е. на левом конце). При $i=2$ доказательство аналогично. Введем семейство числовых функций переменной x вида

$$y(x;\ \gamma) = \bar{y}(x) + \gamma(x - \bar{x}_2),$$

зависящее от параметра $\gamma \in \mathbf{R}$, и функцию двух переменных x_1 и γ

$$F(x_1,\ \gamma) = y(x_1;\ \gamma) - \varphi_1(x_1).$$

Для этой функции выполнено равенство $F(\bar{x}_1, 0) = \bar{y}(\bar{x}_1) - \varphi_1(\bar{x}_1) = 0$, а для производной имеет место неравенство $F_\gamma(\bar{x}_1,\ 0) = \bar{x}_1 - \bar{x}_2 \neq 0$. В этом случае применима теорема о неявной функции [16], согласно которой существуют такие окрестности $U(\bar{x}_1)$, $U(0)$ и непрерывно дифференцируемая на $U(\bar{x}_1)$ функция $\gamma = \Gamma(x_1)$, что $\Gamma(\bar{x}_1) = 0$, $\Gamma(x_1) \in U(0)$ и, кроме того,

$$y(x_1;\ \Gamma(x_1)) = \varphi_1(x_1) \text{ для всех } x_1 \in U(\bar{x}_1), \tag{9.49}$$

$$\Gamma'(\bar{x}_1) = -\frac{F_{x_1}(\bar{x}_1,\ 0)}{F_\gamma(\bar{x}_1,\ 0)} = \frac{\varphi_1'(\bar{x}_1) - \bar{y}'(\bar{x}_1)}{\bar{x}_1 - \bar{x}_2}. \tag{9.50}$$

Рассмотрим семейство функций $y = y(x;\ \Gamma(x_1))$ переменной x при $\Gamma(x_1) \in U(0)$. Согласно равенству (9.49), левый конец графиков этих функций находится на кривой $y = \varphi_1(x)$ (в окрестности $U(\bar{x}_1)$). Правые их концы совпадают с точкой $(\bar{x}_2,\ \varphi_2(\bar{x}_2))$. Тройка $\bar{y}$, $\bar{x}_1$, $\bar{x}_2$ реализует слабый локальный минимум, поэтому на более узком множестве допустимых функций, а именно среди всех функций рассматриваемого семейства*, наименьшее значение функционала (9.46) также реализует функция $y = \bar{y}(x) = y(x;\ 0)$. Иначе говоря, числовая функция

$$v(x_1) = \int_{x_1}^{\bar{x}_2} f(x,\ \bar{y}(x) + \Gamma(x_1)(x - \bar{x}_2),\ \bar{y}'(x) + \Gamma(x_1))\,\mathrm{d}x$$

достигает локального минимума в точке $x_1 = \bar{x}_1$. Отсюда, согласно теореме Ферма, следует, что $v'(\bar{x}_1) = 0$ или

$$v'(\bar{x}_1) = -f(\bar{x}_1,\ \bar{y}(\bar{x}_1),\ \bar{y}'(\bar{x}_1)) +$$
$$+ \Gamma'(\bar{x}_1) \int_{\bar{x}_1}^{\bar{x}_2} [f_y(x,\ \bar{y}(x),\ \bar{y}'(x),\ (x - \bar{x}_2) +$$
$$+ f_{y'}(x,\ \bar{y}(x),\ \bar{y}'(x)]\,\mathrm{d}x = 0. \tag{9.51}$$

* Считаем, что рассматриваемое семейство функций входит в окрестность первого порядка функции $\bar{y}$ из определения слабого локального минимума.

Используя формулу интегрирования по частям, имеем

$$\int\limits_{\bar{x}_1}^{\bar{x}_2} f_{y'}(x,\ \bar{y}(x),\ \bar{y}'(x))\,\mathrm{d}x = (x-\bar{x}_2) f_{y'}(x,\ \bar{y}(x),\ \bar{y}'(x))\Big|_{\bar{x}}^{\bar{x}_2} -$$

$$-\int\limits_{\bar{x}_1}^{\bar{x}_2} (x-\bar{x}_2)\left[\frac{\mathrm{d}}{\mathrm{d}x} f_{y'}(x,\ \bar{y}(x),\ \bar{y}'(x))\right]\mathrm{d}x.$$

Учитывая последнее равенство, из (9.51) получаем

$$-f(\bar{x}_1,\ \bar{y}(\bar{x}_1),\ \bar{y}'(\bar{x}_1)) + \Gamma'(\bar{x}_1)\times$$

$$\times\left\{\int\limits_{\bar{x}_1}^{\bar{x}_2}\left[f_y(x,\ \bar{y}(x),\ \bar{y}'(x)) - \frac{\mathrm{d}}{\mathrm{d}x} f_{y'}(x,\ \bar{y}(x),\ \bar{y}'(x))\right](x-\bar{x}_2)\,\mathrm{d}x +\right.$$

$$\left.+(x-\bar{x}_2) f_{y'}(x,\ \bar{y}(x),\ \bar{y}'(x))\Big|_{\bar{x}_1}^{\bar{x}_2}\right\} = 0.$$

Отсюда, принимая во внимание, что функция $\bar{y}$ удовлетворяет уравнению Эйлера — Лагранжа, и используя равенство (9.50), приходим к равенству (9.48) при $i=1$. ■

Если подвижным является только один конец, например правый, то для функции, реализующей слабый локальный минимум, условие трансверсальности должно быть выполнено только на правом конце (при этом левый конец закреплен: $y(x_1)=y_1$).

Для иллюстрации условий трансверсальности рассмотрим задачу о брахистохроне, в которой, по условию, правый конец подвижен и может перемещаться вдоль гладкой кривой $y=\varphi_2(x)$, причем $\varphi_2(x)>0$ для всех $x\in[0,\ x'']$. Запишем для этой задачи условия трансверсальности при $i=2$:

$$\frac{\sqrt{1+y'^2}}{\sqrt{2gy}} = \frac{y'}{\sqrt{2gy}\,\sqrt{1+y'^2}}(y'-\varphi_2'),$$

откуда

$$y'(\bar{x}_2)\,\varphi_2'(\bar{x}_2) = -1.$$

Геометрически последнее равенство означает, что касательные к брахистохроне и кривой $y=\varphi_2(x)$, проведенные в точке их пересечения, взаимно перпендикулярны (рис. 9.11).

3. **Простейшая изопериметрическая задача** *. До сих пор мы рассматривали вариационные задачи без ограничений, в которых

* Термин *изопериметрическая задача* появился в связи с задачей отыскания замкнутой кривой фиксированной длины (периметра), ограничивающей плоскую область наибольшей возможной площади.

фигурировал безусловный (локальный или глобальный) экстремум. В этом пункте будут сформулированы необходимые условия условного (слабого локального) экстремума в простейшей вариационной задаче с единственным функциональным ограничением в форме равенства. Вариационные задачи с функциональными ограничениями в форме равенств называют *изопериметрическими задачами.*

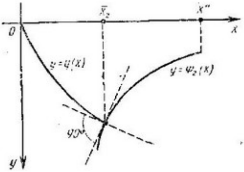

Рис. 9.11

В постановке простейшей изопериметрической задачи имеются два функционала:

$$I(y)=\int_{x_1}^{x_2} f(x,\ y(x),\ y'(x))\,\mathrm{d}x, \tag{9.52}$$

$$J(y)=\int_{x_1}^{x_2} g(x,\ y(x),\ y'(x))\,\mathrm{d}x. \tag{9.53}$$

Первый из них — целевой функционал — подлежит минимизации, функционал (9.53) используется в записи функционального ограничения. Для функционалов I и J считаем выполненными предположения из постановки простейшей вариационной задачи. В частности, функции f и g предполагаются заданными (и, по крайней мере, непрерывными) на связной области $\hat{D}\subset\mathbf{R}^3$.

Множество допустимых функций, на котором минимизируется функционал (9.52), имеет вид

$$Y=\{y\in C_1[x_1,\ x_2]\,|\,J(y)=l,\quad y(x_i)=y_i,\quad i=1,2,$$
$$(x,\ y(x),\ y'(x))\in\widehat{D} \text{ при всех } x\in[x_1,\ x_2]\}, \tag{9.54}$$

где l, x_i, y_i, $i=1, 2$ — заданные числа.

Функция, минимизирующая функционал I на множестве (9.54), реализует глобальный (условный) минимум этого функционала. Понятие локального экстремума здесь определяется обычным образом. Говорят, что функция $\bar{y}\in Y$ реализует *слабый локальный* (условный) *минимум* в простейшей изопериметрической задаче, если найдется такое число $\varepsilon>0$, что неравенство $I(\bar{y})\leqslant I(y)$ имеет место для всех $y\in Y\cap U_\varepsilon^1(\bar{y})$.

Теорема 9.6. *Допустим, что функции f и g непрерывно дифференцируемы на множестве $\hat{D}$. Для того чтобы функция $y\in Y$ реализовала слабый локальный минимум в простейшей изопериметрической задаче, необходимо существование таких не равных одновременно нулю чисел λ_0, λ_1, при которых функция $\bar{y}$ удовлетворяет уравнению Эйлера — Лагранжа вида*

$$L_y-\frac{\mathrm{d}}{\mathrm{d}x}L_{y'}=0, \tag{9.55}$$

где $L=\lambda_0 f(x,\ y(x),\ y'(x)),\ +\lambda_1 g(x,\ y(x),\ y'(x))$.

□ Пусть функция $\bar{y}$, реализующая слабый локальный минимум, является экстремалью функционала J. Тогда она удовлетворяет уравнению Эйлера — Лагранжа вида $g_y - \frac{d}{dx} g_{y'} = 0$, а значит, и уравнению (9.55) при $\lambda_0 = 0$, $\lambda_1 = 1$. Таким образом, можно считать, что функция $\bar{y}$ не является экстремалью функционала J.

Рассмотрим две произвольные функции h_1 и h_2 вида

$$h_i \in C_1[x_1, x_2], \quad h_i(x_1) = h_i(x_2) = 0, \quad i = 1, 2, \tag{9.56}$$

и семейство функций

$$y(x; \gamma_1, \gamma_2) = \bar{y}(x) + \gamma_1 h_1(x) + \gamma_2 h_2(x),$$

зависящее от параметров $\gamma_1, \gamma_2 \in \mathbf{R}$. Если эти параметры взять достаточно близкими к нулю, то функция $y = y(x; \gamma_1, \gamma_2)$ (при фиксированных γ_1, γ_2) будет принадлежать окрестности первого порядка радиуса $\varepsilon > 0$ из определения слабого локального минимума для $\bar{y}$.

Введем функцию двух переменных F с помощью равенства

$$F(\gamma_1, \gamma_2) = J(y(x; \gamma_1, \gamma_2)) - l.$$

Согласно предположениям теоремы, эта функция непрерывно дифференцируема в некоторой окрестности начала координат $(0, 0)$ и для нее выполняется равенство $F(0, 0) = J(\bar{y}) - l = 0$. Дифференцируя функцию F с помощью правила Лейбница и учитывая равенства $h_i(x_1) = h_i(x_2) = 0$, $i = 1, 2$, получаем $F_{\gamma_2}(0, 0) = \delta J(h_2)$, где в правой части равенства находится функционал первой вариации для J (относительно $\bar{y}$). Функция $\bar{y}$ не является экстремалью функционала J, поэтому функцию h_2 можно выбрать так, чтобы имело место неравенство $F_{\gamma_2}(0, 0) = \delta J(h_2) \neq 0$. Функцию h_2, удовлетворяющую этому неравенству, в дальнейшем будем считать зафиксированной.

Согласно теореме о неявной функции [16], существует окрестность $U(0)$ и непрерывно дифференцируемая (на $U(0)$) функция $\gamma_2 = \Gamma(\gamma_1)$ такие, что $\Gamma(0) = 0$,

$$F(\gamma_1, \Gamma(\gamma_1)) = J(y(x; \gamma_1, \Gamma(\gamma_1))) - l = 0 \text{ для всех } \gamma_1 \in U(0), \tag{9.57}$$

$$\Gamma'(0) = -\frac{F_{\gamma_1}(0, 0)}{F_{\gamma_2}(0, 0)} = -\frac{\delta J(h_1)}{\delta J(h_2)}. \tag{9.58}$$

Рассмотрим семейство функций вида $y = y(x; \gamma_1, \Gamma(\gamma_1))$ при достаточно близких к нулю значениях γ_1. Согласно равенству (9.57), все эти функции принадлежат множеству (9.54) (равенства $y(x_i; \gamma_1, \Gamma(\gamma_1)) = y_i$ можно проверить непосредственно, учитывая равенства (9.56)). Более того, верно равенство $y(x; 0, 0) = \bar{y}(x)$ и поэтому при всех γ_1, достаточно близких к нулю, эти функции принадлежат окрестности первого порядка из определения слабого локального минимума для $\bar{y}$. Следовательно, числовая функция

$$v(\gamma_1)=\int\limits_{x_1}^{x_2} f(x,\ \tilde{y}(x),\ \tilde{y}'(x))\,\mathrm{d}x,$$

где $\tilde{y}(x)=\bar{y}(x)+\gamma_1 h_1(x)+\Gamma(\gamma_1)h_2(x)$, в точке $\gamma_1=0$ достигает локального минимума, а значит, по теореме Ферма, имеет место равенство $v'(0)=0$. Дифференцируя функцию v под знаком интеграла, получаем

$$v'(0)=\delta I(h_1)+\Gamma'(0)\,\delta I(h_2)=\delta I(h_1)+\lambda_1\delta J(h_1)=0, \tag{9.59}$$

где

$$\lambda_1=\frac{\Gamma'(0)\,\delta I(h_2)}{\delta J(h_1)}=-\frac{\delta I(h_2)}{\delta J(h_2)}. \tag{9.60}$$

Заметим, что функция h_1 не была зафиксирована, поэтому равенство (9.59) имеет место для любой функции h_1 вида (9.56).

Введем функционал

$$I^*(y)=\int\limits_{x_1}^{x_2}[f(x,\ y(x),\ y'(x))+\lambda_1 g(x,\ y(x),\ y'(x))]\,\mathrm{d}x,$$

определенный на множестве $\hat{D}$. Очевидно,

$$\delta I^*(h_1)=\delta I(h_1)+\lambda_1\delta J(h_1),$$

где функция h_1 имеет вид (9.56). Поэтому равенство (9.59) равносильно тождественному равенству нулю первой вариации функционала I^*: $\delta I^*(h_1)\equiv 0$. Таким образом, функция $\bar{y}$ является решением уравнения Эйлера — Лагранжа (9.55) при $\lambda_0=1$ и λ_1 вида (9.60). ■

Уравнение (9.55) из теоремы 9.6 можно умножать на любое число, отличное от нуля. Поэтому при его решении достаточно ограничиться рассмотрением следующих двух случаев: 1) $\lambda_0=0$, $\lambda_1=1$ и 2) $\lambda_0=1$, $\lambda_1\in\mathbf{R}$.

Числа λ_0 и λ_1 часто называют *множителями Лагранжа*, а метод решения изопериметрической задачи на основании теоремы 9.6 — *методом Лагранжа*. Существует аналогия между способом учета ограничений в нелинейном программировании и вариационном исчислении: в обоих случаях фигурирует функция (функционал) Лагранжа. Понятие функционала Лагранжа неявно было использовано при доказательстве теоремы 9.6 (см. функционал I^*).

Утверждение теоремы 9.6 можно распространить на более широкий класс изопериметрических задач, в которых имеется несколько функциональных ограничений-равенств $J_j(y)=l_j$, $j=1,\ 2,\ \ldots,\ k$ [15]. В этом случае функция L из (9.55) имеет вид

$$L=\lambda_0 f(x,\ y(x),\ y'(x))+\sum_{j=1}^{k}\lambda_j g_j(x,\ y(x),\ y'(x)),$$

где g_j — подынтегральная функция для функционала J_j, $j=1,\ 2,\ \ldots,\ k$.

Пример (задача Дидоны *). Рассмотрим следующую вариационную задачу с неподвижными концами: соединить кривой $y=y(x)\geqslant 0$, $-x_0\leqslant x\leqslant x_0$, имеющей длину l, две точки $(-x_0, 0)$ и $(x_0, 0)$ таким образом, чтобы плоская область, которая ограничена осью абсцисс и этой кривой, имела наибольшую возможную площадь. Математическая формулировка данной задачи приводит к минимизации функционала

$$I(y)=-\int_{-x_0}^{x_0} y(x)\,\mathrm{d}x$$

при ограничении

$$J(y)=\int_{-x_0}^{x_0} \sqrt{1+y'^2(x)}\,\mathrm{d}x=l.$$

При этом $l\geqslant 2x_0$ (в противном случае задача заведомо не имеет решения) и выполняются краевые условия $y(-x_0)=y(x_0)=0$. Запишем для данной задачи уравнение Эйлера — Лагранжа:

$$-\lambda_0-\frac{\mathrm{d}}{\mathrm{d}x}\,\frac{\lambda_1 y'}{\sqrt{1+y'^2}}=0. \tag{9.61}$$

Если $\lambda_0=0$, то отсюда следует, что $y'=\mathrm{const}$, а значит, что $y\equiv 0$ — единственное решение уравнения (9.61), удовлетворяющее краевым условиям. Это возможно только при $l=2x_0$.

Пусть $\lambda_0=1$, $\lambda_1=\lambda\in\mathbf{R}$. В этом случае, интегрируя уравнение (9.61), найдем

$$\frac{\lambda y'}{\sqrt{1+y'^2}}=x-C_1,$$

где $C_1=\mathrm{const}$. Решая полученное уравнение относительно y', получаем

$$y'=\pm\frac{x-C_1}{\sqrt{\lambda^2-(x-C_1)^2}},$$

откуда после интегрирования имеем

$$y-C_2=\pm\sqrt{\lambda^2-(x-C_1)^2},$$

или

$$(x-C_1)^2+(y-C_2)^2=\lambda^2.$$

В результате получено уравнение с е м е й с т в а о к р у ж н о с т е й. Используя краевые условия и ограничение-равенство (запи-

---

* Исторический комментарий к этой задаче см. в [2].

санное для функций найденного семейства), запишем систему уравнений:

$$(-x_0-C_1)^2+C_2^2=\lambda^2,$$

$$(x_0-C_1)^2+C_2^2=\lambda^2,$$

$$\lambda\int\limits_{-x_0}^{x_0}\frac{dx}{\sqrt{\lambda^2-(x-C_1)^2}}=l,$$

откуда находим

$$C_1=0,\ C_2^2=\lambda^2-x_0^2,\ \arcsin\frac{x_0}{\lambda}=\frac{l}{2\lambda}. \tag{9.62}$$

Последнее уравнение относительно λ имеет (и притом единственное с точностью до знака) решение только при $l\leqslant\pi x_0$.

Таким образом, при $2x_0<l\leqslant\pi x_0$ существует единственное решение уравнения Эйлера — Лагранжа. Это дуга окружности радиуса $|\lambda|$ с центром в точке $(0, -|C_2|)$, где C_2 и λ определяются из уравнений (9.62). Из геометрических соображений следует, что решение рассматриваемой задачи существует, поэтому можно сделать вывод, что найденные дуги являются решениями задачи Дидоны. При $l>\pi x_0$ уравнение Эйлера — Лагранжа гладкого решения не имеет.

§ 9.6. БОЛЕЕ ОБЩИЕ ВАРИАЦИОННЫЕ ЗАДАЧИ

1. Классификация задач вариационного исчисления. Задачи вариационного исчисления классифицируются на основе выделения различных классов и типов целевых функционалов, граничных условий и ограничений, которые фигурируют в постановке любой вариационной задачи.

По типу целевого функционала вариационные задачи можно подразделить на:

задачи с интегральным функционалном (задачи Лагранжа)

$$I(y)=\int\limits_{x_1}^{x_2}f(x,\ y(x),\ y'(x))\,dx; \tag{9.63}$$

задачи с терминальным функционалом (задачи Майера)

$$\Psi(y)=\psi[y(x_1),\ y(x_2)]; \tag{9.64}$$

задачи со смешанным функционалом (задачи Больца)

$$B(y)=I(y)+\Psi(y). \tag{9.65}$$

Во всех рассмотренных ранее вариационных задачах допустимые функции $y=y(x)$ считались числовыми. В задачах (9.63) — (9.65) функция $y=y(x)$ может быть векторной: $y(x)=(y_1(x),$

$y_2(x), \dots, y_n(x))$, $n \geqslant 1$. При $n > 1$ говорят о *пространственной (векторной) задаче*, а при $n = 1$ — о *плоской задаче*. Заметим, что при $n > 1$ в формуле (9.63) f — функция $2n+1$ переменной, а $y'(x)$ означает набор функций $y_1'(x), y_2'(x), \dots, y_n'(x)$.

Граничные (краевые) условия в вариационных задачах в общем случае $n \geqslant 1$ можно задать указанием множества из пространства $\mathbf{R}^{2n+2}$, которому должны принадлежать концы графиков допустимых функций $y = y(x)$, т. е. «обобщенная» точка $(x_1, y(x_1), x_2, y(x_2))$. Если это множество состоит из единственного элемента (x_1, y_1, x_2, y_2), то соответствующую вариационную задачу называют *задачей с неподвижными (закрепленными) концами* $y(x_1) = y_1$, $y(x_2) = y_2$; причем отрезок $[x_1, x_2]$ фиксирован. В *задачах с подвижными концами* отрезок $[x_1, x_2]$ (только его левый или правый конец) может быть как фиксированным, так и нефиксированным; при этом подмножество, которому должна принадлежать пара точек $(y(x_1), y(x_2))$, обычно задают в виде множества решений уравнения

$$\Phi(y(x_1), y(x_2)) = 0_s,$$

где Φ — s-мерная вектор-функция $2n$ переменных. Например, в задаче с подвижными концами, рассмотренной в п. 2 § 9.5, Φ — двумерная вектор-функция двух переменных вида $\Phi(y(x_1), y(x_2)) = (y(x_1) - \varphi_1(x_1), y(x_2) - \varphi_2(x_2))$.

Ограничения в вариационных задачах, как правило, имеют вид *функциональных равенств* *

$$J_j(y) = l_j, \quad j = 1, 2, \dots, k, \tag{9.66}$$

где J_j и l_j, $j = 1, 2, \dots, k$, — заданные функционалы и числа, или *дифференциальных равенств*

$$D_j(x, y(x), y'(x)) = 0, \quad j = 1, 2, \dots, p; \quad x \in [x_1, x_2]. \tag{9.67}$$

где $D_1, D_2, \dots, D_p$ — заданные функции $2n+1$ переменной. В частности, равенства (9.67) могут иметь вид

$$y_j' = \varphi_j(x, y), \quad j = 1, \dots, n,$$

где φ_j — функция $n+1$ переменной. Вариационные задачи с ограничениями вида (9.66) называют *изопериметрическими задачами*. В общем случае в вариационной задаче могут присутствовать как функциональные, так и дифференциальные ограничения. В последнее время объектом изучения вариационного исчисления стали также задачи, содержащие кроме указанных ограничений еще и функциональные неравенства.

Все перечисленные вариационные задачи относятся к *задачам первого порядка* в том смысле, что в их формулировках фигурирует только производная первого порядка. В целевые функционалы и ограничения *задач высшего порядка* могут входить производные $y', y'', \dots, y^{(m)}$ до некоторого порядка m включительно.

* Равенства (9.66) по аналогии с задачей нелинейного программирования можно записать в виде $\bar{J}(y) = 0$, где $\bar{J}_j(y) = J_j(y) - l_j$.

В задачах первого порядка допустимые функции $y=y(x)$, как правило, считают гладкими (непрерывно дифференцируемыми). В некоторых случаях удобно рассматривать более широкие классы допустимых функций (например, кусочно-гладкие функции). В задачах высшего порядка требования к допустимым функциям следует усилить, чтобы постановка таких задач была корректной.

2. Пространственная вариационная задача. Рассмотрим пространственную задачу Лагранжа с неподвижными концами (без ограничений) и сформулируем необходимое условие оптимальности. Целевой функционал, подлежащий минимизации, здесь имеет вид (9.63), где

$$y(x)=(y_1(x),\ y_2(x),\dots,\ y_n(x));\ y'(x)=(y_1'(x),\ y_2'(x),\dots,\ y_n'(x)),$$

а f — функция $2n+1$ переменной, определенная на некоторой области $\hat{D}\subset\mathbf{R}^{2n+1}$. Положим

$$D=\{(x,\ y)\in\mathbf{R}^{n+1}\,|\,(x,\ y,\ y')\in\widehat{D}\ \text{при некотором}\ y\in\mathbf{R}^n\}$$

и зафиксируем пару точек $A(x_1,\ y_1)$, $B(x_2,\ y_2)$ множества D, где $y_k=(y_{k1},\ y_{k2},\ \dots,\ y_{kn})$, $k=1,\ 2$, причем $x_1<x_2$.

Линейное пространство n-мерных вектор-функций, каждая компонента которых принадлежит $C_1[x_1,\ x_2]$, будем обозначать через $C_1^n[x_1,\ x_2]$ *. Используя это пространство, множество допустимых вектор-функций в пространственной задаче определим с помощью равенства

$$Y=\{y\in C_1^n[x_1,\ x_2]\,|\,y(x_k)=y_k,\ k=1,\ 2,$$

$$(x,\ y(x),\ y'(x))\in\widehat{D}\ \text{при всех}\ x\in[x_1,\ x_2]\}.$$

Решением пространственной задачи (реализующим глобальный безусловный минимум целевого функционала) является вектор-функция $\bar{y}\in Y$, для которой неравенство $I(\bar{y})\leqslant I(y)$ имеет место при всех $y\in Y$.

Для того чтобы ввести понятие локального решения, пространство $C_1^n[x_1,\ x_2]$ необходимо превратить в нормированное пространство. Норму в $C_1^n[x_1,\ x_2]$ можно ввести согласно любой из следующих формул:

$$\|y\|_0=\max_{x\in[x_1,\ x_2]}\|y(x)\|,$$

$$\|y\|_1=\max_{x\in[x_1,\ x_2]}\|y(x)\|+\max_{x\in[x_1,\ x_2]}\|y'(x)\|,$$

где через $\|\dots\|$ обозначена обычная евклидова норма в пространстве $\mathbf{R}^n$:

$$\|z\|=\sqrt{z_1^2+z_2^2+\dots+z_n^2},\quad z=(z_1,\ z_2,\dots,\ z_n).$$

* В пространстве $C_1^n[x_1,\ x_2]$ сумма элементов и произведение элемента на число определяются в виде покомпонентной суммы и покомпонентного произведения на число.

Определим с помощью введённых норм *ε-окрестности*

$$\mathbf{U}^0_\varepsilon(\bar{y}) = \{y \in C_1^n[x_1, x_2] \mid \|y - \bar{y}\|_0 < \varepsilon\},$$

$$\mathbf{U}^1(\bar{y}) = \{y C_1^n \in [x_1, x_2] \mid \| \bar{y} - y \|_1 < \varepsilon\}$$

нулевого и первого порядка соответственно.

Говорят, что функционал I вида (9.63) достигает *слабого (сильного) локального минимума* на вектор-функции $\bar{y} \in Y$, если найдется такое число $\varepsilon > 0$, что неравенство $I(\bar{y}) \leqslant I(y)$ выполняется при всех $y \in Y \cap \mathbf{U}_\varepsilon^1(\bar{y})$ (соответственно $y \in Y \cap \mathbf{U}_\varepsilon^0(\bar{y})$).

Необходимое условие слабого локального (сильного и глобального) минимума для рассматриваемой пространственной задачи сформулировано в следующем утверждении.

Теорема 9.7. *Предположим, что функция f из* (9.63) *непрерывна вместе со своими частными производными f_{y_i}, $f_{y'_i}$, $i = 1, 2, \ldots, n$, на множестве $\hat{D}$. Для того чтобы вектор-функция $\bar{y} \in Y$ реализовала слабый локальный минимум функционала* (9.63), *необходимо, чтобы она удовлетворяла следующей системе дифференциальных уравнений Эйлера — Лагранжа:*

$$f_{y_i}(x, y, y') - \frac{\mathrm{d}}{\mathrm{d}x} f_{y'_i}(x, y, y'), \quad i = 1, 2, \ldots, n. \tag{9.68}$$

☐ Доказательство проведем сведéнием пространственного случая к плоскому.

Введем множества

$$\hat{D}_i = \{(x, y_i, y'_i) \in \mathbf{R}^3 \mid (x, \bar{y}_1, \ldots, \bar{y}_{i-1}, y_i, \bar{y}_{i+1}, \ldots, \bar{y}_n,$$

$$\bar{y}'_1, \ldots, \bar{y}'_{i-1}, y'_i, \bar{y}'_{i+1}, \ldots, y'_n) \ni \hat{D}\}, \quad i = 1, 2, \ldots, n,$$

на которых определим функционалы

$$I_i(y_i) = \int_{x_1}^{x_2} f(x, \bar{y}_1, \ldots, \bar{y}_{i-1}, y_i, \bar{y}_{i+1}, \ldots, \bar{y}_n,$$

$$\bar{y}'_1, \ldots, \bar{y}'_{i-1}, y'_i, \bar{y}'_{i+1}, \ldots, \bar{y}'_n)\,\mathrm{d}x$$

и зададим граничные условия

$$y_i(x_k) = y_{ki}, \quad i = 1, 2, \ldots, n;\ k = 1, 2.$$

Вектор-функция $\bar{y}$ реализует слабый локальный минимум функционала (9.63), поэтому функция $\bar{y}_i$ реализует слабый локальный минимум функционала I_i, $i = 1, 2, \ldots, n$. Таким образом, имеем n простейших вариационных задач, к каждой из которых применима теорема 9.1. В результате получаем, что функция $\bar{y}_i$ удовлетворяет уравнению (9.68), $i = 1, 2, \ldots, n$. ■

Для нахождения экстремалей в пространственной задаче необходимо решить систему n дифференциальных уравнений (9.68) (общее решение которой в общем случае зависит от $2n$ произвольных постоянных) при $2n$ краевых условиях $y_i(x_k) = y_{ki}$, $i=1, 2, \ldots, n$; $k=1, 2$. Нередко сделать это трудно (а подчас и невозможно), так как система (9.68) может быть очень сложной.

3. Задача высшего порядка. Получим необходимое условие оптимальности для следующей задачи Лагранжа высшего порядка с неподвижными концами (без ограничений). Требуется минимизировать функционал

$$I(y) = \int_{x_1}^{x_2} f(x, y(x), y'(x), \ldots, y^{(m)}(x))\,\mathrm{d}x \tag{9.69}$$

на следующем множестве допустимых функций:

$$Y = \{y \in C_m[x_1, x_2] \mid y^{(k)}(x_i) = y_{ik},\ k=0, 1, \ldots, m-1;\ i=1, 2; \tag{9.70}$$

$$(x, y(x), y'(x), \ldots, y^{(m)}(x)) \in \widehat{D} \text{ при всех } x \in [x_1, x_2]\}.$$

Функция $m+2$ переменных f считается заданной и по крайней мере непрерывной на некоторой области $\widehat{D} \subset \mathbf{R}^{m+2}$; $C_m[x_1, x_2]$ — линейное пространство числовых функций, имеющих на отрезке $[x_1, x_2]$ непрерывные производные до порядка m включительно; x_i, y_{ik}, $k=0, 1, \ldots, m-1$; $i=1, 2$,— заданные числа, определяющие граничные условия.

Функцию $\bar{y} \in Y$, реализующую минимум функционала (9.69) на множестве Y вида (9.70), называют решением (глобальным и безусловным) сформулированной задачи высшего порядка.

Введем по обычным правилам понятие локального решения. Превратим пространство $C_m[x_1, x_2]$ в нормированное пространство, определив на нем норму следующим образом:

$$\|y\|_m = \sum_{k=0}^{m} \max_{x \in [x_1, x_2]} |y^{(k)}(x)|.$$

В частных случаях $m=0$ и $m=1$ из $C_m[x_1, x_2]$ получаем введенные ранее нормированные пространства $C[x_1, x_2]$ (т. е. $C_0[x_1, x_2]$) и $C_1[x_1, x_2]$. Далее определяем *ε-окрестность порядка m* элемента $\bar{y}$:

$$U_\varepsilon^m(\bar{y}) = \{y \in C_m[x_1, x_2] \mid \|y - \bar{y}\|_m < \varepsilon\}.$$

Если можно указать такое число $\varepsilon > 0$, что неравенство $I(\bar{y}) \leqslant I(y)$ имеет место для всех $y \in Y \cap U_\varepsilon^m(\bar{y})$, то говорят, что функционал (9.69) достигает на элементе $\bar{y}$ *слабого локального минимума*.

При выводе необходимого условия оптимальности потребуется вспомогательное утверждение, которое в частном случае $m=1$ совпадает с леммой Дюбуа — Реймона (см. § 9.2).

Лемма 9.5 (обобщенная лемма Дюбуа — Реймона). *Пусть задана функция* $\omega \in C[x_1, x_2]$ *и равенство*

$$\int_{x_1}^{x_2} \omega(x) h^{(m)}(x) \, \mathrm{d}x = 0 \tag{9.71}$$

имеет место для всех функций h, *удовлетворяющих условиям* *

$$h \in C_m[x_1, x_2]; \; h^{(k)}(x_1) = h^{(k)}(x_2) = 0, \; k = 0, 1, \ldots, m-1. \tag{9.72}$$

Тогда существует такой многочлен p_{m-1}, *степень которого не больше чем* $m-1$, *что равенство* $\omega(x) = p_{m-1}(x)$ *верно при всех* $x \in [x_1, x_2]$.

☐ Рассмотрим многочлен вида

$$p_{m-1}(x) = C_0 + C_1(t - x_2) + \ldots + C_{m-1}(t - x_2)^{m-1}$$

и подберем константы $C_0, C_1, \ldots, C_{m-1}$ таким образом, чтобы выполнялись равенства

$$\int_{x_1}^{x_2} (t - x_2)^k (p_{m-1}(t) - \omega(t)) \, \mathrm{d}t = 0, \; k = 0, 1, \ldots, m-1. \tag{9.73}$$

Интегрируя, получаем следующую систему m линейных уравнений относительно $C_0, C_1, \ldots, C_{m-1}$:

$$\frac{(x_2 - x_1)^{k+1}}{k+1} C_0 + \frac{(x_2 - x_1)^{k+2}}{k+2} C_1 + \ldots + \frac{(x_2 - x_1)^{k+m}}{k+m} C_{m-1} =$$

$$= \int_{x_1}^{x_2} (t - x_2)^k \omega(t) \, \mathrm{d}t, \; k = 0, 1, \ldots, m-1.$$

Можно проверить, что определитель этой системы уравнений не равен нулю:

$$\begin{vmatrix} \dfrac{x_2 - x_1}{1} & \dfrac{(x_2 - x_1)^2}{2} & \ldots & \dfrac{(x_2 - x_1)^m}{m} \\ \dfrac{(x_2 - x_1)^2}{2} & \dfrac{(x_2 - x_1)^3}{3} & \ldots & \dfrac{(x_2 - x_1)^{m+1}}{m+1} \\ \cdots & \cdots & \cdots & \cdots \\ \dfrac{(x_2 - x_1)^m}{m} & \dfrac{(x_2 - x_1)^{m+1}}{m+1} & \ldots & \dfrac{(x_2 - x_1)^{2m-1}}{2m-1} \end{vmatrix} =$$

$$= (x_2 - x_1)^{m^2} \begin{vmatrix} 1 & 1/2 & \ldots & 1/m \\ 1/2 & 1/3 & \ldots & 1/(m+1) \\ \cdots & \cdots & \cdots & \cdots \\ 1/m & 1/(m+1) & \ldots & 1/(2m-1) \end{vmatrix} \neq 0,$$

поэтому, согласно теореме Крамера, она имеет (и притом единственное) решение. Пусть $C_0, C_1, \ldots, C_{m-1}$ — решение этой системы.

* В этих условиях использовано стандартное обозначение $h^{(0)} = h$.

Функция $p_{m-1}(t)-\omega(t)$ на отрезке $[x_1, x_2]$ непрерывна, поэтому для доказательства леммы достаточно убедиться в справедливости равенства

$$I_* = \int_{x_1}^{x_2} (p_{m-1}(t)-\omega(t))^2\,\mathrm{d}t = 0.$$

Распишем это равенство более подробно:

$$I_* = \int_{x_1}^{x_2} p_{m-1}(t)(p_{m-1}(t)-\omega(t))\,\mathrm{d}t - \int_{x_1}^{x_2} \omega(t)(p_{m-1}(t)-\omega(t))\,\mathrm{d}t.$$

Последовательно умножим равенства (9.73) на $C_0, C_1, \ldots, C_{m-1}$, а затем полученные равенства почленно сложим. В результате имеем

$$\int_{x_1}^{x_2} p_{m-1}(t)(p_{m-1}(t)-\omega(t))\,\mathrm{d}t = 0.$$

Следовательно,

$$I_* = -\int_{x_1}^{x_2} \omega(t)(p_{m-1}(t)-\omega(t))\,\mathrm{d}t. \tag{9.74}$$

Введем функцию $h_0 \in C_m[x_1, x_2]$ с помощью формулы

$$h_0(x) = \frac{1}{(m-1)!}\int_{x_1}^{x} (t-x)^{m-1}(p_{m-1}(t)-\omega(t))\,\mathrm{d}t.$$

Ее производные имеют вид $(k=0, 1, \ldots, m-1)$

$$h_0^{(k)}(x) = \frac{(-1)^k}{(m-k-1)!}\int_{x_1}^{x} (t-x)^{m-k-1}(p_{m-1}(t)-\omega(t))\,\mathrm{d}t.$$

Очевидно, $h_0^{(k)}(x_1)=0$. Более того, в силу (9.73) выполняются равенства $h^{(k)}(x_2)=0$, $k=0, 1, \ldots, m-1$.

Таким образом, функция h_0 удовлетворяет условиям леммы и ее можно подставить в равенство (9.71) вместо h. Поскольку $h_0^{(m)} = (-1)^{m-1}(p_{m-1}-\omega)$, выполнив эту подстановку, на основании (9.74) получаем требуемое равенство $I_*=0$. ■

Теорема 9.8. *Предположим, что функция $\mathfrak{f}$ непрерывна на множестве $\bar{D}\subset\mathbf{R}^{m+2}$ вместе с частными производными $\mathfrak{f}_y, \mathfrak{f}_{y'}, \ldots, \mathfrak{f}_{y^{(m)}}$. Для того чтобы функция $\bar{y}\in Y$ реализовала слабый локальный минимум в задаче высшего порядка, необходимо, чтобы она удовлетворяла дифференциальному уравнению*

$$f_y(x, y, y', \ldots, y^{(m)}) - \frac{\mathrm{d}}{\mathrm{d}x}\Big[f_{y'}(x, y, y', \ldots, y^{(m)}) - \ldots$$

$$\ldots - \frac{d}{dx}\left[f_{y^{(m-1)}}(x, y, y', \ldots, y^{(m)}) -\right.$$

$$\left. - \frac{d}{dx}[f_{y^{(m)}}(x, y, y', \ldots, y^{(m)})]\right]\ldots\Big] = 0, \qquad (9.75)$$

которое при дополнительном предположении

$$f_{y^{(k)}}(x, y(x), y'(x), \ldots, y^{(m)}(x)) \in C_k[x_1, x_2], \ k = 1, 2, \ldots, m-1,$$

после выполнения операции дифференцирования принимает вид дифференциального уравнения

$$f_y(x, y, y', \ldots, y^{(m)}) - \frac{d}{dx} f_{y'}(x, y, y', \ldots, y^{(m)}) +$$

$$+ \frac{d^2}{dx^2} f_{y''}(x, y, y', \ldots, y^{(m)}) + \ldots + (-1)^m \frac{d^m}{dx^m} f_{y^{(m)}}(x, y, y', \ldots$$

$$\ldots, y^{(m)}) = 0, \qquad (9.76)$$

называемого уравнением Эйлера — Пуассона.

☐ Пусть функция $\bar{y} \in Y$ реализует слабый локальный минимум. Зафиксируем произвольную функцию h, удовлетворяющую условиям (9.72), и рассмотрим семейство функций $\bar{y} + \lambda h$, зависящее от параметра $\lambda \in \mathbf{R}$. Если λ достаточно близко к нулю, то функции этого семейства принадлежат окрестности $U_\varepsilon^m(\bar{y})$ порядка m, которая существует согласно начальному допущению о том, что $\bar{y}$ реализует слабый локальный минимум. Функция $\bar{y}$ реализует также наименьшее значение целевого функционала на указанном множестве рассматриваемого семейства функций. Значит, числовая функция φ, определяемая равенством $\varphi(\lambda) = I(\bar{y} + \lambda h)$, достигает в точке $\lambda = 0$ локального минимума. По теореме Ферма, в результате дифференцирования под знаком интеграла получаем

$$\varphi'(0) = \int_{x_1}^{x_2} \left[\sum_{k=0}^{m} f_{y^{(k)}}(x, \bar{y}(x), \bar{y}'(x), \ldots, \bar{y}^{(m)}(x))\, h^{(k)}(x)\right] dx = 0. \qquad (9.77)$$

Проинтегрируем по частям слагаемое с номером k $m-k$ раз, $k = 0, 1, \ldots, m-1$. Учитывая граничные условия (9.72), придем к равенству

$$\varphi'(0) = \int_{x_1}^{x_2} \omega(x)\, h^{(m)}(x)\, dx = 0, \qquad (9.78)$$

где

$$\omega(x) = (-1)^m \int_{x_1}^{x} \int_{x_1}^{t_1} \ldots \int_{x_1}^{t_{m-1}} f_y(x, \bar{y}(x), \bar{y}'(x), \ldots, \bar{y}^{(m)}(x))\, dx\, dt_1 \ldots$$

$$\ldots \mathrm{d}t_{m-1} + \ldots + \int\limits_{x_1}^{x}\int\limits_{x_1}^{t_1} f_{y^{(m-2)}}(x, \bar{y}(x), \bar{y}'(x), \ldots, \bar{y}^{(m)}(x))\,\mathrm{d}x\,\mathrm{d}t_1 -$$

$$-\int\limits_{x_1}^{x} f_{y^{(m-1)}}(x, \bar{y}(x), \bar{y}'(x), \ldots, \bar{y}^{(m)}(x))\,\mathrm{d}x +$$

$$+ f_{y^{(m)}}(x, \bar{y}(x), \bar{y}'(x), \ldots, \bar{y}^{(m)}(x)).$$

Равенство (9.78) верно для любой функции h вида (9.72); кроме того, $\omega \in C[x_1, x_2]$. Поэтому, используя обобщенную лемму Дюбуа — Реймона, получаем $\omega(x) = p_{m-1}(x)$, $x \in [x_1, x_2]$, где p_{m-1} — многочлен, степень которого не больше чем $m-1$. В результате m-кратного дифференцирования равенства $\omega(x) = p_{m-1}(x)$ получим, что левая часть равенства (9.75) при $y = \bar{y}(x)$, $x \in [x_1, x_2]$ равна нулю. Это означает, что функция $\bar{y}$ удовлетворяет уравнению (9.75).

Если $f_{y^k}(x, \bar{y}(x), \bar{y}'(x), \ldots, \bar{y}^{(m)}(x)) \in C_k[x_1, x_2]$, $k = 0, 1, \ldots, m-1$, то в (9.75) при $y = \bar{y}(x)$ возможно дифференцирование. Дифференцируя, получаем, что функция $y = \bar{y}(x)$ удовлетворяет уравнению Эйлера — Пуассона (9.76). ■

Теорема 9.8 при $m=1$ совпадает с уже доказанной теоремой 9.1 Решения уравнения (9.76) называют *экстремалями* задачи высшего порядка.

Для отыскания экстремалей задачи высшего порядка следует решить дифференциальное уравнение Эйлера — Пуассона (9.76), которое имеет порядок $2m$. Его общее решение в общем случае зависит от $2m$ произвольных постоянных; для их определения используют краевые условия $y^{(k)}(x_i) = y_{ik}$, $k = 0, 1, \ldots, m-1$; $i = 1, 2$, число которых также равно $2m$. В этом смысле система необходимых условий является полной, что, однако, не гарантирует ее разрешимости (тем более однозначной разрешимости).

4. Задача определения формы прогиба балки. Цилиндрическая упругая балка AB длины $2l$ расположена горизонтально. Оба ее конца жестко закреплены. Требуется установить форму прогиба осевой линии балки.

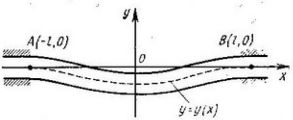

Рис. 9.12

Свяжем с балкой прямоугольную декартову систему координат (рис. 9.12) и будем считать, что параметры s и φ имеют тот же смысл, что и в задаче определения критической нагрузки балки из § 9.4. Обозначим через $y = y(x)$ уравнение оси балки. Потенциальная энергия балки складывается из потенциальной энергии, созданной силами упругости,

$$W_1 = \frac{1}{2}\mu\int\limits_0^{2l}\left(\frac{\mathrm{d}\varphi}{\mathrm{d}s}\right)^2 \mathrm{d}s,$$

и потенциальной энергии, созданной полем тяготения,

$$W_2 = \rho g \int_0^{2l} y \, \mathrm{d}s,$$

где ρ — линейная плотность балки. Перейдем от переменной s к переменной x. Для этого воспользуемся равенством $\mathrm{d}s = \sqrt{1 + y'^2(x)}\,\mathrm{d}x$. Учитывая, что $\mathrm{d}\varphi/\mathrm{d}s$ — это кривизна кривой $y = y(x)$, имеем [16]:

$$\mathrm{d}\varphi/\mathrm{d}s = y''/(1 + y'^2)^{3/2}.$$

Следовательно, полная потенциальная энергия

$$W_п(y) = W_1 + W_2 = \int_{-l}^{l} \left[\frac{1}{2} \mu \frac{y''^2(x)}{(1 + y'^2(x))^{5/2}} + \right.$$

$$\left. + \rho g y(x) \sqrt{1 + y'^2(x)} \right] \mathrm{d}x.$$

Будем считать, что прогиб балки невелик, и упростим полученное выражение, полагая $y'(x) \approx 0$, $x \in [-l, l]$:

$$W_п(y) = \int_{-l}^{l} \left[\frac{1}{2} \mu y''^2(x) + \rho g y(x) \right] \mathrm{d}x. \qquad (9.79)$$

Положению равновесия балки соответствует минимум потенциальной энергии. Поэтому задача сводится к отысканию функции $y \in C_2[-l, l]$, реализующей наименьшее возможное значение функционала (9.76) и удовлетворяющей граничным условиям

$$y(-l) = y(l) = y'(-l) = y'(l) = 0.$$

Эта задача второго порядка. Запишем для нее уравнение Эйлера — Пуассона (9.76). Поскольку $f_y = \rho g$, $f_{y'} = 0$, $f_{y''} = \mu y''$, уравнение (9.76) принимает вид

$$\mu \frac{\mathrm{d}^2 y''}{\mathrm{d}x^2} = -\rho g.$$

Отсюда, интегрируя, получаем общее решение уравнения Эйлера — Пуассона:

$$y'' = -\frac{\rho g}{2\mu} x^2 + C_1 x + C_2,$$

$$y = -\frac{\rho g}{24\mu} x^4 + C_3 x^3 + C_4 x^2 + C_5 x + C_6,$$

где $C_1, C_2, \ldots, C_6$ — произвольные постоянные. В силу симметрии

$C_3=C_5=0$. Значения остальных постоянных находятся из условий

$$y(-l)=y(l)=-\frac{\rho g}{24\mu}l^4+C_4l^2+C_6=0,$$

$$y'(-l)=y'(l)=-\frac{\rho g}{6\mu}l^3+2C_4l=0.$$

А именно, $C_4=\rho gl^2/(12\mu)$, $C_6=-\rho gl^2/(24\mu)$. Окончательно получаем единственную экстремаль

$$y=-\frac{\rho g}{24\mu}(-x^4+2l^2x^2-l^4)=-\frac{\rho g}{24\mu}(x^2-l^2)^2,$$

которая определяет решение исходной задачи.

5. Задача Лагранжа с дифференциальными ограничениями. Рассматриваемую в этом пункте задачу будем называть *задачей Лагранжа в понтрягинской форме*. Эта форма записи вариационной задачи близка к формулировке задачи оптимального управления, которая изучается в последующих главах. Ее основная особенность заключается в «отделении» функции y от производной y' с помощью замены y' на новую функцию u. В результате получается задача, в которой минимизация целевого функционала осуществляется не только подбором функции y, но также выбором функции u, называемой *управлением*.

Минимизируемый функционал типа Лагранжа

$$I(y,\ u)=\int_{x_1}^{x_2} f(x,\ y(x),\ u(x))\,\mathrm{d}\,x \tag{9.80}$$

определен на парах векторных функций вида $y\colon [x_1,\ x_2]\to\mathbf{R}^n$, $u\colon [x_1,\ x_2]\to\mathbf{R}^m$, причем $y\in C_1{}^n[x_1,\ x_2]$, $u\in C^m[x_1,\ x_2]$. Другими словами, функцконал (9.80) определен на декартовом произведении $C_1{}^n[x_1,\ x_2]\times C^m[x_1,\ x_2]$.

Отрезок $[x_1,\ x_2]$ будем считать фиксированным, а функцию $f(n+m+1$ переменных) — непрерывно дифференцируемой на некоторой области $\hat{D}\subset\mathbf{R}^{n+m+1}$.

Множество допустимых функций $\mathbf{Y}$ в рассматриваемой задаче образуют пары $(y,\ u)$ векторных функций, для которых $y\in C_1{}^n[x_1,\ x_2]$ $u\in C^m[x_1,\ x_2]$ и, кроме того, $y'=\varphi(x,\ y(x),\ u(x))$ при всех $x\in[x_1,\ x_2]$,

$$\Phi(y(x_1))=0_{n_1},\ \ \Theta(y(x_2))=0_{n_2}.$$

Здесь $y'=(y'_1,\ y'_2,\ \ldots,\ y'_n)$; векторная функция $\varphi\colon \hat{D}\to\mathbf{R}^n$ предполагается покомпонентно непрерывно дифференцируемой на множестве $\hat{D}$, причем включение $(x,\ y(x),\ u(x))\in\hat{D}$ имеет место при всех $x\in[x_1,\ x_2]$. Векторные функции Φ и Θ (размерности n_1 и n_2 соответственно) в граничных условиях будем считать непрерывно дифференцируемыми на некоторых областях пространства $\mathbf{R}^n$, содержащих соответственно точки $y(x_1)$ и $y(x_2)$.

Допустимую пару функций $(\bar{y},\ \bar{u})$ называют *решением* задачи Лагранжа в понтрягинской форме, если неравенство $I(\bar{y},\ \bar{u})\leqslant$

$\leqslant I(y, u)$ имеет место для всех $(y, u) \in \mathbf{Y}$. Говорят, что допустимая пара функций $(\bar{y}, \bar{u})$ является *слабым локальным решением*, если найдется такое число $\varepsilon > 0$, что неравенство $I(\bar{y}, \bar{u}) \leqslant I(y, u)$ выполняется для всех $(y, u) \in \mathbf{Y} \cap [\mathbf{U}_\varepsilon^1(\bar{y}) \times \mathbf{U}_\varepsilon^0(\bar{u})]$. Здесь $\mathbf{U}_\varepsilon^1(\bar{y})$ — окрестность n-мерной вектор-функции $\bar{y}$, а $\mathbf{U}_\varepsilon^0(\bar{u})$ — окрестность m-мерной вектор-функции $\bar{u}$.

Ограничения в виде дифференциальной связи $y' = \varphi(x, y, u)$ (а также граничные условия) учитывают, вводя *функцию (функционал) Лагранжа* вида

$$\mathbf{L}(y, u; \mu, \lambda_0, \lambda_1, \lambda_2) = \int_{x_1}^{x_2} L(x, y(x), y'(x), u(x), \mu(x), \lambda_0)\, \mathrm{d}x + \\ + \Psi(y; \lambda_1, \lambda_2),$$

где

$$L(x, y, y', u, \mu, \lambda_0) = \lambda_0 f(x, y, u) + \sum_{k=1}^{n} \mu_k (y'_k - \varphi_k(x, y, u)),$$

$$\Psi(y; \lambda_1, \lambda_2) = \langle \lambda_1, \Phi(y(x_1)) \rangle + \langle \lambda_2, \Theta(y(x_2)) \rangle.$$

Здесь $\mu \in C_1^n[x_1, x_2]$, $\lambda_0 \in \mathbf{R}$, $\lambda_1 \in \mathbf{R}^{n_1}$, $\lambda_2 \in \mathbf{R}^{n_2}$ — множители Лагранжа, а нижний индекс k означает k-ю компоненту соответствующего вектора. Множитель μ_k представляет собой функцию, заданную на $[x_1, x_2]$; с его помощью учитывают ограничение $y_k' - \varphi_k(x, y, u) = 0$, которое представимо в виде $G_k(y, u) = 0$, где G_k — оператор, действующий из множества $\mathbf{Y}$ в пространство $C[x_1, x_2]$ (см. п. 1, § 8.4). Функция Лагранжа содержит терминальный член. Следующая ниже теорема показывает, что решение сформулированной условной задачи Лагранжа в понтрягинской форме можно свести к решению безусловной задачи Больца с целевым функционалом $\mathbf{L}$. В этом смысле ее формулировку полезно сравнить с формулировкой теоремы 9.4.

Теорема 9.9. *Для того чтобы пара функций $(\bar{y}, \bar{u})$ реализовала слабый локальный минимум в задаче Лагранжа в понтрягинской форме, необходимо, чтобы существовали такие не равные одновременно нулю* множители Лагранжа $\mu \in C_1^n[x_1, x_2]$, $\bar{\lambda}_0 \in \mathbf{R}$, $\bar{\lambda}_1 \in \mathbf{R}^{n_1}$, $\bar{\lambda}_2 \in \mathbf{R}^{n_2}$, что:*

1) *пара функций $(\bar{y}, \bar{u})$ являлась решением уравнения Эйлера — Лагранжа по y, т. е. для всех $x \in [x_1, x_2]$*

$$\left(L_y - \frac{\mathrm{d}}{\mathrm{d}x} L_{y'}\right)\Bigg|_{\substack{y=\bar{y}(x)\\ u=\bar{u}(x)}} = 0_n; \tag{9.81}$$

---

* Т. е. либо среди $\bar{\lambda}_0$ и компонент векторов $\bar{\lambda}_1$, $\bar{\lambda}_2$ есть хотя бы одно число, отличное от нуля, либо функция $\bar{\mu}(x)$, не равна тождественно нулю на отрезке $[x_1, x_2]$.

2) *пара функций* $(\bar{y}, \bar{u})$ *являлась решением уравнения Эйлера — Лагранжа по* u, *т. е. для всех* $x \in [x_1, x_2]$

$$L_u \Big|_{\substack{y=\bar{y}(x)\\ u=\bar{u}(x)}} = 0_m; \tag{9.82}$$

3) *выполнялись условия трансверсальности*

$$L_{y'} \Big|_{\substack{y=\bar{y}(x)\\ u=\bar{u}(x)}} = \langle \lambda_1, \Phi'(\bar{y}(x_1)) \rangle, \tag{9.83}$$

$$L_{y'} \Big|_{\substack{y=\bar{y}(x)\\ u=\bar{u}(x)}} = -\langle \lambda_2, \Theta'(y(x_2)) \rangle. \tag{9.84}$$

Теорему 9.9 можно получить как следствие теоремы 8.4 [2].

Запишем подробно уравнения Эйлера — Лагранжа и условия трансверсальности, которым должна удовлетворять оптимальная пара функций $(\bar{y}, \bar{u})$:

$$\frac{d\mu_i}{dx} = \lambda_0 f_{y_i}(x, y, u) - \sum_{k=1}^{n} \mu_k \varphi_{ky_i}(x, y, u), \ i = 1, 2, \ldots, n; \tag{9.81'}$$

$$\lambda_0 f_{u_j}(x, y, u) = \sum_{k=1}^{n} \mu_k \varphi_{ku_j}(x, y, u), \ j = 1, 2, \ldots, m; \tag{9.82'}$$

$$\mu_i(x_1) = \sum_{k=1}^{n} \lambda_{1k} \Phi_{ky_i}(y(x_1)), \ i = 1, 2, \ldots, n_1; \tag{9.83'}$$

$$\mu_i(x_2) = -\sum_{k=1}^{n} \lambda_{2k} \Theta_{ky_i}(y(x_2)), \ i = 1, 2, \ldots, n_2, \tag{9.84'}$$

Предположим, что из системы функциональных (не дифференциальных) уравнений (9.82′) удалось однозначно выразить функцию u через y и μ. Подставляя найденное выражение для u в (9.81′) и в уравнение $y' = \varphi(x, y, u)$, получим систему $2n$ дифференциальных уравнений первого порядка относительно y и μ. Эту систему следует решать совместно с уравнениями (9.83′) и (9.84′), а также с уравнениями $\Phi(y(x_1)) = 0_{n_1}$, $\Theta(y(x_2) = 0_{n_2}$. Всего имеем $4n + n_1 + n_2$ уравнений. Неизвестными являются функции $\mu_1, \mu_2, \ldots, \mu_n, y_1, y_2, \ldots, y_n$ и числовые множители Лагранжа, среди которых $n_1 + n_2$ независимых. Следовательно, учитывая, что при интегрировании системы дифференциальных уравнений возникнут $2n$ произвольных постоянных, получаем $4n + n_1 + n_2$ неизвестных. Таким образом, система необходимых условий полная (однако не всегда имеет, и тем более единственное, решение).

В простых задачах рассмотренный способ решения приводит к единственной экстремали, и если оптимальное решение существует, то найденная экстремаль и является решением.

6. Сводимость вариационных задач одного класса к задачам другого класса. Приведенная в п. 1 классификация вариационных задач в значительной степени условная. Это подтверждают следующие рассуждения.

Рассмотрим пространственную изопериметрическую задачу Лагранжа, предполагая для простоты, что имеется только одно ограничение:

$$\int_{x_1}^{x_2} f(x,\ y(x),\ y'(x))\,\mathrm{d}x \to \min,$$

$$\int_{x_1}^{x_2} g(x,\ y(x),\ y'(x))\,\mathrm{d}x = l.$$

Введем дополнительную функцию y_{n+1} и n-мерное управление u с помощью формул

$$y_{n+1}(x) = \int_{x_1}^{x} g(t,\ y(t),\ u(t))\,\mathrm{d}t;\quad u_i(x) = y_i'(x),$$

$$i = 1,\ 2, \ldots, n;\ x \in [x_1,\ x_2].$$

Тогда исходная задача эквивалентна следующей задаче Лагранжа в понтрягинской форме:

$$\int_{x_1}^{x_2} f(x,\ y(x),\ u(x))\,\mathrm{d}x \to \min;$$

$$\dot{y}_i = u_i(x),\ i = 1,\ 2, \ldots, n;$$

$$\dot{y}_{n+1} = g(x,\ y(x),\ u(x))$$

с дополнительными граничными условиями $y_{n+1}(x_1) = 0$, $y_{n+1}(x_2) = l$.

С помощью рассмотренного способа всегда можно «избавиться» от функциональных ограничений-равенств, перейдя к задаче Лагранжа в понтрягинской форме.

Теперь рассмотрим задачу Лагранжа высшего порядка с закрепленными концами:

$$\int_{x_1}^{x_2} f(x,\ y(x),\ y'(x), \ldots, y^{(m)}(x))\,\mathrm{d}x \to \min;$$

$$y^{(k)}(x_1) = y_{1k};\ y^{(k)}(x_2) = y_{2k},\ k = 0,\ 1, \ldots, m-1.$$

Не уменьшая общности, ограничимся плоской задачей, т. е. считаем, что y — скалярная функция. Вводим функции $y_1, \ldots, y_m$ и скалярное управление u с помощью равенств

$$y_1(x) = y(x),\ y_i(x) = y_{i-1}'(x),\ i = 2,\ 3, \ldots, m;\ u(x) = y_m'(x),$$

$$x \in [x_1,\ x_2].$$

Тогда исходная задача высшего порядка равносильна задаче Лагранжа в понтрягинской форме

$$\int_{x_1}^{x_2} f(x,\ y_1(x),\ y_2(x),\dots,y_m(x),\ u(x))\,\mathrm{d}x \to \min;$$

$$y_i' = y_{i+1}(x),\ i=1,\ 2,\dots,m-1,$$

$$y_m' = u$$

с граничными условиями

$$y_{i+1}(x_1) = y_{1i};\ y_{i+1}(x_2) = y_{2i},\ i=0,\ 1,\dots,m-1.$$

Следовательно, переходя к задаче Лагранжа в понтрягинской форме, можно «избавиться» также от производных более высокого порядка.

Выясним, как вариационную задачу с интегральным целевым функционалом можно свести к задаче с терминальным функционалом и дополнительным дифференциальным ограничением-равенством. Для простоты рассмотрим простейшую вариационную задачу. Введем функции y_1, y_2 с помощью равенств

$$y_1(x) = y(x),\ y_2(x) = \int_{x_1}^{x} f(t,\ y(t),\ y'(t))\,\mathrm{d}t,\ x \in [x_1,\ x_2].$$

В этом случае простейшая вариационная задача эквивалентна задаче с терминальным функционалом

$$\Psi = y_2(x_2) \to \min,$$

дифференциальным ограничением $y_2' = f(x,\ y_1,\ y_1')$ и дополнительным граничным условием $y_2(x_1) = 0$.

§ 9.7. ПРИКЛАДНЫЕ ЗАДАЧИ

1. Задача выбора оптимального режима работы термоэлектрического генератора*. Принцип работы термоэлектрического генератора основан на эффекте Зеебека. В *термопаре,* составленной из двух последовательно соединенных разнородных металлических проводников, контакты между которыми имеют различную температуру, возникает электродвижущая сила, называемая *термо-э.д.с.* Значение ее пропорционально разности температур горячего и холодного контактов. Если температуры контактов поддерживаются постоянными с помощью специальных источника и поглотителя теплоты, то термо-э.д.с. имеет постоянное значение.

В термопаре часть теплоты, полученной от источника, превращается в электрическую энергию, а остальная теплота отдается ох-

* См.: *Шехтер Р. С.* Вариационный метод в инженерных расчетах. — М.: Мир, 1971.

ладителю. Если обозначить через dQ количество теплоты, полученной термопарой от источника, то количество выработанной при этом электрической энергии dW можно вычислить по формуле

$$dW=\theta(T_г-T_х)\,dQ. \tag{9.85}$$

Здесь $T_г$ и $T_х$ — температуры горячего и холодного контактов термопары соответственно; θ — коэффициент, характеризующий эффективность преобразования тепловой энергии в электрическую и имеющий размерность $К^{-1}$.

Для практической реализации термоэлектрического эффекта была предложена следующая конструкция (рис. 9.13). Генератор представляет собой помещенную в кожух трубу длины L, которая составлена из чередующихся колец, сделанных из двух раз-

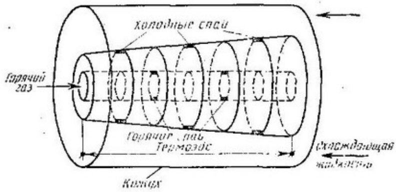

Рис. 9.13

личных металлов. Кольца изолированы друг от друга и имеют электрический контакт (спай) попеременно только с внутренней или внешней стороны трубы. Каждые два соседних кольца можно рассматривать как термопару. Внутри трубы проходит горячий газ, служащий источником теплоты. Между трубой и внешним кожухом протекает охлаждающая жидкость (например, вода), поглощающая теплоту. Температура жидкости постоянна, а температура газа максимальна на входе в генератор и уменьшается при движении газа вдоль трубы. Теплота, отдаваемая газом, преобразуется в электрическую энергию, и на концах трубы возникает э.д.с. Выработанная электрическая энергия, как видно из формулы (9.85), тем больше, чем больше количество теплоты dQ, отдаваемой газом внутренней поверхности бесконечно малого цилиндрического элемента трубы и чем больше разность температур ΔT между горячими и холодными контактами. Эти величины — dQ и ΔT (теплота, отдаваемая газом, и разность температур) зависят от коэффициентов теплоотдачи $\alpha_г$ и $\alpha_ж$, определяющих скорость распространения теплоты от газа к стенке трубы и от нее к жидкости соответственно, а также от удельной теплопроводности Λ материалов, из которых состоит труба, и от толщины стенки $l(x)$.

При фиксированных значениях длины L трубы, внутреннего диаметра D, массового расхода M газа, начальной температуры T_0 газа и температуры $T_ж$ охлаждающей жидкости распределение температуры T_x вдоль трубы зависит только от толщины $l(x)$ стенки трубы как функции координаты x. Изменение вида функции $T(x)$ вызывает изменение количества выработанной генератором электрической энергии, поэтому возникает задача выбора такой функции $l(x)$, чтобы выработанная генератором электрическая энергия W была максимальной. Это типичная вариационная задача.

В самом деле, так как в выражении (9.85) величины T_x, $T_г$ и dQ зависят только от координаты x и эта зависимость, как было установлено, определяется функцией $l(x)$, то полная энергия W, выработанная в генераторе, равна

$$W = \int_0^L F(x,\ l(x),\ l'(x))\,dx. \tag{9.86}$$

Задача максимизации выработанной электрической энергии эквивалентна следующей вариационной задаче: найти непрерывную функцию $l(x)$, для которой функционал (9.86) принимает максимальное значение. Очевидно, на краях интервала $[0,\ L]$ значения функции $l(x)$ могут быть произвольными, поэтому данная вариационная задача является задачей со свободными концами.

Отметим, что выражать все величины в уравнении (9.85) через функцию $l(x)$ неудобно, поэтому будем решать другую, эквивалентную данной задачу: найти такое распределение температуры газа $T(x)$, чтобы величина W была максимальной. При этом необходимо представить W в следующем виде:

$$W = \int_0^L F_1(x,\ T(x),\ T'(x))\,dx. \tag{9.87}$$

Значение функции $T(x)$ на левом конце интервала $[0,\ L]$ фиксировано, поскольку начальная температура газа поддерживается постоянной:

$$T(x)\mid_{x=0} = T_0. \tag{9.88}$$

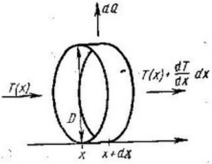

Рис. 9.14

Таким образом, вариационная задача с функционалом (9.87) является комбинированной: левый конец фазовой траектории закреплен, а правый — свободен.

Чтобы определить вид функции F_1 в выражении (9.87), выведем уравнение теплового баланса для бесконечно малого цилиндрического элемента трубы (рис. 9.14). Предполагая, что диаметр трубы велик по сравнению с толщиной стенки, используем для описания процесса распространения теплоты внутри трубы закон Фурье. Согласно этому закону, количество теплоты, которое переносится от внутренней поверхности dS элемента трубы длины dx к его внешней поверхности, определяется выражением

$$dQ = \Lambda\,\frac{T_г - T_x}{l}\,dS, \tag{9.89}$$

где $dS = \pi D dx$ — площадь внутренней поверхности цилиндрического элемента трубы.

Формула (9.89) справедлива в том случае, когда теплота не поглощается стенкой. В рассматриваемом же случае именно за счет поглощенной теплоты и вырабатывается электрическая энергия. Однако эффективность θ такого преобразования очень мала, поэтому в первом приближении можно считать, что количество теплоты, отданной газом, равно количеству теплоты, полученному охлаждающей жидкостью. Тогда значения теплоты в формулах (9.85) и (9.89) равны.

Выразив разность $T_г - T_х$ из (9.89) и подставляя в (9.85), получаем

$$dW = \frac{\theta l}{\Lambda dS} (dQ)^2. \tag{9.90}$$

Теплота dQ, уходящая из газа при прохождении рассматриваемого цилиндрического элемента трубы, определяется выражением

$$dQ = -Mc_p \frac{dT}{dx} dx, \tag{9.91}$$

где c_p — удельная теплоемкость газа при постоянном давлении, $T = T(x)$ — температура газа в точке x.

Уравнение, описывающее теплопередачу во всей системе газ — стенка трубы — жидкость, имеет вид

$$dQ = k(T - T_ж) dS, \tag{9.92}$$

где $T_ж$ — температура охлаждающей жидкости, k — коэффициент теплопередачи всей системы, зависящий от коэффициентов теплоотдачи $\alpha_г$ и $\alpha_ж$, а также от удельной теплопроводности стенок. Коэффициент k можно вычислить по формуле

$$k = \frac{1}{\frac{1}{\alpha_г} + \frac{l}{\Lambda} + \frac{1}{\alpha_ж}}. \tag{9.93}$$

Подставляя (9.93) в (9.92), а (9.92) в (9.90), получаем выражение для dW:

$$dW = \frac{\theta l}{\Lambda} \left(\frac{T - T_ж}{\frac{1}{\alpha_г} + \frac{l}{\Lambda} + \frac{1}{\alpha_ж}} \right)^2 dS. \tag{9.94}$$

Выразим в этой формуле $l(x)$ через $T(x)$. Сравнив уравнения (9.91) и (9.92) и учтя (9.93), запишем:

$$\frac{T - T_ж}{\frac{1}{\alpha_г} + \frac{l}{\Lambda} + \frac{1}{\alpha_ж}} \pi D dx = -Mc_p \frac{dT}{dx} dx.$$

Решим это уравнение относительно l:

$$l(x) = -\frac{\Lambda\pi D(T-T_ж)}{Mc_p\frac{dT}{dx}} - \Lambda\left(\frac{1}{\alpha_г} + \frac{1}{\alpha_ж}\right). \tag{9.95}$$

Подставляя (9.95) в формулу (9.94) и интегрируя получившееся выражение по x от 0 до L, найдем

$$W = \int_0^L \theta\left[-Mc_p\frac{dT}{dx}(T-T_ж) - \frac{1}{\pi D}\left(\frac{1}{\alpha_г} + \frac{1}{\alpha_ж}\right)\left(Mc_p\frac{dT}{dx}\right)^2\right]dx. \tag{9.96}$$

Это выражение соответствует (9.87). Чтобы найти функцию $T(x)$, которая максимизирует W, выпишем уравнение Эйлера, которое для данного случая имеет вид

$$\frac{d^2T}{dx^2} = 0. \tag{9.97}$$

На левом конце отрезка [0, 2] значение T фиксировано условием (9.88), на правом получаем естественное граничное условие

$$\left[C_1(T-T_ж) + 2C_2\frac{dT}{dx}\right]\bigg|_{x=L} = 0, \tag{9.98}$$

где $C_1 = -\theta Mc_p$, $C_2 = -\frac{\theta}{\pi D}\left(\frac{1}{\alpha_г} + \frac{1}{\alpha_ж}\right)(Mc_p)^2$.

Общее решение уравнения (9.97), очевидно, имеет вид

$$T(x) = a_1 + a_2 x. \tag{9.99}$$

Постоянные a_1 и a_2 определяются из граничных условий (9.88) и (9.98). Из (9.88) следует

$$a_1 = T_0. \tag{9.100}$$

Подставив (9.99) в условие (9.98), получаем уравнение для a_2

$$C_1(T_0 + a_2L - T_ж) + 2C_2a_2 = 0$$

из которого находим

$$a_2 = -\frac{C_1(T_0-T_ж)}{C_1L + 2C_2} = -\frac{T_0-T_ж}{L + 2\frac{C_2}{C_1}} =$$

$$= -\frac{T_0-T_ж}{L + \frac{2Mc_p}{\pi D}\left(\frac{1}{\alpha_г} + \frac{1}{\alpha_ж}\right)}. \tag{9.101}$$

Определим теперь количество выработанной электрической энергии, которое соответствует оптимальному режиму, задаваемому распределением (9.99). Подставляя (9.99) в (9.96) и интегрируя, получаем

$$W_{\text{опт}}=\int_0^L \theta\left[-Mc_p a_2(T_0+a_2x-T_{\text{ж}})-\frac{1}{\pi D}\left(\frac{1}{\alpha_{\text{г}}}+\frac{1}{\alpha_{\text{ж}}}\right)(Mc_p a_2)^2\right]\mathrm{d}x=$$

$$=-\theta\left[\frac{Mc_p a_2^2 L^2}{2}+Mc_p a_2(T_0-T_{\text{ж}})L+\frac{1}{\pi D}\left(\frac{1}{\alpha_{\text{г}}}+\frac{1}{\alpha_{\text{ж}}}\right)(Mc_p a_2)^2 L\right].$$

Используя выражение (9.101), окончательно находим

$$W_{\text{опт}}=-\theta\left[\frac{1}{2}\frac{(T_0-T_{\text{ж}})^2 LMc_p}{L+\dfrac{2Mc_p}{\pi D}\left(\dfrac{1}{\alpha_{\text{г}}}+\dfrac{1}{\alpha_{\text{ж}}}\right)}-\right.$$

$$\left.-\frac{(T_0-T_{\text{ж}})^2 LMc_p}{L+\dfrac{2Mc_p}{\pi D}\left(\dfrac{1}{\alpha_{\text{г}}}+\dfrac{1}{\alpha_{\text{ж}}}\right)}\right]=\frac{\theta Mc_p(T_0-T_{\text{ж}})^2}{2+\left(\dfrac{1}{\alpha_{\text{г}}}+\dfrac{1}{\alpha_{\text{ж}}}\right)\dfrac{4Mc_p}{\pi DL}}. \qquad (9.102)$$

Видно, что количество выработанной в оптимальном режиме электрической энергии увеличивается с увеличением значения коэффициентов теплоотдачи $\alpha_{\text{г}}$ и $\alpha_{\text{ж}}$ и разности температур $T_0-T_{\text{ж}}$ и с увеличением эффективности преобразования энергии.

Подставив $T(x)$ в формулу (9.95), получим искомую зависимость толщины стенки трубы от координаты:

$$l(x)=-\frac{\Lambda\pi D(T-T_{\text{ж}})}{Mc_p\dfrac{\mathrm{d}T}{\mathrm{d}x}}-\Lambda\left(\frac{1}{\alpha_{\text{г}}}+\frac{1}{\alpha_{\text{ж}}}\right)=-\left[\frac{\Lambda\pi D(T_0-T_{\text{ж}})}{Mc_p a_2}+\right.$$

$$\left.+\Lambda\left(\frac{1}{\alpha_{\text{г}}}+\frac{1}{\alpha_{\text{ж}}}\right)\right]-\frac{\Lambda\pi D}{Mc_p}x=\left[\Lambda\left(\frac{1}{\alpha_{\text{г}}}+\frac{1}{\alpha_{\text{ж}}}\right)+\frac{\Lambda\pi D}{Mc_p}L\right]-\frac{\Lambda\pi D}{Mc_p}x. \qquad (9.103)$$

Таким образом, $l(x)$ — убывающая линейная функция, т. е. на входе стенки трубы должна иметь наибольшую толщину, а на выходе — наименьшую толщину (в рамках допущения $l\ll D$).

Чтобы проверить, действительно ли при этом имеет место максимум, а не минимум функционала (9.86), используем необходимое условие Лагранжа. В рассматриваемом случае

$$F_{\dot u\dot u}=-\frac{2}{\pi D}\left(\frac{1}{\alpha_{\text{г}}}+\frac{1}{\alpha_{\text{ж}}}\right)(Mc_p)^2<0.$$

Распределение толщины стенки трубы (9.103) действительно дает максимальное количество выработанной электрической энергии, которое определяется выражением (9.102).

В рассмотренной вариационной задаче необходимые условия экстремума позволили найти решение, поскольку из простых физических соображений следует существование единственного экстремума. Однако такая ситуация встречается далеко не во всех технических задачах. Возьмем, например, вариационную задачу с тем же функционалом W, определяемым равенством (9.96), но предположим, что коэффициенты теплоотдачи $\alpha_г$ и $\alpha_ж$ велики по сравнению с коэффициентом $Mc_p/\pi D$. Тогда вторым слагаемым в интеграле (9.96) можно пренебречь:

$$W = -\theta Mc_p \int_0^L \frac{dT}{dx}(T - T_ж)\, dx. \qquad (9.104)$$

Уравнение Эйлера для такого функционала обращается в тождество

$$\frac{dT}{dx} - \frac{dT}{dx} = 0.$$

В данном случае это означает, что функционал не имеет экстремума. Этот факт можно установить непосредственно из рассмотрения выражения (9.104). Производя интегрирование по частям, получаем

$$W = -\theta Mc_p \left[T(T - T_ж) \Big|_0^L - \int_0^L \frac{dT}{dx} T\, dx \right] =$$

$$= \theta Mc_p \left(TT_ж - \frac{1}{2} T^2 \right) \Big|_0^L.$$

Для всех функций $T(x)$, имеющих одинаковые граничные условия, функционал $W(T)$ имеет одно и то же значение.

2. Задача выбора оптимального профиля ребер радиатора, предназначенного для охлаждения нагретой стенки *. Рассматриваемый радиатор применяется для увеличения потока теплоты от нагретой стенки в окружающую среду (воздух). При конструировании такого радиатора возникают задачи об оптимальном выборе числа ребер, материала, из которого они изготовлены, и др. Рассмотрим простейшую задачу оптимизации профиля ребра радиатора, считая заданным материал ребра и его массу.

На рис. 9.15 изображено сечение ребра радиатора. Будем считать, что перенос теплоты в направлении, перпендикулярном плоскости сечения, отсутствует. Выберем такую шкалу температур, в которой температура стенки равна единице, а температура воздуха — нулю.

Выведем уравнение теплового баланса. Для этого рассмотрим бесконечно малый элемент ребра (рис. 9.16). Количество теплоты,

* См.: *Шехтер Р. С.* Вариационный метод в инженерных расчетах. — М.: Мир, 1971.

распространяющейся в направлении оси Ox (вследствие теплопроводности внутри ребра), описывается формулой

$$Q = -\Lambda 2D(x)\frac{dT}{dx},$$

где Λ — коэффициент теплопроводности материала, из которого сделано ребро, $D(x)$ — функция, описывающая профиль ребра, $T(x)$ — распределение температуры в ребре. Изменение количест-

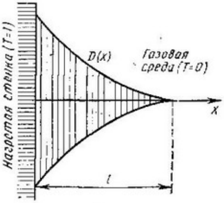

Рис. 9.15

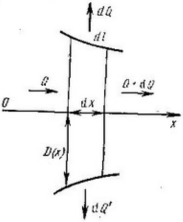

Рис. 9.16

ва теплоты dQ на промежутке dx может быть выражено следующим образом:

$$dQ = \frac{dQ}{dx}\,dx = -2\Lambda\frac{d}{dx}\left(D(x)\frac{dT}{dx}\right)dx. \tag{9.105}$$

Согласно закону сохранения энергии, полное изменение количества теплоты на промежутке dx должно быть равно нулю, т. е.

$$dQ + 2\,dQ' = 0, \tag{9.106}$$

где dQ' — количество теплоты, отдаваемое поверхностью dl ребра; это количество теплоты выразим соотношением

$$dQ' = \alpha\,dl\cdot T(x) = \alpha\sqrt{1+\left(\frac{dD}{dx}\right)^2}\,dx\,T(x), \tag{9.107}$$

где α — коэффициент теплоотдачи от ребра к воздуху. Подставляя (9.105) и (9.107) в (9.106), получим уравнение теплового баланса

$$\frac{d}{dx}\left[D(x)\frac{dT}{dx}\right] = \frac{\alpha}{\Lambda}\sqrt{1+\left(\frac{dD}{dx}\right)^2}\,T(x).$$

Предположим, что величина $\left(\frac{dD}{dx}\right)^2$ мала по сравнению с единицей, т. е. поверхность ребра является достаточно покатой. Тогда

$$\frac{d}{dx}\left[D(x)\frac{dT}{dx}\right] = \frac{\alpha}{\Lambda}T(x). \tag{9.108}$$

К этому уравнению необходимо добавить два граничных условия: 1) температура нагретой стенки фиксирована:

$$T(x)\mid_{x=0}=T(0)=1; \tag{9.109}$$

2) поперечные размеры ребра конечны:

$$D(x)\mid_{x=l}=0. \tag{9.110}$$

Здесь величина l является искомой. Полное количество теплоты, отдаваемое поверхностью ребра, равно

$$Q(T)=2\int_0^l \frac{dQ}{dx}\,dx=2\int_0^l \alpha T(x)\,dx. \tag{9.111}$$

Таким образом, количество отдаваемой теплоты Q можно рассматривать как интегральный функционал, заданный на множестве функций $T(x)$, удовлетворяющих уравнению (9.108) и условию (9.109).

Условие постоянства массы ребра радиатора запишем в виде

$$M=2\rho\int_0^l D(x)\,dx=\text{const}, \tag{9.112}$$

где ρ — плотность материала, из которого сделано ребро.

Необходимо найти максимум функционала (9.111) при условии (9.112). Применяя метод множителей Лагранжа, можно свести эту задачу к задаче отыскания максимума функционала

$$I(T,\ D)=\int_0^l [2\alpha T(x)+2\lambda\rho D(x)]\,dx. \tag{9.113}$$

Однако в данном случае введение множителя Лагранжа λ не свело условную вариационную задачу к безусловной, поскольку кроме условия (9.112) имеется еще ограничение в виде дифференциального уравнения (9.108). Задача (9.113) с ограничением (9.108) по-прежнему является вариационной задачей на условный экстремум. Поскольку уравнение (9.108) определяет явную связь между функциями $T(x)$ и $D(x)$, выразим в функционале (9.113) $D(x)$ через $T(x)$ с помощью (9.108) и получим безусловную вариационную задачу с функционалом, заданным на функциях $T(x)$.

Интегрируя уравнение теплового баланса (9.108), найдем

$$-D(x)\frac{dT}{dx}=\frac{\alpha}{\Lambda}\int_x^l T(x')\,dx'. \tag{9.114}$$

Здесь предполагается, что количество теплоты в точке $x=l$ (край ребра) равно нулю, т. е. использовано граничное условие (9.110).

Выразим из (9.114) функцию $D(x)$ и подставим ее в (9.113). В результате получим функционал

$$I_1(T)=2\int_0^l\left[\alpha T-\frac{\lambda\rho\alpha}{\Delta}\left(\frac{dT}{dx}\right)^{-1}\int_x^l T\,dx'\right]dx. \tag{9.115}$$

Отметим, что в выражениях (9.113) и (9.115) величина l не является, вообще говоря, фиксированной она может изменяться при изменении функции $D(x)$, определяющей профиль ребра. Но так как функция $D(x)$ исключена из (9.115), можно считать l фиксированным и искать оптимальное распределение $T(x)$ при данном l. После решения такой задачи можно будет, кроме того, выбрать оптимальное l (из всех l, удовлетворяющих условию постоянства массы ребра).

Итак, будем решать безусловную вариационную задачу с функционалом (9.115). Единственным ограничением на функции $T(x)$ является граничное условие (9.109): $T(0)=1$ (граничное условие (9.110) было использовано при выводе соотношения (9.114)). Отличие этой задачи от простейшей вариационной задачи (см. п. 1 § 9.1) состоит в том, что функционал (9.115) содержит первообразную от функции $T(x)$:

$$I(y)=\int_{x_0}^{x_1}F\left(x,\ y,\ \dot{y},\ \int_x^{x_1}y(x')\,dx'\right)dx. \tag{9.116}$$

Однако для функционала такого вида процедура вывода уравнения Эйлера совершенно аналогична той, что была применена для функционала, включающего производные функции $y(x)$. Уравнение Эйлера имеет вид

$$F_y-\frac{d}{dx}F_{\dot{y}}+\int_{x_1}^{x_1}F_z\,dx''=0, \tag{9.117}$$

где $z=\int_x^{x_1}y(x')\,dx'$, т. е. $F_z(x,\ y,\ \dot{y},\ z)$ обозначает частную производную функции F по четвертой переменной

К уравнению (9.117) необходимо добавить естественное граничное условие на правом конце

$$F_{\dot{y}}\mid_{x=x_1}=0. \tag{9.118}$$

Для вариационной задачи с функционалом (9.115) уравнение (9.117) будет иметь вид

$$\alpha+\frac{\lambda\rho\alpha}{\Delta}\left[\frac{2\dfrac{d^2T}{dx^2}}{\left(\dfrac{dT}{dx}\right)^3}\int_x^l T\,dx'+\frac{T}{\left(\dfrac{dT}{dx}\right)^2}-\int_0^x\frac{dx'}{\dfrac{dT}{dx}}\right]=0 \tag{9.119}$$

(в рассматриваемом случае $x_0=0$, $x_1=l$). Естественное граничное условие (9.118) примет вид

$$\left[\frac{\lambda\rho\alpha}{\Lambda}\left(\frac{dT}{dx}\right)^{-2}\int_x^l T\,dx'\right]_{x=l}=0.$$

Оно выполнено для любой функции $T(x)$, поэтому нет необходимости учитывать его в дальнейшем.

Общее решение уравнения (9.119) найти трудно, но можно указать простое частное решение:

$$T(x)=1-\sqrt{-\frac{\lambda\rho}{\Lambda}}\,x. \tag{9.120}$$

Поскольку в данном случае линейное распределение температуры по длине ребра наиболее естественно с физической точки зрения, это распределение является оптимальным. Соответствующую оптимальную форму ребра можно найти с помощью соотношения (9.114):

$$D(x)=-\left[\left(\frac{dT}{dx}\right)^{-1}\int_x^l T(x')\,dx'\right]\frac{\alpha}{\Lambda}=\frac{\alpha}{\Lambda a}\left(\frac{ax^2}{2}+x-\frac{al^2}{2}-l\right), \tag{9.121}$$

где использовано обозначение $a=-\sqrt{\lambda\rho/\Lambda}$.

Неопределенный множитель λ находится из условия (9.112) постоянства массы ребра:

$$M=2\rho\int_0^L D(x)\,dx=-2\rho\left(\frac{\alpha l^3}{3\Lambda}+\frac{\alpha l^2}{2\Lambda a}\right).$$

Отсюда находим величину a:

$$a=-\frac{\Lambda}{\rho}\left(\frac{M\Lambda}{\rho\alpha l^2}+\frac{2l}{3}\right)^{-2}. \tag{9.122}$$

Из (9.122) легко получить значение множителя Лагранжа:

$$\lambda=-\frac{\Lambda}{\rho}\left(\frac{M\Lambda}{\rho\alpha l^2}+\frac{2l}{3}\right)^{-2}. \tag{9.123}$$

Для завершения решения задачи осталось определить оптимальное значение параметра l. Из всех физически осмысленных l (таких, что $D(x)>0$ при $0<x<l$) надо найти то значение, которое соответствует максимуму рассеиваемой энергии. Энергия, рассеиваемая в среду, определяется формулой (9.111):

$$Q=2\alpha\int_0^l T(x)\,dx.$$

Используя выражения (9.120) и (9.123), получаем после интегрирования

$$Q=2\alpha\left(l-\frac{1}{2}\,\frac{3\rho\alpha l^4}{3M\Lambda+2\rho\alpha l^3}\right).$$

Максимум Q получится при значении l, удовлетворяющем уравнению

$$\frac{\partial Q}{\partial l}=2\alpha\left[1-\frac{6\rho\alpha l^3}{3M\Lambda+2\rho\alpha l^3}+\frac{9\rho^2\alpha^2 l^6}{(3M\Lambda+2\rho\alpha l^3)^2}\right]=0,$$

откуда

$$l=\sqrt[3]{3M\Lambda/(\rho\alpha)}. \tag{9.124}$$

3. Оптимизация режима работы электродвигателя постоянного тока. Рассмотрим работу электродвигателя, обеспечивающего поворот платформы экскаватора на заданный угол. Режим перемещения должен быть выбран таким образом, чтобы в начальный и конечный моменты времени угловая скорость вращения платформы была равна нулю; кроме того, необходимо наложить ограничение на значение тока в двигателе, которое исключало бы возможность перегрева двигателя. При этих ограничениях требуется найти зависимости силы тока и угловой скорости вращения двигателя от времени, обеспечивающие оптимальный режим движения платформы экскаватора в смысле минимальных потерь энергии в электродвигателе.

Уравнение моментов для двигателя постоянного тока имеет вид

$$C_\Phi I\Phi_{\text{дв}}=J\frac{d\omega}{dt}+M_c, \tag{9.125}$$

где I — сила тока якоря электродвигателя, $\Phi_{\text{дв}}$ — результирующий магнитный поток, пронизывающий контур якоря; C_Φ — коэффициент пропорциональности; J — суммарный момент инерции электродвигателя и исполнительного механизма по отношению к валу электродвигателя; ω — угловая частота вращения якоря; M_c — момент сил сопротивления (будем считать $M_c=\text{const}$).

Для упрощения уравнения (9.125) перейдем к относительным единицам, приняв за единицы силы тока якоря, магнитного потока, скорости и моменты соответствующие номинальные значения этих физических величин:

$$i=\frac{I}{I_{\text{н}}},\ \ \Phi=\frac{\Phi_{\text{дв}}}{\Phi_{\text{н}}},\ \ \nu=\frac{\omega}{\omega_{\text{н}}},\ \ \mu=\frac{M_c}{M_{\text{н}}},\ \ \tau=\frac{t}{T_{\text{м}}},$$

где в качестве единицы времени введена постоянная времени, численно равная времени разгона двигателя до номинальной частоты вращения $\omega_{\text{н}}$ под действием номинального вращающего момента $M_{\text{н}}$:

$$T_{\text{м}}=J\omega_{\text{н}}/M_{\text{н}}.$$

Положим, кроме того, для простоты $C_\Phi=1$, запишем уравнение (9.125) в виде

$$i\Phi=\frac{d\nu}{d\tau}+\mu. \tag{9.126}$$

Для электродвигателей независимого возбуждения можно считать с хорошей степенью точности $\Phi=1$, тогда (9.126) примет вид

$$i=\frac{d\nu}{d\tau}+\mu, \quad \mu=\text{const}. \tag{9.127}$$

За единицу угла поворота возьмем угол, на который поворачивается вал электродвигателя за время $\Delta t=T_{\text{м}}$ при частоте вращения, равной $\omega_{\text{н}}$. В этих относительных единицах угол поворота платформы

$$\alpha=\int\limits_0^T \nu(\tau)\,d\tau. \tag{9.128}$$

Значение этого интеграла является фиксированным по условию задачи.

Введем функционал, задающий величину потерь энергии в якоре электродвигателя. За единицу этой величины примем потери энергии за время $t=T_{\text{м}}$ при силе тока $I=I_{\text{н}}$. Тогда в относительных единицах полная энергия Q, которая теряется в якоре, будет выражаться интегралом

$$Q=\int\limits_0^T i^2(\tau)\,d\tau. \tag{9.129}$$

Таким образом, строгая формулировка вариационной задачи оптимизации режима работы электропривода имеет следующий вид: требуется определить функции $\nu(\tau)$ и $i(\tau)$, удовлетворяющие уравнению (9.127) и обеспечивающие минимум интегралу (9.129) при фиксированном значении интеграла (9.128) и заданных граничных условиях $\nu(0)=\nu(T)=0$ (выражающих тот факт, что скорость платформы в начальный и конечный моменты времени должна быть равна нулю).

Возможна и другая постановка вариационной задачи. Найти функции $i(\tau)$, $\nu(\tau)$, удовлетворяющие тем же граничным условиям и уравнению (9.127) и обеспечивающие максимум интегралу (9.128) при заданном значении интеграла (9.129). Решение такой задачи определяет режим, соответствующий максимальному углу перемещения платформы при заданной величине потерь в якоре.

Подставим выражение (9.127) для i в интеграл (9.129). Первая задача на условный экстремум сведется к простой изопериметрической задаче: найти функцию $\nu(\tau)$, обеспечивающую минимум интегралу

$$Q=\int\limits_0^{T_2}\left(\frac{d\nu}{d\tau}+\mu\right)^2 d\tau$$

при заданном значении интеграла $\int\limits_0^T \nu d\tau = \alpha$ и заданных граничных условиях $\nu(0) = \nu(T) = 0$. С помощью метода множителей Лагранжа (см. § 9.5) эта задача сводится к безусловной вариационной задаче с функционалом

$$L = \int\limits_0^T \left[\left(\frac{d\nu}{d\tau} + \mu\right)^2 + \lambda\nu\right] d\tau.$$

Уравнение Эйлера для такой задачи имеет следующий вид:

$$2\frac{d^2\nu}{dt^2} - \lambda = 0. \tag{9.130}$$

Общее решение уравнения (9.130) записывается так:

$$\nu = C_1 + C_2\tau + \frac{\lambda}{4}\tau^2.$$

Произвольные постоянные C_1, C_2, λ определяются из условий $\nu(0) = 0$, $\nu(T) = 0$, $\int\limits_0^T \nu d\tau = \alpha$. Из первого условия следует $C_1 = 0$. Остальные два условия образуют систему уравнений относительно C_2 и λ:

$$0 = C_2 T + \frac{\lambda}{4} T^2,$$

$$\alpha = \frac{C_2 T^2}{2} + \frac{\lambda T^3}{12}.$$

Находим

$$C_2 = \frac{6\alpha}{T^2}, \quad \lambda = -\frac{24\alpha}{T^3}.$$

В результате получаем решение рассматриваемой вариационной задачи:

$$\nu(\tau) = \frac{6\alpha}{T^2}\left(\tau - \frac{\tau^2}{T}\right),$$

$$i(\tau) = \mu + \frac{6\alpha}{T^2} + \frac{12\alpha}{T^3}\tau.$$

Величину потерь энергии в найденном оптимальном режиме можно найти, подставив i в (9.129):

$$Q = \frac{12\alpha^3}{T^3} + \mu^2 T.$$

Для проверки того факта, что мы получили именно минимум функционала, используем условие Лежандра

$$L_{\dot{u}\dot{u}} = 2 > 0.$$

Более подробно с задачами оптимизации электрических приводов можно познакомиться в кн.: *Чистов В. П., Бондаренко В. И., Святославский В. А.* Оптимальное управление электрическими приводами. — М.: Энергия, 1968.

Глава 10
ОПТИМАЛЬНОЕ УПРАВЛЕНИЕ

Истоки теории оптимального управления лежат в вариационном исчислении. Поэтому между задачами оптимального управления и вариационного исчисления имеется тесная связь, но есть и существенные различия.

Задача оптимального управления заключается в оптимизации функционала при дифференциальных ограничениях специального вида. Одним из важнейших методов исследования и решения этой задачи основан на необходимом условии оптимальности в форме принципа максимума Л. С. Понтрягина. Формулировка, вывод и обсуждение принципа максимума и составляют основное содержание этой главы. Рассмотрены различные классы задач оптимального управления автономными и неавтономными системами, задачи с закрепленными и подвижными концами, с фиксированным и нефиксированным временем управления. Отмечена связь задачи оптимального управления с соответствующей задачей для дискретной системы. Подробно изучена задача оптимального быстродействия.

§ 10.1. НАЧАЛЬНЫЕ ПОНЯТИЯ ТЕОРИИ УПРАВЛЯЕМЫХ СИСТЕМ

1. Задача о мягкой посадке ракеты на Луну. Пусть в начальный момент времени $t=0$ ракета, которую мы отождествим с материальной точкой массы m, находится на высоте H над поверхностью Луны и имеет скорость V, направленную вертикально вниз. Задача заключается в выборе такого режима работы двигателя, чтобы в некоторый (не заданный заранее) момент времени $t=T$ ракета достигла поверхности Луны и при этом ее скорость была равна нулю (мягкая посадка). В системе координат, связанной с поверхностью Луны (ось Ox направлена вертикально вверх), уравнение движения ракеты имеет следующий вид:

$$m\frac{d^2x(t)}{dt^2}=-mg_л+F(t), \tag{10.1}$$

где $x=x(t)$ — высота ракеты над поверхностью Луны; $g_л$ — ускорение свободного падения на Луне, $F(t)$ — переменная сила двигателя. Очевидно, значение силы $F(t)$ не может быть сколь угодно большим; оно ограничено техническими возможностями двигателя, т. е.

$$0\leqslant F(t)\leqslant F_{max}=\text{const},\quad t\in[0, T]. \tag{10.2}$$

При этих условиях возможны несколько различных режимов работы двигателя, обеспечивающих мягкую посадку. Нужно выбрать тот, который с определенной точки зрения является наиболее выгодным. Примем в качестве критерия такой выгодности минимум расходуемого при посадке топлива. Обозначим расход топлива в единицу времени через $q(t)$. Учитывая, что сила $F(t)$ пропорциональна расходу топлива в единицу времени, можно записать: $q(t)=\alpha F(t)$, где α — коэффициент пропорциональности. Общий расход топлива за время T выражается интегралом $\int\limits_0^T q(t)\,\mathrm{d}t$, поэтому критерий выгодности (оптимальности) режима работы двигателя можно задать с помощью следующего функционала:

$$I(F)=\alpha\int\limits_0^T F(t)\,\mathrm{d}t. \tag{10.3}$$

Теперь сформулируем задачу о мягкой посадке: найти кусочно-непрерывную функцию $F(t)$, подчиненную неравенствам (10.2), при которой решение $x(t)$ уравнения (10.1) удовлетворяет при $t=0$ заданным начальным условиям

$$x(0)=H,\quad \left.\frac{\mathrm{d}x}{\mathrm{d}t}\right|_{t=0}=V,$$

а при некотором $t=T$ — условиям

$$x(t)|_{t=T}=0,\quad \left.\frac{\mathrm{d}x}{\mathrm{d}t}\right|_{t=T}=0,$$

причем такое, что функционал (10.3) принимает минимальное возможное значение на множестве таких функций F (считаем, что класс кусочно-непрерывных функций является наиболее широким классом технически реализуемых функций $F(t)$).

Сформулированная задача является простейшей задачей оптимального управления. Управлением служит функция F.

2. Основные понятия. Предположим, что поведение некоторого объекта может описать системой обыкновенных дифференциальных уравнений первого порядка

$$\frac{\mathrm{d}y_i}{\mathrm{d}t}=f_i(y_1,\ldots,y_n,\ u_1,\ldots,u_m),\quad 0\leqslant t\leqslant T,\quad i=1,\ldots,n. \tag{10.4}$$

Пусть момент времени $t_0=0$ является *начальным*, а момент $t_1=T$ — *конечным*. Правая часть системы (10.4) не зависит явно от t, поэтому, считая $t_0=0$ начальным моментом, мы не ограничиваем общности дальнейших рассуждений.

Функции f_i, $i=1, 2, \ldots, n$, от $n+m$ переменных, входящих в уравнения (10.4), предполагаются непрерывными на $\mathbf{R}^{n+m}$ и непрерывно дифференцируемыми по первым n переменным.

Функции $y_1, y_2, \ldots, y_n$ в (10.4) называют *фазовыми координатами* объекта. Они в совокупности образуют вектор-функцию $y(t)=$

$= (y_1(t), ..., y_n(t))^{\mathrm{T}}$, которая при каждом $t \in [0, T]$ задает в пространстве $\mathbf{R}^n$ *(фазовом пространстве) точку,* соответствующую состоянию объекта в момент t.

Функции $u_1, u_2, ..., u_m$, заданные на отрезке $[0, T]$, называют *управляющими переменными.* В совокупности они образуют вектор-функцию $u(t) = (u_1(t), ..., u_m(t))^{\mathrm{T}} \in \mathbf{R}^m$ при $t \in [0, T]$. В дальнейшем вектор-функцию u будем называть просто *управлением.* Предположим, что вектор-функция u может принимать произвольные значения из некоторого множества U, $U \in \mathbf{R}^m$, которое назовем *областью управления,* т. е. $u(t) \in U$ для всех $t \in [0, T]$. Функции u_i, $i = 1, ..., m$, в дальнейшем будем считать кусочно-непрерывными.

Произвольное управление u с кусочно-непрерывными компонентами, удовлетворяющее условию $u(t) \in U$ при всех $t \in [0, T]$, будем называть *допустимым управлением.*

Рассмотрим произвольное допустимое управление u и зафиксируем начальное условие $y(0) = y^{(0)}$, или в координатной форме

$$y_i(0) = y_i^{(0)}, \quad i = 1, 2, ..., n. \tag{10.5}$$

После того как выбрано $u = u(t)$, подставляя вектор-функцию $u(t)$ в правую часть системы (10.4), получаем задачу Коши для системы дифференциальных уравнений

$$\frac{dy_i}{dt} = f_i(y_1, ..., y_n, u_1(t), ..., u_m(t)), \quad i = 1, 2, ..., n, \tag{10.6}$$

с начальным условием (10.5). Функции $u_1, .., u_m$ кусочно-непрерывные, поэтому функции, входящие в правую часть равенства (10.6), вообще говоря, не являются непрерывными, а значит, нельзя воспользоваться теоремой о существовании и единственности решения задачи Коши, известной из курса дифференциальных уравнений. Укажем, каким образом понимать решение системы (10.4) при выбранном управлении u. Кусочно-непрерывно дифференцируемая* на отрезке $[0, T]$ вектор-функция y является *решением* системы (10.6) с начальным условием $y^{(0)}$ если при подстановке ее в равенства (10.6) они превращаются в тождества для всех точек $t \in [0, T]$, в которых все функции $u_1, ..., u_m$ непрерывны и, кроме того, верно начальное условие (10.5). В дальнейшем будем считать, что каждое допустимое управление u однозначно определяет решение y системы (10.4) с начальным условием (10.5)**. Это решение называют *фазовой траекторией,* соответствующей управлению u; при этом управление u определяет *поведение* системы во времени от $t = 0$ до $t = T$. Значение управления в точках разрыва u_i в определении решения системы уравнений не играет роли, поэтому будем далее предполагать, что все функции u_i непрерывны слева в точках разрыва.

---

* См. с. 237.

** Согласно теореме о существовании и единственности решения задачи Коши, при сделанных выше предположениях существует, и притом единственное, решение только в малой окрестности каждой точки $t \in [0, T]$, в которой все функции $u_1, \dots, u_m$ непрерывны (но не на всем отрезке $[0, t)$.

Будем считать, что управление $u(t)$, заданное на отрезке $[0, T]$, *переводит систему* из состояния $y^{(0)}=\left(y_1^{(0)}, y_2^{(0)},\ldots, y_n^{(0)}\right)^{\tau}$ в состояние $y^{(1)}=\left(y_i^{(1)}, y_2^{(1)},\ldots, y_n^{(1)}\right)^{\tau}$, если соответствующее этому управлению решение y удовлетворяет начальным условиям (10.5), и в момент $t=T$ имеет значение $y(T)=\left(y_1^{(1)},\ldots, y_n^{(1)}\right)^{\tau}$, т. е. удовлетворяет условиям

$$y_i(T)=y_i^{(1)},\quad i=1,\ldots, n.$$

3. Постановка задачи оптимального управления с закрепленными концами. Сформулируем две вспомогательные задачи, которые условимся называть задачами управления (в отличие от задач оптимального управления).

I. *Задача с фиксированной продолжительностью управления.* В фазовом пространстве заданы две точки $A=\left(y_1^{(0)}, y_2^{(0)},\ldots, y_n^{(0)}\right)$ и $B=\left(y_1^{(1)}, y_2^{(1)},\ldots, y_n^{(1)}\right)^{\tau}$. Среди всех допустимых управлений найти такое управление, которое переводит систему (10.4), находящуюся в начальный момент $t=0$ в состоянии A, к заданному моменту $t=T$ в состояние B.

II. *Задача с нефиксированной продолжительностью управления.* Среди всех допустимых управлений, заданных на $[0, \tau]$ при $\tau>0$, найти такое управление, которое переводит систему (10.4), находящуюся в начальный момент $t=0$ в состоянии A, в состояние B к некоторому заранее не заданному моменту времени T, $0\leqslant T\leqslant\tau$. При этом момент времени T также подлежит определению.

Возникает важный вопрос о том, существует ли управление u, переводящие систему из состояния A в состояние B. Если установлено, что такого управления не существует, то решать задачу управления бессмысленно. В общем случае установить существование решения задачи управления трудно. Однако если ограничиться рассмотрением линейных систем, то можно дать ответ на поставленный вопрос. Введем необходимые определения. Систему дифференциальных уравнений вида

$$\frac{dy}{dt}=Ay+Bu, \tag{10.7}$$

где $dy/dt=(dy_1/dt, \ldots, dy_n/dt)^{\tau}$, A и B — числовые матрицы размера $n\times n$ и $n\times m$ соответственно, называют *линейной системой управления.* Эта система *полностью управляема,* если для любой пары точек $y^{(0)}, y^{(1)}\in\mathbf{R}^n$ существует кусочно-непрерывная вектор-функция $u(t)=(u_1(t), \ldots, u_m(t))^{\tau}\in\mathbf{R}^m$, переводящая систему (10.7) из состояния $y^{(0)}$ в состояние $y^{(1)}$ за некоторое время T. Имеет место следующее утверждение (см.: *Ли Э. Б., Маркус Л.* Основы теории оптимального управления. — М.: Наука, 1972).

Система вида (10.7) *полностью управляема тогда и только тогда, когда ранг $(n\times nm)$-матрицы, составленной из матриц B, AB, A^2B, ..., $A^{n-1}B$, равен n.*

Если ранг указанной матрицы меньше n, то существуют пары точек $y^{(0)}$ и $y^{(1)}$, для которых невозможно подобрать управление u, переводящее систему из одной точки в другую за конечный промежуток времени.

Говоря о системе (10.4) как об управляемой системе, обычно считают, что допустимая кусочно-непрерывная управляющая вектор-функция, переводящая систему из заданного начального состояния в конечное, существует.

Для большинства прикладных задач управления, соответствующих реальным объектам, существует не единственный набор функций $u_1, \ldots, u_m$, являющихся решением задачи I или задачи II. Поэтому возникает возможность выбора из всех решений задачи управления такого решения, которое является в каком-то смысле наиболе выгодным (оптимальным). В этом случае приходим к задаче оптимального управления. В ней требуется, чтобы искомое управление не только решало задачу I или II, но, кроме того, минимизировало (максимизировало) некоторый интегральный функционал, который называется *критерием оптимальности (критерием качества)* и определяется равенством

$$I(u)=\int_0^T f_0(y_1(t),\ldots, y_n(t),\ u_1(t),\ldots, u_m(t))\,\mathrm{d}t, \tag{10.8}$$

где $f_0(y_1, \ldots, y_n, u_1, \ldots, u_m)$ — фиксированная непрерывная на $\mathbf{R}^{n+m}$ функция, непрерывно дифференцируемая по первым n переменным. Верхний предел интегрирования в зависимости от типа задачи может быть заранее фиксированным (задача I) или нефиксированным (задача II). Во втором случае целевой функционал зависит не только от u, но и от T, и поэтому было бы естественно писать $I(u, T)$. Однако для краткости T будем опускать, специально оговаривая, что имеется в виду задача с нефиксированным временем управления.

Таким образом, *задача оптимального управления* заключается в минимизации функционала (10.8) на подмножестве множества допустимых управлений, переводящих систему (10.4) из заданной точки A в заданную точку B. Управление $u^*=(u^*_1(t), \ldots, u^*_m(t))^{\mathrm{T}}$, являющееся решением этой задачи, называют *оптимальным управлением*, а соответствующую ему согласно (10.4) траекторию $y^*=(y^*_1(t), \ldots, y^*_n(t))^{\mathrm{T}}$ — *оптимальной траекторией*.

Отметим важный частный случай поставленной задачи оптимального управления. В *задаче оптимального быстродействия* функция f_0 тождественно равна единице. В этом случае функционал (10.8) имеет вид $I(u)=T$ и оптимизация состоит в нахождении допустимого управления u, минимизирующего время перехода системы из состояния A в состояние B (это случай задачи II).

Как правило, на искомое оптимальное управление накладывают ограничения, связанные с его технической реализацией, например

$$|u_i(t)|\leqslant \alpha_i,\ \ i=1,\ldots, m,\ \text{при}\ t\in[0, T].$$

В общем случае значения вектор-функции $u(t)=(u_1(t), \ldots, u_m(t))^{\mathrm{T}}$ должны принадлежать замкнутой (незамкнутой) области управления $U \subset \mathbf{R}^m$. Замкнутость области управления является причиной того, что задачи оптимального управления часто называют неклассическими вариационными задачами. Отличие сформулированных задач оптимального управления от задачи Лагранжа в понтрягинской форме (см. п. 5 § 9.6) главным образом состоит в дополнительном ограничении вида $u(t) \in U$ при всех $t \in [0, T]$. Если множество U замкнутое (что имеет место в большинстве прикладных задач управления), то методы вариационного исчисления «не работают» (в частности, неприменима теорема 9.9). Для задач оптимального управления приходится разрабатывать новые методы решения. Один из таких методов основан на применении принципа максимума Понтрягина, который можно рассматривать как обобщение метода множителей Лагранжа в классическом вариационном исчислении. Используют и другие методы, например метод динамического программирования.

4. Простейшие свойства допустимых управлений и соответствующих траекторий. Пусть допустимое управление $u(t)=(u_1(t), u_2(t), \ldots, u_m(t))^{\mathrm{T}}$ переводит систему (10.4) из начального состояния, соответствующего фазовой точке A в момент времени $t=0$, в конечное состояние, соответствующее фазовой точке B в момент времени $t=T$, и функционал (10.8) принимает на этом управлении значение I_0. Функция f из (10.4) и функция f_0 из (10.8) явно не зависят от t, поэтому имеет место следующее утверждение.

При любом $h \in \mathbf{R}$ управление вида $(u_1(t+h), \ldots, u_m(t+h))^{\mathrm{T}}$, $-h \leqslant t \leqslant T-h$, является допустимым, переводит систему из состояния, соответствующего фазовой точке A в момент $t=-h$, в состояние, соответствующее фазовой точке B в момент $t=T-h$, и функционал (10.8) *принимает на нем значение I_0.* Таким образом в рассматриваемой задаче возможно «смещение» начального момента времени в любую точку временно́й оси.

Справедливо следующее свойство. *Пусть имеется набор точек $A_1, A_2, \ldots, A_k$ фазового пространства и для каждого $i=1, 2, \ldots, k-1$ существует допустимое управление $u^{(i)}(t)$, которое переводит систему* (10.4) *из состояния, соответствующего точке A_i, в состояние, соответствующее точке A_{i+1}, за время Δt_i, $\sum\limits_{i=1}^{k-1} \Delta t_i = T$, и функционал* (10.8) *принимает на нем значение I_i. Тогда существует допустимое управление $u(t)$, $0 \leqslant t \leqslant T$, переводящее систему* (10.4) *из состояния, соответствующего точке A_1, в состояние, соответствующее точке A_k, причем на этом управлении целевой функционал принимает значение, равное $I_1+I_2+\ldots+I_k$.* В самом деле, поскольку управление можно «сдвигать» по оси времени, будем считать, что отрезки, на которых заданы управления $u^{(1)}(t)$, $u^{(2)}(t)$, ..., $u^{(k-1)}(t)$, «примыкают» друг к другу, т. е. функции $u^{(i)}(t)$ заданы на отрезках $[t_i, t_{i+1}]$, где $0=t_1<t_2<\ldots<t_k=T$. Обозначим через $u(t)$ управление, заданное на отрезке $[0, T]$ и совпадающее на каждом частичном

промежутке $[t_i, t_{i+1}]$ с управлением $u^{(i)}(t)$. (рис. 10.1). Тогда управление $u(t)$ является допустимым и переводит систему из точки A_1 в точку A_k. Кроме того, значение целевого функционала на этом управлении в силу аддитивности интеграла (10.8) равно $I_1+I_2+\ldots$ $\ldots+I_k$.

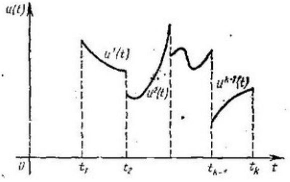

Рис. 10.1

Принцип оптимальности Беллмана является обобщением соответствующего свойства для дискретных систем управления (см. гл. 4). Это свойство состоит в ом, что всякий «кусок» оптимального управления, рассматриваемого на частичном отрезке, также является оптимальным управлением на этом частичном отрезке. *Пусть оптимальное управление $u(t)$, определенное на отрезке $[0, T]$, переводит систему из начального состояния A в конечное состояние B и $y(t)$ — соответствующая ему оптимальная траектория. Выберем две произвольные точки на временной оси $t', t'' \in$ $\in (0, T)$ вида $t' < t''$. Управление $u(t)$, рассматриваемое на отрезке $[t', t'']$, является оптимальным управлением, переводящим систему из состояния $y(t')$ в состояние $y(t'')$, а сответствующий «кусок» оптимальной траектории является оптимальной траекторией для этого перехода.* Для доказательства справедливости этого утверждения предположим противное. Пусть управление $u(t)$, рассматриваемое на отрезке $[t', t'']$, не является оптимальным. Тогда существует управление $u'(t)$, переводящее систему из положения $y(t')$ в положение $y(t'')$ по траектории, отличной от соответствующей части траектории $y(t)$, которая соединяет точки A и B (рис. 10.2) и доставляет целевому функционалу значение I_2', меньшее, чем значение I_2 этого же функционала на управлении $u(t)$ в пределах от t' до t''. В противном случае «новое» управление, переводящее систему из A в B, можно составить из трех

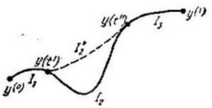

Рис. 10.2

«кусков»: исходного управления, переводящего систему из состояния A в точку $y(t')$, управления $u'(t)$, осуществляющего переход из $y(t')$ в $y(t'')$ и управления $u(t)$, рассматриваемого на отрезке $[t'', T]$ и переводящего систему из $y(t'')$ в точку B. На этом «новом» управлении функционал (10.8) принимает значение $I_1+I_2'+I_3$ меньшее чем $I_1+I_2+I_3$, что противоречит условию оптимальности исходного управления $u(t)$, $0 \leqslant t \leqslant T$.

§ 10.2. ПРИНЦИП МАКСИМУМА ПОНТРЯГИНА ДЛЯ ЗАДАЧ С ЗАКРЕПЛЕННЫМИ КОНЦАМИ

1. Принцип максимума для задачи с нефиксированным временем управления. Рассмотрим задачу оптимального управления системой (в векторной форме)

$$\frac{dy}{dt}=f(y, u), \tag{10.9}$$

где $dy/dt=(dy_1/dt, \ldots, dy_n/dt)^{\mathrm{T}}$, $f=(f_1, \ldots, f_n)^{\mathrm{T}}$,

$$y=y(t)=(y_1(t), \ldots, y_n(t))^{\mathrm{T}}, \quad u=u(t)=(u_1(t), \ldots, u_m(t)^{\mathrm{T}}$$

с нефиксированным временем управления T, начальным состоянием $y^{(0)}$, конечным состоянием $y^{(1)}$, критерием качества

$$I(u)=\int_0^T f_0(y(t)),\ u(t))\, dt \to \min \tag{10.10}$$

и областью управления $U\subset \mathbf{R}^m$. При этом для функций y, u, f будем считать выполненными все предположения предыдущего параграфа.

Введем функцию $H(y, \psi, u)$ типа $\mathbf{R}^{2n+1}\times U\to\mathbf{R}$, называемую *функцией Понтрягина*, которая играет в теории оптимального управления роль, аналогичную роли функции (функционала) Лагранжа в вариационном исчислении:

$$H(y, \boldsymbol{\psi}, u)=\sum_{i=0}^{n} \psi_i f_i(y, u), \tag{10.11}$$

где $\boldsymbol{\psi}=(\psi_0, \psi_1, \ldots, \psi_n)^{\mathrm{T}}$. Обозначим $\psi=(\psi_1, \psi_2, \ldots, \psi_n)^{\mathrm{T}}$. Справедлива [20] следующая теорема.

Теорема 10.1. *Пусть вектор-функция $u^*(t)$ является оптимальным управлением, а вектор-функция $y^*(t)$ — соответствующей оптимальной траекторией в сформулированной задаче оптимального управления. Тогда существуют непрерывная вектор-функция $\psi^*(t)=(\psi_1^*(t), \ldots, \psi_n^*(t))^{\mathrm{T}}$ и число $\psi_0^*\leqslant 0$ такие, что:*

1) *вектор-функция вида $\boldsymbol{\psi}^*=(\psi_0^*,\ \psi^*(t))^{\mathrm{T}}$, $0\leqslant t\leqslant T$, является ненулевой;*

2) *вектор-функция $\psi^*(t)$ является решением системы дифференциальных уравнений*

$$\frac{d\psi_i}{dt}=-\frac{\partial H(y, \psi, u)}{\partial y_i}\bigg|_{\substack{u=u^*(t)\\ y=y^*(t)}}=-\sum_{j=0}^{n} \psi_j \frac{\partial f_j(y^*, u^*)}{\partial y_i}, \quad i=1, 2, \ldots, n; \tag{10.12}$$

3) *при каждом $t\in[0, T]$ функция $H(y^*(t),\ \boldsymbol{\psi}^*(t),\ u)$ векторной переменной $u=(u_1, u_2, \ldots, u_m)^{\mathrm{T}}$ достигает максимума на множестве U при $u=u^*(t)$*, т. е.

$$H(y^*(t), \boldsymbol{\psi}^*(t),\ u^*(t))=\max_{u\in U} H(y^*(t), \boldsymbol{\psi}^*(t),\ u),\ 0\leqslant t\leqslant T; \tag{10.13}$$

4) *при каждом* $t \in [0, T]$ *выполняется равенство*

$$H(y^*(t), \psi^*(t), u^*(t)) = 0. \qquad (10.14)$$

В формулировке теоремы 10.1 главным является условие максимума (10.13). Именно поэтому теорему и называют *принципом максимума.* Условие 1) теоремы исключает случай $H \equiv 0$ и делает равенство (10.13) содержательным.

С помощью функции Понтрягина правую часть системы (10.9) *можно представить в виде* $f_i(y, u) = \partial H(y, \psi, u)/\partial\psi_i$, $i = 1, 2, ..., n$. Следовательно, исходную систему (10.9) вместе с линейной системой (10.12), которую называют *сопряженной системой,* часто записывают в симметричном виде:

$$\frac{dy_i}{dt} = \frac{\partial H(y, \psi, u)}{\partial \psi_i}, \qquad i = 1, 2, ..., n.$$

$$\frac{d\psi_i}{dt} = -\frac{\partial H(y, \psi, u)}{\partial y_i},$$

Если в конкретной задаче удается показать, что $\psi_0^* \neq 0$, то всегда можно положить $\psi_0^* = -1$, так как равенства (10.12)—(10.14) не изменяются при умножении их на произвольное положительное число. В общем случае возможен и случай, когда $\psi_0^* = 0$. Тогда функция Понтрягина не включает f_0, а значит, необходимые условия теоремы 10.1 не содержат информации о целевом функционале $I(u)$. Следовательно, заменяя исходный функционал другим, получаем для новой задачи те же самые условия оптимальности в форме принципа максимума, что и для исходной. Задачи оптимального управления подобного типа называют *особыми* (или *вырожденными*).

Равенство (10.14) служит в основном для определения конечного момента времени T и включено в число необходимых условий оптимальности потому, что здесь рассматривается задача с нефиксированным временем управления. Если функции $u^*(t)$ и $\psi^*(t)$ удовлетворяют равенствам (10.12), (10.13), то $H(y^*(t), \psi^*(t), u^*(t)) = \text{const}$ и поэтому при решении конкретной задачи равенство (10.14) достаточно проверить [20] только для одного произвольного фиксированного момента $t \in [0, T]$.

Принцип максимума дает только необходимые условия оптимальности. Поэтому если некоторое допустимое управление удовлетворяет этому принципу, то оно не обязательно является оптимальным. Нередко, желая подчеркнуть это обстоятельство, функцию $u^*(t)$, удовлетворяющую всем условиям теоремы 10.1, называют *экстремалью Понтрягина.* Лишь для некоторых классов задач (например, линейных задач оптимального быстродействия) принцип максимума является как необходимым, так и достаточным условием оптимальности (см. § 10.4).

В прикладных задачах принцип максимума нередко позволяет однозначно определить оптимальное управление. Так, если заранее известно, что оптимальное управление существует и, кроме того, найдено единственное допустимое управление, удовлетворяющее этому принципу, то это единственное управление является оптимальным.

2. Принцип максимума для задачи с фиксированным временем управления. Обратимся к задаче оптимального управления, отличающейся от задачи, рассмотренной в предыдущем пункте, только тем, что здесь время управления T будем считать фиксированным, т. е. заданным с самого начала. Принцип максимума для этой задачи нетрудно вывести из теоремы 10.1. Для этого введем дополнительную переменную y_{n+1} по правилу

$$\frac{dy_{n+1}}{dt}=1, \; y_{n+1}(0)=0. \tag{10.15}$$

Очевидно, что $y_{n+1}(t)\equiv t$, т. е. включение дополнительного уравнения с начальным условием (10.15) в систему (10.9) эквивалентно введению t в число фазовых координат управляемой системы.

Учитывая дополнительную переменную, переформируем задачу с фиксированным временем управления (в новой формулировке время управления уже не фиксировано). В фазовом пространстве $\mathbf{R}^{n+1}$ имеется «расширенная» система управления

$$\begin{cases} \dfrac{dy}{dt}=f(y,u), & (10.16)\\[2ex] \dfrac{dy_{n+1}}{dt}=1. & (10.17)\end{cases}$$

Задача состоит в выборе такого допустимого управления, при котором система (10.16)—(10.17) из точки $(y^{(0)}, 0)\in\mathbf{R}^{n+1}$ перешла бы в точку $(y^{(1)}, T)\in\mathbf{R}^{n+1}$ и при этом был бы минимизирован критерий качества (10.10). Время управления здесь можно считать нефиксированным, так как переход системы в фазовую точку $(y^{(1)}, T)$ уже обеспечивает заданное значение T времени управления.

Запишем функцию Понтрягина для сформулированной задачи

$$\widehat{H}=\sum_{i=0}^{n}\psi_i f_i(y,u)+\psi_{n+1}\cdot 1=H(y,\psi,u)+\psi_{n+1},$$

где $H(y, \psi, u)$ имеет вид (10.11), и применим теорему 10.1. Равенства (10.13) и (10.14) принимают соответственно вид

$$H(y^*(t),\psi^*(t),u^*(t))+\psi^*_{n+1}(t)=$$
$$=\max_{u\in U}[H(y^*(t),\psi^*(t),u)+\psi^*_{n+1}(t)], \quad 0\leqslant t\leqslant T, \tag{10.18}$$

$$H(y^*(t),\psi^*(t),u^*(t))+\psi^*_{n+1}(t)=0, \quad 0\leqslant t\leqslant T, \tag{10.19}$$

где $\psi_0^*\leqslant 0$ и $(\psi^*(t), \psi^*_{n+1}(t))^{\mathrm{T}}$ — решение сопряженной системы

$$\frac{d\psi_i}{dt}=-\sum_{j=0}^{n}\psi_j\frac{\partial f_j(y^*,u^*)}{\partial y_i}, \quad i=1,2,\ldots,n+1. \tag{10.20}$$

Если все величины ψ_0^*, $\psi_1^*(t)$, ..., $\psi_n^*(t)$, $\psi^*_{n+1}(t)$ тождественно равны нулю, то из (10.19) получим тождество $\psi^*_{n+1}(t)\equiv 0$. Но это невозможно, так как, согласно теореме 10.1, вектор-функция $(\psi_0^*, \psi_1^*, ..., \psi_n^*, \psi^*_{n+1})^{\mathrm{T}}$ является ненулевой. Следовательно, $(\psi_0^*, \psi(t))^{\mathrm{T}}$ — ненулевая вектор-функция.

Слагаемое $\psi^*_{n+1}(t)$ в равенстве (10.18) сокращается, а значит, функция $\psi^*_{n+1}(t)$ при формулировке необходимого условия оптимальности не нужна. При этом как бессодержательные следует отбросить $(n+1)$-е уравнение системы (10.20) и условие (10.19). В результате имеем следующее утверждение.

Теорема 10.2. *Пусть вектор-функция $u^*(t)$ является оптимальным управлением, а $y^*(t)$ — соответствующая ему оптимальная траектория в задаче оптимального управления с фиксированным временем управления T. Тогда существует непрерывная вектор-функция $\psi^*(t)=(\psi_1^*(t), ..., \psi_n^*(t))^{\mathrm{T}}$ и число $\psi_0^*\leqslant 0$ такие, что выполняются условия* 1)—3) *теоремы* 10.1.

Единственное отличие условий оптимальности теоремы 10.2 от условий оптимальности теоремы 10.1 состоит в отсутствии равенства (10.14). В рассматриваемой здесь задаче время T задано, поэтому становится лишним условие для его определения.

3. Схема применения принципа максимума. Сначала рассмотрим задачу оптимального управления с фиксированным временем управления.

Обычно нахождение экстремали Понтрягина начинают с центрального условия принципа максимума

$$H(y, \psi, u)\to\max_{u\in U}, \tag{10.21}$$

из которого при каждом фиксированном наборе y, ψ определяют управление u, являющееся функцией параметров y и ψ, т. е.

$$u=u(y, \psi). \tag{10.22}$$

В общем случае это сделать трудно, однако для некоторых классов задач управления функцию (10.22) удается записать в явном виде.

Пусть, например,

$$f_i(y, u)=f_i(y)+\sum_{k=1}^{m} f_{ik}(y)u_k, \quad i=0, 1,..., n,$$

$$U=\{u\in \mathrm{R}^m \mid a_k\leqslant u_k\leqslant b_k, \quad k=1, 2,..., m\},$$

где a_k и b_k — заданные числа. В этом случае функция Понтрягина имеет вид

$$H=\sum_{i=0}^{n}\psi_i f_i(y)+\sum_{k=1}^{m}\left(\sum_{i=0}^{n}\psi_i f_{ik}(y)\right)u_k$$

и достигает максимума (благодаря линейности по u) только в граничных точках множества U, а именно при

$$u_k = \begin{cases} b_k, \text{ если } \sum_{i=0}^{n} \psi_i f_{ik}(y) > 0, \\ a_k, \text{ если } \sum_{i=0}^{n} \psi_i f_{ik}(y) < 0. \end{cases}$$

Предположим, что функция (10.22) найдена. Подставим ее в исходную и сопряженную ей системы. Имеем

$$\frac{\mathrm{d}y}{\mathrm{d}t} = f(y, u(y, \psi)),$$

$$\frac{\mathrm{d}\psi}{\mathrm{d}t} = -\frac{\partial H(y, \psi, u(y, \psi))}{\partial y}, \quad 0 \leqslant t \leqslant T, \tag{10.23}$$

где

$$\mathrm{d}\psi/\mathrm{d}t = (\mathrm{d}\psi_1/\mathrm{d}t, \mathrm{d}\psi_2/\mathrm{d}t, \ldots, \mathrm{d}\psi_n/\mathrm{d}t)^{\mathrm{T}},$$

$$\partial H/\partial y = (\partial H/\partial y_1, \partial H/\partial y_2, \ldots, \partial H/\partial y_n)^{\mathrm{T}}.$$

Получена система $2n$ дифференциальных уравнений относительно неизвестных функций $y_1, y_2, \ldots, y_n, \psi_1, \psi_2, \ldots, \psi_n$, общее решение которой содержит $2n$ произвольные постоянные. После нахождения общего решения эти произвольные постоянные определяют из $2n$ краевых условий:

$$y(0) = y^{(0)}, \quad y(T) = y^{(1)}. \tag{10.24}$$

В результате получают некоторые функции $\hat{y}(t)$ и $\hat{\psi}(t)$. Для определения ψ_0 достаточно, как отмечено выше, рассмотреть два случая $\psi_0 = 0$ и $\psi_0 = -1$ и установить, какой из них имеет место в действительности. При этом следует учитывать, что вектор-функция $(\hat{\psi}_0, \hat{\psi}(t))^{\mathrm{T}}$ должна быть ненулевой.

Пусть, наконец, функции $\hat{y}(t)$, $\hat{\psi}(t)$ найдены. Подставим их в (10.22):

$$\hat{u} = u(\hat{y}(t), \hat{\psi}(t)), \; 0 \leqslant t \leqslant T,$$

и предположим, что функция $\hat{u}$ оказалась кусочно-непрерывной, причем $\hat{u} = u(\hat{y}(t), \hat{\psi}(t)) \in U$ при каждом $t \in [0, T]$. Тогда эта функция является экстремалью Понтрягина, а значит, входит в число управлений, «подозрительных» на оптимальные. Если известно, что решение задачи оптимального управления существует и доказана единственность экстремали Понтрягина $\hat{u}$, то она является искомым оптимальным управлением.

Итак, применение принципа максимума сводится к задаче использования условия максимума (10.21) и решения системы дифференциальных уравнений (10.23) с краевыми условиями (10.24). Это краевая задача принципа максимума. В тех случаях, когда аналитически решить ее не удается, используют различные численные методы [8].

Схема применения принципа максимума для решения задачи оптимального управления с нефиксированным временем управления T та же, что и в задаче с фиксированным временем. При этом время T определяют из равенства (10.14) при $t=T$.

4. Решение задачи о мягкой посадке на Луну. Полагая $x=y_1$ $dx/dt=y_2$, $F(t)=u(t)$, $F_{\max}=r$, $g_л=g$, задачу о мягкой посадке, сформулированную в § 10.1, запишем в следующем виде: найти кусочно-непрерывную функцию $u(t)$, удовлетворяющую неравенствам $0\leqslant u(t)\leqslant r$, $0\leqslant t\leqslant T$, для которой решение $y=(y_1(t),\ y_2(t))^{\mathrm{T}}$ системы

$$\frac{dy_1}{dt}=y_2,$$
$$\frac{dy_2}{dt}=-g+\frac{u(t)}{m} \qquad (10.25)$$

удовлетворяет краевым условиям

$$y_1(0)=H,\ \ y_1(T)=0;\ \ y_2(0)=V,\ \ y_2(T)=0, \qquad (10.26)$$

причем функционал

$$I(u)=\alpha\int_0^T u(t)\,dt$$

достигает своего наименьшего значения. Конечный момент времени T заранее не задан и также подлежит определению.

Составим функцию Понтрягина для этой задачи:

$$H(y,\psi,\ u)=\alpha\psi_0 u+\psi_1 y_2+\psi_2\left(-g+\frac{u}{m}\right) \qquad (10.27)$$

и запишем сопряженную систему (10.12):

$$\frac{d\psi_1}{dt}=0,\quad \frac{d\psi_2}{dt}=-\psi_1.$$

Общее решение этой системы имеет вид

$$\psi_1^*=\text{const},\ \ \psi_2^*=C_1t+C_2,\ \ C_1=-\psi_1^*,$$

где C_1, C_2 — постоянные. Функция H (10.27) линейная по u, поэтому условие максимума (10.13) для нее выполняется только при

$$u^*(t)=\begin{cases}0, \text{ если } \alpha\psi_0^*+\dfrac{\psi_2^*}{m}<0,\\ r, \text{ если } \alpha\psi_0^*+\dfrac{\psi_2^*}{m}>0.\end{cases} \qquad (10.28)$$

Это, в частности, означает, что оптимальным управлением может быть только кусочно-постоянная функция, принимающая значение 0 или r. Исследуем уравнения движения ракеты в каждом из этих случаев.

Если $u^*(t) \equiv 0$ при $t \in [\tau, \tau_1] \subset [0, T]$ (что соответствует режиму свободного падения) система уравнений (10.25) имеет решение

$$y_1(t) = y_1(\tau) + y_2(\tau)(t-\tau) - \frac{g(t-\tau)^2}{2},$$

$$y_2(t) = y_2(\tau) - g(t-\tau). \tag{10.29}$$

Если же $u^*(t) \equiv r$ при $t \in [\tau, \tau_1]$ (что соответствует режиму максимального торможения), то аналогично получаем

$$y_1(t) = y_1(\tau) + y_2(\tau)(t-\tau) + \frac{(-g + r_0/m)(t-\tau)^2}{2},$$

$$y_2(t) = y_2(\tau) + (-g + r/m)(t-\tau). \tag{10.30}$$

Из физических соображений ясно, что на каком-то завершающем участке полета, соответствующем отрезку времени $[\tau\ T]$, $\tau < T$, двигатель должен работать, причем, согласно (10.28), в режиме максимального торможения. Этот факт можно установить и аналитически. Действительно, предположим обратное и, подставив $t = \tau_1 = T$ в (10.29), с учетом (10.26) найдем

$$0 = y_1(\tau) + y_2(\tau)(T-\tau) - \frac{g(T-\tau)^2}{2},$$

$$0 = y_2(\tau) - g(T-\tau).$$

Если исключить из этой системы T, то получим уравнение $y_1(\tau) + y_2^2(\tau) = g = 0$, что возможно лишь при $y_1(\tau) < 0$. Полученное неравенство соответствует положению ракеты в момент времени $t = \tau$ под поверхностью Луны, что невозможно.

Учитывая установленный факт, а также то обстоятельство, что в силу линейности функции $\alpha\psi_0^* + \psi_2^*/m = \alpha\psi_0^* + (C_1 t + C_2)/m$ по t управление u^* (10.28) не может иметь более одной точки разрыва (точки переключения), имеем

$$u^*(t) = \begin{cases} 0, \text{ если } t \in [0, \tau], \\ r, \text{ если } t \in (\tau, T), \end{cases} \tag{10.31}$$

при некотором $\tau \in [0, T]$. Здесь и далее τ — это момент переключения управления на режим максимального торможения.

Рассмотрим движение ракеты на завершающем участке, соответствующем отрезку времени $[\tau, T]$. Подставляя $t = \tau_1 = T$ в (10.30) и учитывая (10.26), найдем

$$0 = y_1(\tau) + y_2(\tau)s + \frac{(-g + r/m)s^2}{2},$$

$$0 = y_2(\tau) + (-g + r/m)s,$$

где $s = T - \tau$ — продолжительность режима максимального торможения. Исключая из полученной системы равенств параметр s, придем к соотношению

$$y_1(\tau)=\frac{y_2^2(\tau)}{g-r/m}, \tag{10.32}$$

где $y_1(\tau)$ — высота, $y_2(\tau)$ — скорость в момент переключения управления. Если $H=V^2/(g-r/m)$, то мягкая посадка может быть обеспечена только управлением $u^*(t)\equiv r$, $0\leqslant t\leqslant T$, т. е. в режиме максимального торможения с $t=0$ до $t=T$. В случае $H<V^2/(g-r/m)$ мягкая посадка невозможна, так как добиться нулевой скорости на поверхности Луны не удастся даже с помощью режима максимального торможения в пределах $0\leqslant t\leqslant T$. Если же $H>V^2/(g-r/m)$, то при $0\leqslant t\leqslant\tau$ осуществляется режим свободного падения ($u^*(t)\equiv 0$), а затем при $\tau<t\leqslant T$ — режим максимального торможения ($u^*(t)\equiv$ $\equiv r$). Заметим, что момент переключения τ может быть найден из (10.29) при $\tau=0$, $t=\tau$ и (10.32).

Обозначим через $y^*(t)$ решение системы (10.25), соответствующее управлению (10.31) и удовлетворяющее граничным условиям (10.26). Запишем применительно к функции Понтрягина (10.27) равенство (10.14) при $t=0$ и $t=\tau$:

$$H(y^*, \psi^*, u^*)|_{t=0}=\psi_1^* V-\psi_2^*(0)\,g=0, \tag{10.33}$$

$$H(y^*, \psi^*, u^*)|_{t=\tau}=\psi_1^* y_2(\tau)-\psi_2^*(\tau)\,g=0. \tag{10.34}$$

Согласно (10.28), в момент переключения τ выполняется равенство

$$\alpha\psi_0^*+\frac{\psi_2^*(\tau)}{m}=0. \tag{10.35}$$

Предполагая, что $\psi_0^*=0$, из (10.35) получим, что $\psi_2^*(\tau)=0$. С другой стороны, при $\psi_0^*=0$ из равенства $H(y^*, \psi^*, u^*)\,|_{t=T}=0$, следует, что $\psi_2^*(T)=0$. Так как $\psi_2^*(t)=C_1t+C_2$, имеем $\psi_2^*(t)\equiv 0$. Тогда из (10.33) получаем $\psi_1^*=0$. Окончательно $(\psi_0^*, \psi_1^*, \psi_2^*(t))^T\equiv 0$, что, согласно условию 1) теоремы 10.1, невозможно. Следовательно, $\psi_0^*<0$ и можно положить $\psi_0^*=-1$.

Используя обозначение $\psi_2^*=C_1t+C_2$ и равенства (10.35), равенства (10.33) и (10.34) запишем в виде

$$\psi_1^* V-C_2 g=0,$$

$$\psi_1^*(V-g\tau)-\alpha m g=0.$$

Отсюда находим

$$\psi_1^*=\frac{\alpha m g}{V-g\tau};\quad C_2=\frac{\alpha m V}{V-g\tau}.$$

Поскольку $C_1=-\psi_1^*$, окончательно получаем

$$\psi_2^*(t)=-\frac{\alpha m g t}{V-g\tau}+\frac{\alpha m V}{V-g\tau}.$$

Как отмечалось ранее, проверив равенство $H(y^*(t), \psi^*(t), u^*(t))=0$ для одного значения $t\in[0, T]$, нет необходимости убеждаться в его справедливости при остальных значениях $t\in[0, T]$. Значит, равенство (10.14) для функций $y^*(t)$, $\psi^*(t)$, $u^*(t)$ в данном случае выполнено.

Подведем итог проведенному исследованию. Найдено единственное управление (10.31), удовлетворяющее всем условиям принципа максимума. Из физического смысла задачи ясно, что она должна иметь решение. Значит, найденное управление является оптимальным. Искомое время управления T можно определить из условия $H(y^*, \psi^*, u^*)\,|_{t=T}=0$.

§ 10.3. ЗАДАЧА ОПТИМАЛЬНОГО УПРАВЛЕНИЯ С ПОДВИЖНЫМИ КОНЦАМИ

1. Постановка задачи. На практике нередко возникают задачи, в которых один или оба конца траекторий, отвечающих допустимым управлениям, не являются «жестко закрепленными», а могут находиться в пределах заданных множеств, содержащих более чем один элемент. Такие задачи управления называют задачами с подвижными концами.

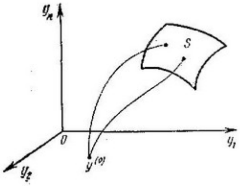

Рис. 10.3

Рассмотрим задачу оптимального управления с подвижным правым концом. В этой задаче целевой функционал подлежит минимизации на множестве всех допустимых управлений, переводящих управляемую систему из заданной точки фазового пространства на заданное множество $S\subset\mathbf{R}^n$ (рис. 10.3). Возможен и случай, когда $S=\mathbf{R}^n$; тогда имеет место *задача оптимального управления со свободным правым концом.*

Далее будем считать, что S — множество решений ситемы уравнений, т. е.

$$S=\{y\in\mathbf{R}^n \mid \varphi_i(y)=0, \quad i=1, 2,\ldots, k\}, \qquad (10.36)$$

где $\varphi_1, \varphi_2, \ldots, \varphi_k$ — заданные функции, непрерывно дифференцируемые на пространстве $\mathbf{R}^n$. В частном случае $\varphi_i(y)=0$, $i=1, 2, \ldots, k$, из равенства (10.36) следует, что $S=\mathbf{R}^n$.

Задача оптимального управления с подвижным правым концом формулируется следующим образом. Дана управляемая система (10.9)

$$\frac{dy}{dt}=f(y, u)$$

и задано множество S вида (10.36). Требуется найти такое допустимое управление $u(t)$, $0\leqslant t\leqslant T$, чтобы точка, описывающая состояние системы, из заданного начального состояния $y^{(0)}$ к конечно-му моменту времени $t=T$ «попала на множество S», т. е. чтобы выполнялись равенства

$$\varphi_i(y(T))=0, \quad i=1, 2, \ldots, k, \tag{10.37}$$

и при этом был достигнут минимум функционала (10.10)

$$I(u)=\int_0^T f_0(y(t), u(t))\,dt.$$

Время управления T может быть фиксированным или нефиксированным.

Пусть пара функций u^*, y^* является решением сформулированной задачи, причем траектория $y^*(t)$ начинается в точке $y^{(0)}$ и заканчивается в точке $y^{(1)}\in S$. Очевидно, управление u^* и соответствующая ему траектория y^* оптимальны и в задаче с системой (10.9) и функционалом (10.10), но с закрепленными концами $y^{(0)}$ и $y^{(1)}$, поскольку все траектории, соединяющие точки $y^{(0)}$ и $y^{(1)}$, входят в множество траекторий, которые соединяют точку $y^{(0)}$ с множеством S. Отсюда следует, что для пары функций u^*, y^*, представляющей собой решение задачи с подвижным правым концом, также выполнен принцип максимума в форме теоремы 10.1 или теоремы 10.2 в зависимости от того, не фиксировано или фиксировано время управления T.

2. Условия трансверсальности. Изменение формулировки необходимых условий оптимальности в задаче с подвижным правым концом заключается в замене граничных условий на правом конце оптимальной траектории на граничные условия для вспомогательных функций ψ_i на том же конце. Последние называют *условиями трансверсальности.* Они заключаются в требовании ортогональности вектора $\psi^*(T)$ множеству S в точке $y^*(T)$ [8, 20].

Рассмотрим эти условия подробно. Уравнение гиперплоскости P_i, касательной в точке $y^*(T)$ к поверхности, которая задана уравнением $\varphi_i(y)=0$, имеет вид

$$\langle \nabla\varphi_i(y^*(T)),\ y-y^*(T)\rangle=0.$$

Пересечение $P=\bigcap_{i=1}^{k} P_i$ представляет собой линейное множество, касательное к S в точке $y^*(T)$, и условие трансверсальности заключается в ортогональности вектора $\psi^*(T)$ этому линейному множеству, т. е.

$$\langle \psi^*(T),\ y-y^*(T)\rangle=0 \text{ для всех } y\in P.$$

Иначе говоря, в точке $y=y^*(T)$ (как и в любой другой точке множества P) достигает минимума линейная функция

$$c(y)=\langle \psi^*(T),\ y\rangle-\langle \psi^*(T),\ y^*(T)\rangle$$

при линейных ограничениях-равенствах

$$\langle \nabla\varphi_i(y^*(T)),\ y\rangle-\langle \nabla\varphi_i(y^*(T)),\ y^*(T)\rangle=0, \quad i=1, 2, \ldots, k.$$

Последнее, согласно теоремам 2.5 и 2.7, имеет место тогда и только тогда, когда найдутся такие числа $\lambda_1, \lambda_2, ..., \lambda_k$, что

$$\psi^*(T) + \sum_{i=1}^{k} \lambda_i \nabla \varphi_i (y^*(T)) = 0_n.$$

Полагая $\alpha_i = -\lambda_i$, $i = 1, 2, ..., k$, условие трансверсальности окончательно сформулируем следующим образом: существуют числа $\alpha_1, \alpha_2, ..., \alpha_k$ такие, что

$$\psi^*(T) = \sum_{i=1}^{k} \alpha_i \nabla \varphi_i (y^*(T)). \tag{10.38}$$

При $\varphi_i(y) \equiv 0$, $i = 1, 2, .., k$, отсюда вытекает условие трансверсальности для задачи со свободным правым концом:

$$\psi^*(T) = 0_n. \tag{10.39}$$

Условие трансверсальности принимает простой вид в случае, когда S — линейное множество, параллельное некоторым координатным осям $Oy_{l_1}, Oy_{l_2}, ..., Oy_{l_p}$, $1 \leqslant l_p \leqslant n$. Напомним, что линейное множество S параллельно координатной оси Oy_i, если вместе с каждой своей точкой $(\bar{y}_1, ..., \bar{y}_{i-1}, \bar{y}_i, \bar{y}_{i+1}, ..., \bar{y}_n) \in S$ оно содержит всю прямую $(\bar{y}_1, ..., \bar{y}_{i-1}, y_i, \bar{y}_{i+1}, ..., \bar{y}_n)$, $-\infty < y_i < \infty$, проходящую через эту точку. Очевидно, что вектор $\boldsymbol{\psi}^*(T)$ ортогонален линейному множеству, параллельному оси Oy_i тогда и только тогда, когда равенство

$$\langle \psi^*(T), (\bar{y}_1, ..., \bar{y}_{i-1}, y_i, \bar{y}_{i+1}, ..., \bar{y}_n)^{\mathrm{T}} - \boldsymbol{y}^*(T) \rangle = 0$$

выполняется для всех $y_i \in \mathbf{R}$. Это имеет место в том и только в том случае, если $\psi_i^*(T) = 0$. Поэтому условие трансверсальности для линейного множества S, параллельного нескольким координатным осям $Oy_{l_1}, Oy_{l_2}, ..., Oy_{l_p}$, принимает вид $\psi^*_{l_i}(T) = 0$, $i = 1, 2, ..., p$.

Для задачи с подвижным левым концом $y^*(0) \in S$, где S имеет вид (10.36), условие трансверсальности формулируется аналогично: существуют такие числа $\alpha_1, \alpha_2, ..., \alpha_k$, что

$$\psi^*(0) = \sum_{i=1}^{k} \alpha_i \nabla \varphi_i (y^*(0)),$$

$$\varphi_i(y^*(0)) = 0, \quad i = 1, 2, ..., k.$$

Если в задаче подвижными являются оба конца, то условия трансверсальности нужно аналогично записать как для левого, так и для правого концов.

Теорема 10.3. *Пусть вектор-функция $u^*(t)$ является оптимальным управлением, а вектор-функция $y^*(t)$ — соответствующей оптимальной траекторией в сформулированной выше задаче с подвижным правым концом, где S имеет вид* (10.36). *Тогда существуют* *непрерывная вектор-функция* $\psi^*(t)=(\psi_1^*(t), \psi_2^*(t), ..., \psi_n^*(t))^T$ *и число* $\psi_0^*\leqslant 0$ *такие, что выполняются условия* 1)—3) *теоремы* 10.1 *и условие трансверсальности* (10.38) *вместе с равенствами* (10.37) *при* $y=y^*$. *Для задачи с нефиксированным временем управления, кроме того, имеет место условие* 4) *теоремы* 10.1.

Аналогично можно сформулировать необходимые условия оптимальности в виде принципа максимума для задачи с левым подвижным концом, а также для задачи с обоими подвижными концами.

Схема применения принципа максимума остается прежней. Для задачи с подвижным правым концом отличие состоит в том, что $2n$ произвольные постоянные, возникающие при решении системы (10.23), вместе с неизвестными числами $\alpha_1, \alpha_2, ..., \alpha_k$ (всего $2n+k$ неизвестных) необходимо определять, исходя из выполнения начального условия $y(0)=y^{(0)}$ и условий трансверсальности (10.37)—(10.38) (всего $2n+k$ равенств). Система перечисленных условий является полной в том смысле, что число уравнений совпадает с числом неизвестных, а значит, все неизвестные в принципе могут быть найдены.

3. Доказательство принципа максимума для задачи со свободным правым концом. Установим справедливость теоремы 10.3 в случае $S=\mathbf{R}_n$ при фиксированном T. Пусть вектор-функция u^* является оптимальным управлением, а вектор-функция y^* — соответствующей оптимальной траекторией. Как отмечено в п. 2, условие трансверсальности для задачи со свободным правым концом имеет вид (10.39) и теорема 10.3 утверждает, что существует ненулевая непрерывная вектор-функция $(\psi_0^*, \psi^*(t))^T$, удовлетворяющая уравнениям (10.12), условиям $\psi_0^*\leqslant 0$, (10.39) и такая, что для нее выполняется условие максимума (10.13).

Если предположить, что $\psi_0^*=0$, то (10.12) превращается в однородную систему линейных дифференциальных уравнений, имеющую единственное решение $\psi^*(t)\equiv 0_n$, удовлетворяющее граничному условию 10.39. Требуется установить существование ненулевой вектор-функции $(\psi_0^*, \psi^*(t))^T$, поэтому случай $\psi_0^*=0$ следует исключить и можно положить $\psi_0^*=-1$. Тогда сопряженную систему (10.12) можно записать в виде

$$\frac{d\psi_i}{dt}=-\sum_{j=1}^{n}\psi_j\frac{\partial f_j(y^*, u^*)}{\partial y_i}+\frac{\partial f_0(y^*, u^*)}{\partial y_i}, \quad i=1, 2, \ldots, n. \qquad (10.40)$$

Это линейная система дифференциальных уравнений, и для нее задача Коши с граничным условием $\psi(T)=0_n$ всегда имеет, и притом единственное, решение. Таким образом, для доказательства теоремы 10.3 в случае $S=\mathbf{R}^n$ достаточно убедиться в том, что для решения указанной задачи Коши, которое мы обозначим через $\psi^*(t)$, выполняется равенство (10.13) для всех $t\in[0, T]$.

Сначала установим справедливость равенства (10.13) для любой точки $\tau\in(0, T)$, в которой функция u^* непрерывна. Рассмотрим произвольный вектор $v\in U$ и семейство допустимых управлений вида

$$u_\lambda(t)=\begin{cases} v & \text{при } t\in[\tau-\lambda, \tau),\\ u^*(t) & \text{при } t\notin[\tau-\lambda, \tau),\end{cases}$$

где параметр λ положителен и мал настолько, что $[\tau-\lambda, \tau)\subset[0, T]$. Каждому допустимому управлению $u_\lambda(t)$ отвечает определенное решение системы (10.9) при $u=u_\lambda(t)$ с начальным условием $y(0)=y^{(0)}$. Обозначим это решение через $y_\lambda(t)$.

Так как u^* — оптимальное управление, то $I(u_\lambda)-I(u^*)\geqslant 0$ для всех достаточно малых положительных λ. Следовательно,

$$\lim_{\lambda\to+0}\frac{1}{\lambda}\left[I(v_\lambda)-I(u^*)\right]\geqslant 0.$$

Используя конкретный вид функционала I, а также равенства $y_\lambda(t)=y^*(t)$, $u_\lambda(t)=u^*(t)$, $0\leqslant t<\tau-\lambda$, перепишем полученное неравенство в виде

$$\lim_{\lambda\to+0}\frac{1}{\lambda}\int_{\tau-\lambda}^{\tau}\left[f_0(y_\lambda(t),\ v)-f_0(y^*(t),\ u^*(t))\right]dt+$$

$$+\lim_{\lambda\to+0}\int_{\tau}^{T}\frac{1}{\lambda}\left[f_0(y_\lambda(t),\ u^*(t))-f_0(y^*(t),\ u^*(t))\right]dt\geqslant 0. \tag{10.41}$$

Функция $f_0(y,\ u)$ непрерывна по y и u, поэтому первый предел, согласно следствию первой интегральной теоремы о среднем [16], равен $f_0(y^*(\tau),\ v)-f_0(y^*(\tau),\ u^*(\tau))$. Функция $f_0(y,\ u)$ дифференцируема по первым n аргументам, поэтому можно записать

$$\frac{1}{\lambda}\left[f_0(y_\lambda(t),\ u^*(t))-f_0(y^*(t),\ u^*(t))\right]=$$

$$=\sum_{i=1}^{n}\frac{\partial f_0(y^*(t),\ u^*(t))}{\partial y_i}\,\frac{1}{\lambda}\left[y_{\lambda i}(t)-y_i^*(t)\right]+O(\lambda)=$$

$$=\left\langle\frac{\partial f_0(y^*(t),\ u^*(t))}{\partial y},\ \frac{1}{\lambda}\left[y_\lambda(t)-y^*(t)\right]\right\rangle+O(\lambda),$$

где $\partial f_0/\partial y=(\partial f_0/\partial y_1,\ \ldots,\ \partial f_0/\partial y_n)^{\tau}$. Таким образом, неравенство (10.41) в результате осуществления предельного перехода принимает вид

$$f_0(y^*(\tau),\ v)-f_0(y^*(\tau),\ u^*(\tau))+$$

$$+\int_{\tau}^{T}\left\langle\frac{\partial f_0(y^*(t),\ u^*(t))}{\partial y},\ z(t)\right\rangle dt\geqslant 0, \tag{10.42}$$

где

$$z(t)=\lim_{\lambda\to+0}\frac{1}{\lambda}\left[y_\lambda(t)-y^*(t)\right],\ t\in[\tau,\ T].$$

Рассмотрим этот предел. Так как τ — точка непрерывности управления u^*, то можно указать достаточно малое положительное λ, при котором функция u^* непрерывна и в некоторой λ-окрестности этой точки. Следовательно, при сделанных предположениях относительно f_0 (см. § 10.1) функция y^* дифференцируема в окрестности точки τ и, учитывая (10.9), можно записать

$$y^*(\tau)=y^*(\tau-\lambda)+\lambda\frac{d\,y^*(\tau-\lambda)}{dt}+o_1(\lambda)=y^*(\tau-\lambda)+$$

$$+\lambda f(y^*(\tau-\lambda),\ u^*(\tau-\lambda))+o_1(\lambda),$$

где через $dy^*(\tau-\lambda)/dt$ обозначен n-мерный вектор из соответствующих производных. Аналогично, для достаточно малых положительных λ выполняется равенство

$$y_\lambda(\tau) = y_\lambda(\tau-\lambda) + \lambda \frac{dy_\lambda(\tau-\lambda)}{dt} + o_2(\lambda) = y^*(\tau-\lambda) +$$

$$+ \lambda f(y^*(\tau-\lambda),\ v) + o_2(\lambda).$$

Здесь использован тот факт, что $y_\lambda(\tau-\lambda) = y^*(\tau-\lambda)$, вытекающий из непрерывности функции $y_\lambda(t)$ и условия $y_\lambda(t) = y^*(t)$, $0 \leqslant t < \tau-\lambda$. Таким образом, верно равенство

$$y_\lambda(\tau) - y^*(\tau) = \lambda[f(y^*(\tau-\lambda),\ v) - f(y^*(\tau-\lambda),\ u^*(\tau-\lambda))] + o(\lambda),$$

а значит,

$$z(\tau) = \lim_{\lambda\to+0} \frac{1}{\lambda}[y_\lambda(\tau) - y^*(\tau)] = f(y^*(\tau),\ v) - f(y^*(\tau),\ u^*(\tau)). \tag{10.43}$$

Выразим вектор производных dz/dt через производные $\partial f_i/\partial y_j$. Из уравнения $dy/dt = f(y,\ u^*)$, которому удовлетворяют на отрезке $[\tau,\ T]$ вектор-функции $y^*(t)$ и $y_\lambda(t)$, следует

$$y_\lambda(t) = y_\lambda(\tau) + \int_\tau^t f(y_\lambda(s),\ u^*(s))\,ds,$$

$$y^*(t) = y^*(\tau) + \int_\tau^t f(y^*(s),\ u^*(s))\,ds.$$

Здесь интеграл от вектор-функции означает вектор из интегралов соответствующих компонент. Имеем

$$\frac{1}{\lambda}[y_\lambda(t) - y^*(t)] = \frac{1}{\lambda}[y_\lambda(\tau) - y^*(\tau)] +$$

$$+ \frac{1}{\lambda}\int_\tau^t [f(y_\lambda(s),\ u^*(s)) - f(y^*(s),\ u^*(s))]\,ds. \tag{10.44}$$

Перейдем в этом равенстве к пределу при $\lambda \to +0$. Предел первого слагаемого в правой части выражения равен $z(\tau)$. Для второго слагаемого получаем

$$\lim_{\lambda\to+0} \frac{1}{\lambda}\int_\tau^t [f(y_\lambda(s),\ u^*(s)) - f(y^*(s),\ u^*(s))]\,ds =$$

$$= \frac{d}{d\lambda}\int_\tau^t f(y_\lambda(s),\ u^*(s))\,ds \,\Big|_{\lambda=+0} =$$

$$= \int_\tau^t \frac{\partial f(y^*(s),\ u^*(s))}{\partial y} \quad \frac{dy_\lambda(s)}{d\lambda}\Big|_{\lambda=+0} ds,$$

где $\partial f/\partial y$ — матрица n-го порядка из частных производных вида $\partial f_i/\partial y_j$. Так как

$$\frac{dy_\lambda(s)}{d\lambda}\Big|_{\lambda=+0} = \lim_{\lambda\to+0} \frac{1}{\lambda}[y_\lambda(s) - y^*(s)] = z(s),$$

то второе слагаемое принимает вид

$$\int_{\tau}^{t} \frac{\partial f(y^*(s),\ u^*(s))}{\partial y} z(s)\,\mathrm{d}s.$$

В соответствии с полученным из (10.44) следует равенство

$$z(t) = z(\tau) + \int_{\tau}^{t} \frac{\partial f(y^*(s),\ u^*(s))}{\partial y} z(s)\,\mathrm{d}s, \quad \tau \leqslant t \leqslant T,$$

которое означает, что функция $z(t)$ на $[\tau, T]$ удовлетворяет дифференциальному уравнению

$$\frac{\mathrm{d}z}{\mathrm{d}t} = \frac{\partial f(y^*(t),\ u^*(t))}{\partial y} z \tag{10.45}$$

с начальным условием (10.43).

Теперь, принимая во внимание равенства (10.45) и (10.40), находим

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle\bar\psi^*,\ z\rangle = \left\langle \frac{\mathrm{d}\bar\psi^*}{\mathrm{d}t},\ z\right\rangle + \left\langle \bar\psi^*,\ \frac{\mathrm{d}z}{\mathrm{d}t}\right\rangle =$$

$$= -\left\langle \frac{\partial f(y^*(t),\ u^*(t))}{\partial y}\bar\psi^*(t),\ z(t)\right\rangle + \left\langle \frac{\partial f_0(y^*(t),\ u^*(t))}{\partial y},\ z(t)\right\rangle +$$

$$+ \left\langle \bar\psi^*(t),\ \frac{\partial f(y^*(t),\ u^*(t))}{\partial y} z(t)\right\rangle = \left\langle \frac{\partial f_0(y^*(t),\ u^*(t))}{\partial y},\ z(t)\right\rangle.$$

где $\bar\psi^*(t) = (\bar\psi^*_1(t),\ \bar\psi^*_2(t),\ \ldots,\ \bar\psi^*_n(t))^{\tau}$. Отсюда, интегрируя и учитывая при этом, что $\bar\psi^*(T) = 0_n$, получаем

$$\langle\bar\psi^*(t),\ z(t)\rangle = -\int_{t}^{T}\left\langle \frac{\partial f_0(y^*(s),\ u^*(s))}{\partial y},\ z(s)\right\rangle \mathrm{d}s.$$

Полагая $t = \tau$, на основании (10.43) имеем

$$\int_{\tau}^{T}\left\langle \frac{\partial f_0(y^*(s),\ u^*(s))}{\partial y},\ z(s)\right\rangle \mathrm{d}s = \langle\bar\psi^*(\tau),\ f(y^*(\tau),\ u^*(\tau)) -$$

$$- f(y^*(\tau),\ v)\rangle.$$

Перепишем неравенство (10.42), учитывая последнее равенство, в виде

$$f_0(y^*(\tau),\ v) - f_0(y^*(\tau),\ u^*(\tau)) + \langle\bar\psi^*(\tau),\ f(y^*(\tau),$$

$$u^*(\tau)) - f(y^*(\tau),\ v)\rangle \geqslant 0,$$

или

$$\langle\bar\psi^*(\tau),\ f(y^*(\tau),\ u^*(\tau))\rangle - f_0(y^*(\tau),\ u^*(\tau)) \geqslant$$

$$\geqslant \langle\bar\psi^*(\tau),\ f(y^*(\tau),\ v)\rangle - f_0(y^*(\tau),\ v).$$

Так как $\psi_0^* = -1$, то последнее неравенство представляет собой неравенство $H(y^*(\tau),\ \psi^*(\tau),\ u^*(\tau)) \geqslant H(y^*(\tau),\ \psi^*(\tau),\ v)$. В силу произвольности выбора

$v \in U$ это означает справедливость равенства (10.13) в каждой точке непрерывности функции u^*.

Теперь предположим, что t — произвольная точка разрыва функции u^*. Очевидно, существует последовательность точек $\{\tau_k\}_{k=1}^{\infty}$ вида $\tau_k \to t-0$, $\tau_k \in [0, T]$, $k=1, 2, \ldots$, причем в каждой точке τ_k функция u^* непрерывна. На основании доказанного выше, можно записать

$$H(y^*(\tau_k), \psi^*(\tau_k), u^*(\tau_k)) \geqslant H(y^*(\tau_k), \psi^*(\tau_k), v)$$

для каждого $v \in U$. Переходя в этом неравенстве к пределу при $k \to \infty$ и учитывая непрерывность функций H, y^*, φ^*, а также непрерывность слева функции u^* (см. § 10.1), получаем требуемое:

$$H(y^*(t), \psi^*(t), u^*(t)) \geqslant H(y^*(t), \psi^*(t), v) \text{ для всех } v \in U.$$

Замечание. В приведенном доказательстве имеется «пробел» — существование предела для $z(t)$ установлено не было. Обычно существование этого предела доказывают с помощью теоремы о непрерывной дифференцируемости решения дифференциального уравнения по начальным данным [2].

4. Принцип максимума для неавтономных систем. Рассмотрим задачу оптимального управления системой

$$\frac{dy}{dt} = f(t, y, u) \tag{10.46}$$

на временно́м промежутке $[t_0, T]$ с заданным начальным $y^{(0)}$ и конечным $y^{(1)}$ состояниями:

$$y(t_0) = y^{(0)}, \quad y(T) = y^{(1)}. \tag{10.47}$$

В качестве критерия оптимальности, подлежащего минимизации, возьмем функционал

$$I(u) = \int_{t_0}^{T} f_0(t, y(t), u(t))\, dt. \tag{10.48}$$

Система (10.46) отличается от рассмотренной ранее системы (10.9) тем, что правая ее часть, т. е. вектор-функция f, явно зависит от t. Такую систему называют *неавтономной системой* в отличие от *автономной системы* (10.9). Заметим, что в функционале (10.48) подынтегральная функция также явно зависит от t. В частном случае, когда указанные функции не зависят от t, получаем задачу оптимального управления, рассмотренную в § 10.2.

Начальный момент считаем заранее заданным и зафиксированным, а конечный — нефиксированным и подлежащим определению. Предполагаем, что для функций y, u, f, f_0 выполнены допущения, аналогичные сделанным при формулировке задачи управления автономной системой и, кроме того, считаем, что функции $f_0, f_1, f_2, \ldots, f_n$ непрерывно дифференцируемы по t при $t \geqslant t_0$. Область управления представляет собой фиксированное множество; ее, как и раньше, будем обозначать через $U \subset \mathbf{R}^m$.

Принцип максимума для сформулированной задачи можно вывести из принципа максимума для автономных систем с подвижным правым концом. Для этого введем дополнительную фазовую переменную y_{n+1} с помощью равенств $dy_{n+1}/dt=1$, $y_{n+1}(t_0)=t_0$, из которых следует, что $y_{n+1}(t)\equiv t$. В соответствии с этим исходную задачу запишем в виде

$$\begin{cases} \dfrac{dy}{dt}=f(y_{n+1},\ y,\ u), \\ \qquad\qquad\qquad\qquad t_0\leqslant t\leqslant T, \\ \dfrac{dy_{n+1}}{dt}=1, \end{cases} \tag{10.49}$$

$$I(u)=\int_{t_0}^{T} f_0(y_{n+1}(t),\ y(t),\ u(t))\,dt\to\min.$$

Это уже задача управления автономной системой с фиксированным начальным состоянием $(y_1^{(0)},\ y_2^{(0)},\dots,\ y_n^{(0)},t_0)\in\mathbf{R}^{n+1}$ и подвижным конечным состоянием, принадлежащим прямой из $\mathbf{R}^{n+1}$ вида

$$L=\{y\in\mathbf{R}^{n+1}\,|\,y_i=y_i^{(1)},\ i=1,2,\dots,n;\ -\infty<y_{n+1}<\infty\}.$$

Пусть u^* и y^* — оптимальное управление и соответствующая ему оптимальная траектория в исходной задаче оптимального управления неавтономной системой с закрепленными концами. Тогда эта пара функций является оптимальной в автономной задаче (10.49) с подвижным правым концом. Согласно теореме 10.3, существуют непрерывная вектор-функция $(\psi_1^*(t),\ \dots,\ \psi_n^*(t),\ \psi^*_{n+1}(t))^{\mathrm{T}}$ и число $\psi_0^*\leqslant 0$ такие, что указанная вектор-функция является решением системы

$$\begin{cases} \dfrac{d\psi_i}{dt}=-\displaystyle\sum_{j=0}^{n}\psi_j\frac{\partial f_j(t,\ y^*,\ u^*)}{\partial y_i}, \quad i=1,2,\dots,n, \\ \dfrac{d\psi_{n+1}}{dt}=-\displaystyle\sum_{j=0}^{n}\psi_j\frac{\partial f_j(t,y^*,u^*)}{\partial t}, \end{cases} \tag{10.50}$$

и выполняется условие максимума

$$H(t,\ y^*(t),\ \psi^*(t),\ u^*(t))+\psi^*_{n+1}(t)=$$
$$=\max_{u\in U}\,[H(t,\ y^*(t),\ \psi^*(t),u))+\psi^*_{n+1}(t)], \tag{10.51}$$

где

$$H(t,\ y,\ \psi,\ u)=\sum_{i=0}^{n}\psi_i f_i(t,\ y,\ u). \tag{10.52}$$

В равенстве (10.51) функция $\psi^*_{n+1}(t)$ сокращается, и оно окончательно имеет вид

$$H(t,\ y^*(t),\ \psi^*(t),u^*(t))=\max_{u\in U}H(t,\ y^*(t),\ \psi^*(t),\ u),\ t_0\leqslant t\leqslant T. \tag{10.53}$$

Условие 4) теоремы 10.1 записывается в форме

$$H(t,\ y^*(t),\ \psi^*(t),\ u^*(t))+\psi^*_{n+1}(t)=0,\quad t_0\leqslant t\leqslant T, \tag{10.54}$$

а условие трансверсальности на правом конце означает ортогональность вектора $(\psi_1{}^*(T),\ \ldots,\ \psi_n{}^*(T),\ \psi^*{}_{n+1}(T))^{\text{т}}$ прямой L, т. е. $\varphi^*{}_{n+1}(T)=0$, так как прямая L параллельна оси Oy_{n+1}. Используя последнее равенство, из второго соотношения (10.50) при $\psi_i=\psi_i{}^*$, $i=0,\ 1,\ \ldots,\ n$, получаем

$$\psi^*_{n+1}(t)=-\int_T^t\sum_{j=0}^n\psi_j^*(s)\,\frac{\partial f_j(s,\ y^*(s),\ u^*(s))}{\partial t}\,\mathrm{d}s,$$

а значит, (10.54) можно записать в виде

$$H(t,\ y^*(t),\ \psi^*(t),\ u^*(t))=$$

$$=\int_T^t\sum_{j=0}^n\psi_j^*(s)\frac{\partial f_j(s,\ y^*(s),\ u^*(s))}{\partial t}\,\mathrm{d}s,\ \ t_0\leqslant t\leqslant T, \tag{10.55}$$

где $\psi_0{}^*(s)\equiv\psi_0{}^*$.

Предположение о том, что $(\psi_0{}^*,\ \psi_1{}^*(t),\ \ldots,\ \psi_n{}^*(t))^{\text{т}}\equiv0$, из (10.54) влечет $\psi^*{}_{n+1}(t)\equiv0$. Но в силу теоремы 10.3 это невозможно; следовательно, вектор-функция $(\psi_0{}^*,\ \psi_1{}^*(t),\ \ldots,\ \psi_n{}^*(t))^{\text{т}}$ ненулевая.

Теорема 10.4. *Пусть u^* оптимальное управление, а y^* — соответствующая оптимальная траектория в задаче оптимального управления неавтономной системой* (10.46) *с граничными условиями* (10.47) *и критерием оптимальности* (10.48) *на временном промежутке $[t_0,\ T]$, где конечный момент времени не фиксирован. Тогда существуют непрерывная вектор-функция $\psi^*(t)=(\psi_1{}^*(t),\ \psi_2{}^*(t),\ \ldots,\ \psi_n{}^*(t))^{\text{т}}$ и число $\psi_0{}^*\leqslant0$ такие, что:*

1) *вектор-функция вида $(\psi_0{}^*,\ \psi^*(t))^{\text{т}}$, $t_0\leqslant t\leqslant T$ является ненулевой;*

2) $\psi^*(t)$ — *решение системы дифференциальных уравнений*

$$\frac{\mathrm{d}\psi_i}{\mathrm{d}t}=-\sum_{j=0}^n\psi_j\frac{\partial f_j(t,\ y^*,\ u^*)}{\partial y_i},\quad i=1,\ 2,\ldots,\ n;$$

3) *выполняется условие максимума* (10.53), *где H имеет вид* (10.52);

4) *выполняется равенство* (10.55).

Как и в автономном случае, из справедливости равенства (10.55) в фиксированный момент времени $t\in[t_0,\ T]$ следует его выполнение для всех $t\in[t_0,\ T]$ [20]. Поэтому справедливость этого равенства можно проверять, например, при $t=T$; в этом случае оно принимает вид $H(T,\ y^*(T),\ \psi^*(T),\ u^*(T))^{\text{т}}=0$. Отсюда обычно определяют искомое время T.

Для задачи оптимального управления неавтономной системой с фиксированным временем T также справедливы условия оптимальности в форме принципа максимума, приведенные в теореме 10.4. В этом случае отпадает необходимость определения момента времени T и поэтому условие 4) лишнее.

Для неавтономных систем возможна постановка задачи оптимального управления с подвижными концами. Для таких задач имеет место теорема 10.4 и, кроме того, выполняются условия трансверсальности, которые записываются точно так же, как и для автономных систем (см. п. 2).

5. Задача терминального управления. Пусть дана неавтономная система (10.46) при $t_0 \leqslant t \leqslant T$. Рассмотрим задачу оптимального управления этой системой с закрепленным левым концом ($y(t_0) = y^{(0)}$), свободным правым концом и нефиксированным временем T. При этом в качестве критерия оптимальности, подлежащего минимизации, возьмем терминальный функционал вида

$$I(u) = \Phi(y(T)),$$

где $\Phi\colon \mathrm{R}^n \to \mathrm{R}$ — заданная непрерывно дифференцируемая функция. Этот функционал задан на множестве всех допустимых управлений (т. е. кусочно-непрерывных функций u вида $u(t) \in U \subset \mathrm{R}^m$ при всех $t \in [t_0, T]$), для которых $y^{(0)}$— начальная точка соответствующих им траекторий.

В частном случае, когда

$$\Phi\ (y(T)) = \sum_{i=1}^{n} c_i y_i(T), \tag{10.56}$$

где c_i — постоянная, рассматриваемую задачу можно свести к задаче с интегральным функционалом. В самом деле, из (10.46) следует равенство

$$y_i(T) = \int_{t_0}^{T} f_i(t,\ y(t),\ u(t))\,\mathrm{d}t + y_i^{(0)}, \quad i = 1, 2, \ldots, n.$$

Поэтому

$$I(u) = \Phi(y(T)) = \int_{t_0}^{T} \sum_{i=1}^{n} c_i f_i(t,\ y(t),\ u(t))\,\mathrm{d}t + \sum_{i=1}^{n} y_i^{(0)}.$$

Так как второе слагаемое в правой части равенства — константа, то задача оптимального управления с терминальным функционалом вида (10.56) равносильна задаче с интегральным функционалом

$$\int_{t_0}^{T} \sum_{i=1}^{n} c_i f_i(t,\ y(t),\ u(t))\,\mathrm{d}t.$$

В общем случае необходимые условия оптимальности в форме принципа максимума для задачи терминального управления полностью аналогичны условиям теоремы 10.4 (см. [8]), где

$$H(t, y, \psi, u) = \sum_{i=1}^{n} \psi_i f_i(t, y, u),$$

Если конечный момент T фиксирован, то условие 4) следует исключить. Изменяется условие трансверсальности; оно принимает вид

$$\psi^*(T) = \psi_0^* \Phi_y(y^*(T)),$$

где

$$\Phi_y = (\partial\Phi/\partial y_1, \partial\Phi/\partial y_2, \ldots, \partial\Phi/\partial y_n)^{\mathrm{т}}.$$

Более общие постановки задач терминального управления и соответствующие им версии принципа максимума можно найти в [2, 8].

6. Сведе́ние задачи оптимального управления к дискретной задаче. Будем считать, что момент времени T фиксирован, левый конец закреплен ($y(t_0) = y^{(0)}$), а правый — свободен.

Разделим временной отрезок $[t_0, T]$ на N равных частей точками $t_k = kh$, $k = 0, 1, 2, \ldots, N$; $Nh = T$. В каждой точке t промежутка $[t_k, t_{k+1})$ заменим вектор производных $\mathrm{d}y/\mathrm{d}t$, вектор-функцию $f(t, y(t), u(t))$ и функцию $f_0(t, y(t), u(t))$ приближенными выражениями:

$$\frac{\mathrm{d}y}{\mathrm{d}t} \approx \frac{y(t_{k+1}) - y(t_k)}{h},$$

$$f(t, y(t), u(t)) \approx f(t_k, y(t_k), u(t_k)), \quad t_k \leqslant t < t_{k+1}.$$

$$f_0(t, y(t), u(t)) \approx f_0(t_k, y(t_k), u(t_k)),$$

В результате вместо системы дифференциальных уравнений (10.46) получаем векторное равенство

$$y(t_{k+1}) = y(t_k) + hf(t_k, y(t_k), u(t_k)), \quad k = 0, 1, \ldots, N,$$

которое после введения обозначений

$$y(t_k) = y^{(k)}, \ u(t_k) = u^{(k+1)}, \ f(t_k, y(t_k), u(t_k)) = f^{(k)}(y^{(k)}, u^{(k+1)})$$

принимает вид

$$y^{(k)} = y^{(k-1)} + hf^{(k-1)}(y^{(k-1)}, u^{(k)}), \quad k = 1, 2, \ldots, N, \tag{10.57}$$

и задает дискретную систему управления.

Критерий оптимальности, используя обозначение $f_0(t_k, y(t_k), u(t_k)) = f_0^{(k)}(y^{(k)}, u^{(k+1)})$, запишем в виде

$$I = \int_{t_0}^{T} f_0(t, y(t), u(t))\,\mathrm{d}t \approx \sum_{k=0}^{N-1} f_0(t_k, y(t_k), u(t_k))h =$$

$$= h \sum_{k=1}^{N} f_0^{(k-1)}(y^{(k-1)}, u^{(k)}).$$

В результате приходим к задаче оптимального управления дискретной системой (10.57) с критерием оптимальности

$$h \sum_{k=1}^{N} f_0^{(k-1)}(y^{(k-1)}, u^{(k)}),$$

заданным начальным состоянием $y(t_0)=y^{(0)}$ и условием $u^{(k)} \in U$, $k=1, 2, \ldots, N$. Решая эту задачу, например, с помощью метода, описанного в гл. 4, или же используя алгоритмы нелинейного программирования, определяем последовательность управлений $\bar{u}^{(1)}, \bar{u}^{(2)}, \ldots, \bar{u}^{(N)}$. С помощью найденной последовательности приближенное оптимальное управление для исходной задачи (10.46), (10.48) можно записать в виде

$$\bar{u}_h(t)=\begin{cases} \bar{u}^{(1)}, \text{ если } t_0 \leqslant t \leqslant t_1, \\ \bar{u}^{(2)}, \text{ если } t_1 < t \leqslant t_2, \\ \dots\dots\dots\dots \\ \bar{u}^{(N)}, \text{ если } t_{N-1} < t \leqslant T. \end{cases}$$

Обозначим через $\bar{y}_h(t)$ траекторию, соответствующую управлению $\bar{u}_h(t)$. Можно доказать (см. кн.: *Ермольев Ю. М.* и др. Математические методы исследования операций.— Киев: Вища школа, 1979), что, выбирая шаг разбиения h достаточно малым положительным, норму разности $\|\bar{y}_h - y^*\|_0$, где $y^*(t)$ — оптимальная траектория исходной задачи, можно также сделать сколь угодно малой. Это означает, что полученное указанным «конечномерным» способом приближенное решение с уменьшением шага h будет все точнее аппроксимировать оптимальное решение исходной бесконечномерной задачи оптимального управления.

7. Принцип максимума и условия оптимальности вариационного исчисления. Принцип максимума применим и для решения вариационных задач. Из него, как следствия, вытекают уравнения Эйлера — Лагранжа, условие Лежандра, а также условия трансверсальности, рассмотренные в предыдущей главе. Это означает, что принцип максимума как бы «объемлет» все перечисленные условия оптимальности вариационного исчисления.

Рассмотрим простейшую задачу вариационного исчисления с подвижным правым концом:

$$I=\int_{x_0}^{x_1} f(x, y(x), y'(x))\,dx \to \min;$$

$$y(x_0)=y^{(0)}, \; y(x_1)=g(x_1).$$

Считаем, что функция f определена на пространстве $\mathbf{R}^3$ и обладает всеми необходимыми в дальнейшем свойствами, а g — заданная гладкая функция.

Пусть непрерывно дифференцируемая функция y^* представляет собой решение сформулированной задачи. Согласно следствию 9.1

из леммы о скруглении углов, эта функция также является решением расширенной задачи, в которой операция минимизации осуществляется в пространстве кусочно-непрерывно дифференцируемых функций. Заменим x на t и сформулируем расширенную задачу в виде следующей задачи оптимального управления с подвижным правым концом:

$$I(u) = \int_{t_0}^{T} f(t, y(t), u(t))\,\mathrm{d}t \to \min;$$

$$\frac{\mathrm{d}y}{\mathrm{d}t} = u(t); \quad U = \mathbf{R},$$

$$y(t_0) = y^{(0)}, \ \varphi(T, y(T)) \equiv y(T) - g(T) = 0,$$

где конечный момент времени T считаем нефиксированным. Здесь x_0 и x_1 также заменены соответственно на t_0 и T. Мы перешли к расширенной задаче, поэтому управление u принадлежит классу кусочно-непрерывных функций, а значит, можно воспользоваться теоремой 10.4. Составим функцию Понтрягина:

$$H(t, y, \psi_0, \psi_1, u) = \psi_0 f(t, y, u) + \psi_1 u.$$

Очевидно, управление вида $u^*(t) = \mathrm{d}y^*(t)/\mathrm{d}t$ является оптимальным для сформулированной выше задачи оптимального управления. По теореме 10.4, существует число $\psi_0^* \leqslant 0$ и функция $\psi_1^*(t)$ такие, что имеет место условие максимума

$$H(t, y^*(t), \psi_0^*, \psi_1^*(t), u^*(t)) = \\ = \max_{u \in \mathbf{R}} H(t, y^*(t), \psi_0^*, \psi_1^*(t), u), \ t_0 \leqslant t \leqslant T. \tag{10.58}$$

Отсюда следует, что производная функции Понтрягина по u при $u = u^*$ должна обращаться в нуль:

$$H_u|_{u=u^*} = \psi_0^* f_u(t, y^*(t), u^*(t)) + \psi_1^*(t) = 0, \ \ t_0 \leqslant t \leqslant T. \tag{10.59}$$

Если предположить, что $\psi_0^* = 0$, то из равенства (10.59) имеем $\psi_1^*(t) \equiv 0$, что невозможно. Следовательно, $\psi_0^* < 0$ и можно положить $\psi_0^* = -1$.

Функция ψ_1^*, согласно теореме 10.4, удовлетворяет сопряженной системе, т. е.

$$\frac{\mathrm{d}\psi_1^*}{\mathrm{d}t} = -\psi_0^* f_y(t, y^*(t), u^*(t)) = f_y(t, y^*(t), u^*(t)).$$

Поэтому

$$\psi_1^*(t) = \int_{t_0}^{t} f_y(s, y^*(s), u^*(s))\,\mathrm{d}s + \psi_1^*(t_0),$$

а значит, равенство (10.59) можно записать в виде

$$-f_u(t,\ y^*(t),\ u^*(t))+\int\limits_{t_0}^{t} f_y(s,\ y^*(s),\ u^*(s))\,\mathrm{d}s+\psi_1^*(t_0)=0.$$

Дифференцируя это равенство по t, получаем

$$-\frac{\mathrm{d}}{\mathrm{d}t} f_u(t,\ y^*(t),\ u^*(t))+f_y(t,\ y^*(t),\ u^*(t))=0,$$

или

$$f_y(t,\ y^*(t),\ u^*(t))-\frac{\mathrm{d}}{\mathrm{d}t} f_u(t,\ y^*(t),\ u^*(t))=0,\quad t_0\leqslant t\leqslant T.$$

Таким образом, функция y^* вместе со своей производной u^* удовлетворяет уравнению Эйлера — Лагранжа из § 9.2.

Из условия максимума (10.58) следует неположительность второй производной функции H по u, т. е.

$$H_{uu}|_{u=u^*}=-f_{uu}(t,\ y^*(t),\ u^*(t))\leqslant 0,\quad t_0\leqslant t\leqslant T,$$

или $f_{uu}(t,\ y^*(t),\ u^*(t))\geqslant 0$. Итак, условие Лежандра из § 9.3 также выполняется.

Выведем условие трансверсальности. Функция $\varphi=\varphi(t,\ y)$, фигурирующая в условии на правом конце траектории, имеет более общий вид, чем рассмотренная в п. 2, так как зависит еще и от t: $\varphi(t,\ y)=y-g(t)$. Поэтому непосредственно использовать условия трансверсальности из п. 2 нельзя. Однако если воспользоваться приемом включения времени в список фазовых координат (см. п. 4), то на основе условий трансверсальности из п. 2 можно вывести аналогичные условия для более общего случая:

$$\psi^*(T)=\sum_{i=1}^{k}\alpha_i\nabla_y\varphi_i(T,\ y^*(T)),$$

$$H(T,\ y^*(T),\ \psi^*(T),\ u^*(T))=-\sum_{i=1}^{k}\alpha_i\varphi_{it}(T,\ y^*(T)).$$

В данной задаче эти условия имеют вид

$$\psi_1^*(T)=\alpha_1,$$

$$H(T,\ y^*(T),\ \psi_0^*,\ \psi_1^*(T),\ u^*(T))=-\alpha_1 g'(T).$$

Отсюда, принимая во внимание конкретный вид функции H и используя равенство $\psi_1^*(t)=f_u(t,\ y^*(t),\ u^*(t))$ при $t=T$, получаем

$$-f(T,\ y^*(T),\ u^*(T))+f_u(T,\ y^*(T),\ u^*(T))\,u^*(T)=$$
$$=-f_u(T,\ y^*(T),\ u^*(T))\,g'(T),$$

или

$$f(T,\ y^*(T),\ u^*(T))=f_u(T,\ y^*(T),\ u^*(T))(u^*(T)-g'(T)),$$

что с точностью до обозначений совпадает с условием трансверсальности на правом конце из теоремы 9.5.

В рассмотренной задаче $U=\mathbf{R}$. В вариационных задачах более общего вида множеством U может быть произвольное открытое множество. Принцип максимума применим как для открытого, так и замкнутого множества U. В реальных задачах множество U нередко замкнуто, значит, методы вариационного исчисления использовать нельзя. В таких случаях принцип максимума оказывается незаменимым.

§ 10.4. ЛИНЕЙНАЯ ЗАДАЧА ОПТИМАЛЬНОГО БЫСТРОДЕЙСТВИЯ

1. Постановка задачи и формулировка принципа максимума. В векторной форме линейная автономная управляемая система задается дифференциальным уравнением вида

$$\frac{\mathrm{d}y}{\mathrm{d}t}=Ay+Bu, \tag{10.60}$$

где $\mathrm{d}y/\mathrm{d}t=(\mathrm{d}y_1/\mathrm{d}t,\ \mathrm{d}y_2/\mathrm{d}t,\ldots,\mathrm{d}y_n/\mathrm{d}t)^{\mathrm{T}}$,

$$y=(y_1,\ y_2,\ldots,y_n)^{\mathrm{T}},\ u=(u_1,\ u_2,\ldots,u_m)^{\mathrm{T}},$$

$$A=\begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix},\quad B=\begin{pmatrix} b_{11} & b_{12} & \ldots & b_{1m} \\ b_{21} & b_{22} & \ldots & b_{2m} \\ \cdot & \cdot & \cdot & \cdot \\ b_{n1} & b_{n2} & \ldots & b_{nm} \end{pmatrix},$$

а a_{ij} и b_{ij} — заданные числовые коэффициенты. Допустимыми управлениями здесь являются кусочно-непрерывные функции $u(t)$ определенные на различных промежутках вида $[0,\ T]$, где $0\leqslant T\leqslant\tau$ (τ — фиксированное положительное число), и удовлетворяющие включению $u(t)\in U\subset\mathbf{R}^m$ для всех $t\in[0,\ T]$. Начальным состоянием системы является точка $y^{(0)}\in\mathbf{R}^n$, а конечным — точка $y^{(1)}\in\mathbf{R}^n$. Каждому допустимому управлению $u(t)$, согласно (10.60), отвечает соответствующая кусочно-непрерывно дифференцируемая траектория $y(t)$ (см. § 10.1).

Линейная задача оптимального быстродействия состоит в определении среди всех допустимых управлений, переводящих систему (10.60) из начального состояния $y^{(0)}$ в состояние $y^{(1)}$, такого управления, которое осуществляет этот переход за наименьшее возможное время, т. е. управления, которое минимизирует функционал

$$I(u)=\int_0^T \mathrm{d}t=T.$$

Минимальное время управления также подлежит определению. Эта задача представляет собой частный случай задачи оптимального управления с закрепленными концами и нефиксированным временем (см. § 10.1), когда система управления является линейной и в формуле (10.8) $f_0(y,\ u)\equiv1$.

Согласно теореме 10.1, запишем принцип максимума для сформулированной задачи. Составим функцию Понтрягина:

$$H=\psi_0+\langle\psi,\ Ay+Bu\rangle.$$

Пусть пара функций u^*, y^* является оптимальным решением линейной задачи оптимального быстродействия. Сопряженная система (10.12) в данном случае принимает вид

$$\frac{d\psi_i}{dt}=-\sum_{j=1}^{n}\psi_j a_{ji},\ i=1,\ 2,\dots,n,$$

или в векторной форме

$$\frac{d\psi}{dt}=-A^{\mathrm{T}}\psi, \tag{10.61}$$

где A^{T} — транспонированная матрица A.

Равенство (10.14) приобретает соответственно вид

$$\psi_0^*+\langle\psi^*(t),\ Ay^*(t)+Bu^*(t)\rangle=0,\ 0\leqslant t\leqslant T. \tag{10.62}$$

Если $\psi^*(t)\equiv 0_n$, то из равенства (10.62) следует, что $\psi_0^*=0$, т. е. $\boldsymbol{\psi}^*(t)\equiv 0_{n+1}$. Согласно условию 1) теоремы 10.1, это не должно иметь места и поэтому $\psi^*(t)\not\equiv 0_n$. Учитывая, что $\psi_0^*\leqslant 0$, равенство (10.62) запишем в виде

$$\langle\psi^*(t),\ Ay^*(t)+Bu^*(t)\rangle\geqslant 0,\ 0\leqslant t\leqslant T. \tag{10.63}$$

Условие максимума (10.13) принимает вид

$$\psi_0^*+\langle\psi^*(t),\ Ay^*(t)\rangle+\langle\psi^*(t),\ Bu^*(t)\rangle=$$
$$=\max_{u\in U}\left[\psi_0^*+\langle\psi^*(t),\ Ay^*(t)\rangle+\langle\psi^*(t),\ Bu\rangle\right],$$

что равносильно равенству

$$\langle\psi^*(t),\ Bu^*(t)\rangle=\max_{u\in U}\langle\psi^*(t),\ Bu\rangle,\ 0\leqslant t\leqslant T. \tag{10.64}$$

Сформулируем принцип максимума применительно к рассматриваемой линейной задаче оптимального быстродействия. *Если u^* и y^* — оптимальное решение линейной задачи оптимального быстродействия, то существует ненулевое решение* $\psi^*(t)=(\psi_1^*(t),\ \psi_2^*(t),\ \dots,\ \psi_n^*(t))^{\mathrm{T}}$ *линейной однородной системы дифференциальных уравнений* (10.61) *такое, что выполняется условие максимума* (10.64), *а также имеет место неравенство* (10.63).

2. Принцип максимума — достаточное условие оптимальности. Рассмотрим задачу наискорейшего перевода линейной системы (10.60) из точки $y^{(0)}$ в положение равновесия системы $dy/dt=Ay$, т. е. будем считать, что $y^{(1)}=0_n$. В качестве области управления U возьмем многомерный параллелепипед

$$U=\{u\in\mathrm{R}^m\mid \alpha_i\leqslant u_i\leqslant\beta_i,\ i=1,\ 2,\dots,m\}, \tag{10.65}$$

где α_i, β_i — заданные числа. Будем считать, что начало координат

0_m принадлежит параллелепипеду (10.65), но не является его вершиной, т. е. существует отрезок в U, параллельный какой-нибудь координатной оси пространства $\mathbf{R}^m$, внутри которого содержится начало координат.

Теорема 10.5. *Дополнительно к сделанным предположениям будем считать выполненным условие общности положения, согласно которому для каждого $i=1, 2, ..., m$ является линейно независимой система векторов вида*

$$B\,\mathrm{e}^{(i)},\ AB\,\mathrm{e}^{(i)}, \ldots, A^{n-1}B\,\mathrm{e}^{(i)}, \tag{10.66}$$

где $\mathrm{e}^{(i)}$ — вектор с нулевыми компонентами, среди которых только i-я отлична от нуля и равна единице. Если допустимое управление u^* переводит систему (10.60) из состояния $y^{(0)}$ в состояние $y^{(1)}=0_n$ и вместе с соответствующей ему траекторией y^* и некоторой вектор-функцией ψ^* удовлетворяет принципу максимума, то u^* — оптимальной управление, а y^* — оптимальная траектория в линейной задаче оптимального быстродействия.

□ Предположим противное: допустимое управление u^* вместе с соответствующей траекторией y^* и ненулевой вектор-функцией ψ^* удовлетворяют всем условиям принципа максимума, но u^* не является оптимальным управлением. Это означает, что существует допустимое управление $\hat{u}$, которое переводит систему (10.60) из состояния $y^{(0)}$ в начало координат 0_n за время $\hat{T}$ меньшее, чем время T, соответствующее управлению u^*: $\hat{T}<T$. Обозначим через $\hat{y}$ траекторию, отвечающую управлению $\hat{u}$.

Поскольку u^* и y^* удовлетворяют уравнению (10.60), а функция ψ^* имеет вид (10.61), находим

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle\psi^*, y^*\rangle = \left\langle\frac{\mathrm{d}\psi^*}{\mathrm{d}t}, y^*\right\rangle + \left\langle\psi^*, \frac{\mathrm{d}y^*}{\mathrm{d}t}\right\rangle =$$
$$= -\langle A^{\mathrm{T}}\psi^*, y^*\rangle + \langle\psi^*, Ay^*+Bu^*\rangle = \langle\psi^*, Bu^*\rangle. \tag{10.67}$$

Отсюда, интегрируя в пределах от 0 до $\hat{T}$, получаем

$$\psi^*(\hat{T})\,y^*(\hat{T}) - \psi^*(0)\,y^*(0) = \int\limits_0^{\hat{T}}\langle\psi^*(t), Bu^*(t)\rangle\,\mathrm{d}t.$$

Аналогично можно найти

$$\psi^*(\hat{T})\,\hat{y}(\hat{T}) - \psi^*(0)\,\hat{y}(0) = \int\limits_0^{\hat{T}}\langle\psi^*(t), B\hat{u}(t)\rangle\,\mathrm{d}t.$$

Следовательно, учитывая равенство $\hat{y}(\hat{T})=0_n$, можно записать

$$\psi^*(\hat{T})\,y^*(\hat{T}) = [\psi^*(\hat{T})\,y^*(\hat{T}) - \psi^*(0)\,y^*(0)] - [\psi^*(\hat{T})\,\hat{y}(\hat{T}) -$$
$$-\psi^*(0)\,\hat{y}(0)] = \int\limits_0^{\hat{T}}[\langle\psi^*(t), Bu^*(t)\rangle - \langle\psi^*(t), B\hat{u}(t)\rangle]\,\mathrm{d}t. \tag{10.68}$$

Так как для функции u^* выполняется условие максимума, то

$$\langle \psi^*(t),\ Bu^*(t)\rangle \geqslant \langle \psi^*(t),\ B\hat{u}(t)\rangle,\ 0 \leqslant t \leqslant \hat{T}.$$

Поэтому подынтегральная функция в (10.68) не отрицательна, а значит, не отрицателен и сам интеграл, т. е.

$$\psi^*(\hat{T})\,y^*(\hat{T}) \geqslant 0. \tag{10.69}$$

С другой стороны, интегрируя (10.67) в пределах от $\hat{T}$ до T, получаем

$$\psi^*(\hat{T})\,y^*(\hat{T}) = \psi^*(\hat{T})\,y^*(\hat{T}) - \psi^*(T)\,y^*(T) = -\int\limits_{\hat{T}}^{T} \langle \psi^*(t),\ Bu^*(t)\rangle\,\mathrm{d}t. \tag{10.70}$$

Из выполнения для функции u^* условия максимума (10.64) при $u = 0_m \in U$ следует

$$\langle\psi^*(t),\ Bu^*(t)\rangle \geqslant \langle \psi^*(t),\ B\cdot 0_m\rangle = 0,\ 0 \leqslant t \leqslant T. \tag{10.71}$$

Поэтому равенство (10.70) влечет неравенство $\psi^*(\hat{T})\,y^*(\hat{T}) \leqslant 0$, которое вместе с (10.69) дает

$$\psi^*(\hat{T})\,y^*(\hat{T}) = \int\limits_{\hat{T}}^{T} \langle \psi^*(t),\ Bu^*(t)\rangle\,\mathrm{d}t = 0.$$

Здесь подынтегральная функция согласно неравенству (10.71) не отрицательна, а значит, равенство интеграла нулю влечет

$$\langle \psi^*(t),\ Bu^*(t)\rangle = 0,\ \hat{T} \leqslant t \leqslant T.$$

Отсюда и из равенства (10.64) следует, что линейная функция $\langle \psi^*(t), Bu\rangle$ переменного u в точке $u = 0_m$ достигает на множестве U своего наибольшего значения, равного нулю.

Согласно предположениям теоремы, существует отрезок, принадлежащий U, который параллелен некоторой координатной оси и содержит 0_m внутри себя. Обозначим концы этого отрезка через u' и u'' и будем считать, что он параллелен i-й координатной оси. В силу доказанного линейная функция $c(u) = \langle \psi^*(t),\ Bu\rangle$ в точке $u = 0_m$ достигает максимума, равного нулю, $\hat{T} \leqslant t \leqslant T$. В этом случае $c(u) = 0$ для всех точек отрезка, соединяющего точки u' и u'', в частности для $u = u'$, $u = u''$:

$$\langle \psi^*(t),\ Bu'\rangle = \langle \psi^*(t),\ Bu''\rangle = 0,\ \hat{T} \leqslant t \leqslant T.$$

Отсюда

$$\langle \psi^*(t),\ B(u' - u'')\rangle = 0,\ \hat{T} \leqslant t \leqslant T.$$

Так как у вектора $u' - u''$ только i-я компонента отлична от нуля, то из последнего равенства получаем

$$\langle \psi^*(t),\ B\,\mathrm{e}^{(i)}\rangle = 0,\ \hat{T} \leqslant t \leqslant T. \tag{10.72}$$

Дифференцируя последнее равенство по t, находим

$$\left\langle \frac{d\psi^*(t)}{dt},\ Be^{(i)} \right\rangle = -\langle A^{\mathrm{T}}\psi^*(t),\ Be^{(i)}\rangle = 0,$$

или

$$\langle \psi^*(t),\ ABe^{(i)}\rangle = 0,\ \hat{T} \leqslant t \leqslant T. \tag{10.73}$$

Аналогично, последовательно дифференцируя, получаем равенства

$$\langle \psi^*(t),\ A^2Be^{(i)}\rangle = 0,\ \hat{T} \leqslant t \leqslant T,$$

$$\dots\dots\dots\dots\dots\dots$$

$$\langle \psi^*(t),\ A^{n-1}Be^{(i)}\rangle = 0,\ \hat{T} \leqslant t \leqslant T. \tag{10.74}$$

Равенства (10.72) — (10.74) означают, что вектор $\psi^*(t)$ ортогонален каждому вектору системы векторов (10.66) при всех $t \in [\hat{T}, T]$. Но данная система векторов образует базис в пространстве $\mathbf{R}^n$; следовательно, $\psi^*(t) \equiv 0_n, \hat{T} \leqslant t \leqslant T$. Это противоречит тому, что ψ^* — ненулевое решение системы (10.61). ■

Замечание. При доказательстве теоремы не было использовано условие (10.63), поскольку это неравенство в рассматриваемой задаче выполняется автоматически. Действительно, из (10.71) при $t=T$, так как $y^{(1)} = y^*(T) = 0_n$, получаем

$$\langle \psi^*(T),\ Ay^*(T) + Bu^*(T)\rangle \geqslant 0,$$

т. е. неравенство (10.63) выполняется в конечный момент времени $t=T$. Но выполнение условия 4) теоремы 10.1 при $t=T$ влечет его справедливость для всех $t \in [0, T]$ (см. п. 1, § 10.2). Поэтому из полученного вытекает неравенство (10.63). Таким образом, для рассматриваемой задачи условие (10.63) из формулировки принципа максимума можно исключить, так как оно является следствием остальных условий.

3. Существование, единственность и структура оптимального управления. При решении задач оптимального управления важно заранее знать, что найденная с помощью принципа максимума экстремаль Понтрягина (т. е. управление, удовлетворяющее всем условиям принципа максимума) единственна и, кроме того, что оптимальное управление существует. В таком случае процедура поиска оптимального управления упрощается: нужно найти какую-нибудь экстремаль Понтрягина; она и будет оптимальным управлением. В следующем ниже утверждении [5, 20] сформулированы условия, гарантирующие единственность экстремали Понтрягина и существование оптимального управления в линейной задаче оптимального быстродействия; кроме того, раскрывается структура оптимального управления.

Теорема 10.6. *Будем считать, что выполнены все предположения теоремы* 10.5 *и, кроме того, начало координат* 0_m *является внутренней точкой параллелепипеда* U *вида* (10.65) (*последнее заведомо выполняется, если* $\alpha_i < 0$, $\beta_i > 0$, $i = 1, 2, \dots, m$).

I. *Пусть для* $y^{(0)} \in \mathbf{R}^n$ *существует некоторое допустимое управление, переводящее систему* (10.60) *из этой точки в начало координат* 0_n. *Тогда существует единственное допустимое управление* u^*, *которое переводит систему из* $y^{(0)}$ *в* 0_n *и удовлетворяет условиям принципа максимума; это управление* u^* *является оптимальным. Каждая компонента вектор-функции* u^* *представляет собой кусочно-постоянную функцию, принимающую значения* α_i *или* β_i *(см. равенство* (10.65)). *При этом если все собственные числа матрицы* A *оказываются вещественными, то компоненты вектор-функции* u^* *имеют не более чем* n *промежутков постоянства.*

II. *Пусть вещественные части всех собственных чисел матрицы* A *отрицательны. Тогда для каждого начального состояния* $y^{(0)} \in \mathbf{R}^n$ *существует единственное допустимое управление* u^*, *обладающее всеми свойствами, указанными в первой части теоремы.*

В первой части теоремы утверждается, что если систему (10.60) из точки $y^{(0)}$ можно перевести в начало координат 0_n, то из $y^{(0)}$ возможен и оптимальный в смысле быстродействия переход в начало координат. При этом оптимальное управление может быть найдено с помощью принципа максимума: отыскав какую-нибудь экстремаль Понтрягина u^*, переводящую систему из $y^{(0)}$ в 0_n, можно быть уверенным в том, что других, отличных от u^* экстремалей нет и, кроме того, что эта экстремаль является оптимальным управлением. Более того, в теореме описана структура оптимального управления — это вектор-функция с кусочно-постоянными компонентами, имеющими конечное число точек разрыва. Эти точки разрыва называют *точками переключения* управления. В них происходит скачкообразное изменение какой-нибудь компоненты управления от наименьшего возможного до наибольшего возможного значения или наоборот. Например, i-я компонента может изменять свое значение с α_i на β_i или же с β_i на α_i. Далее, если все собственные числа матрицы A вещественны, то, как утверждает теорема, у каждой компоненты вектор-функции u^* точек переключения не может быть больше чем $n-1$.

При выполнении всех предположений первой части теоремы в фазовом пространстве $\mathbf{R}^n$ могут быть точки, из которых перевести систему (10.60) в начало координат невозможно. Однако если вещественные части всех собственных чисел матрицы A отрицательны, то в силу второй части теоремы из любой точки фазового пространства система может быть переведена в начало координат оптимальным в смысле быстродействия образом. При этом оптимальное управление находят также с помощью принципа максимума и оно имеет указанную выше структуру.

§ 10.5. ПРИКЛАДНЫЕ ЗАДАЧИ

1. Задача на минимум критических размеров ядерного реактора на тепловых нейтронах. Схематически представим ядерный реактор в виде пространственной области Ω, которая заполнена замедлителем. Распределение замедлителя считаем однородным. В области Ω произвольным образом распределены атомы урана (ядерное горючее), которые при ядерных превращениях испускают нейтроны. Одни из этих нейтронов сталкиваются с ядрами урана, вызывая их расщепление, другие же при движении в замедлителе теряют значительную часть своей энергии и в дальнейшем либо поглощаются ядрами урана (не вызывая их расщепления), либо продолжают перемещаться в замедлителе. Нейтроны последнего типа называются тепловыми.

Установившийся процесс диффузии тепловых нейтронов в ядерном реакторе описывается уравнением

$$\Delta n(M)+\alpha^2 n(M)=0, \tag{10.75}$$

где $M\in\Omega$, $n(M)$ — плотность тепловых нейтронов в точке M внутри реактора. Согласно физическому смыслу, $n(M)\geqslant 0$ при $M\in\Omega$ и $n(M)$ должна быть ограничена. Величину α^2 называют лапласианом или материальным параметром. Рассмотрим задачу на минимум критических размеров простейшего реактора, имеющего форму шара. Учитывая сферическую симметрию задачи, поместим начало координат в центр шара. Уравнение (10.75) в сферических координатах имеет вид

$$\frac{1}{r^2}\frac{\mathrm{d}}{\mathrm{d}r}\left(r^2\frac{\mathrm{d}n}{\mathrm{d}r}\right)+\alpha^2 n=0. \tag{10.76}$$

Предположим, что за пределами области Ω, занимаемой реактором, находится вакуум, т. е. тепловые нейтроны беспрепятственно уходят с границы S области Ω. Поэтому граничное условие для уравнения (10.76) имеет вид

$$n(R)=0, \tag{10.77}$$

где R — радиус шарового реактора. Условие (10.77) является выражением непрерывности функции $n(M)$ во всем пространстве.

Параметр α^2 в уравнениях (10.75) и (10.76) зависит от коэффициента размножения нейтронов. Этот коэффициент, в свою очередь, пропорционален концентрации источников нейтронов (горючего в реакторе), поэтому можно положить $\alpha^2\equiv u(r)$, где функция $u(r)$ пропорциональна концентрации горючего в точке r. Концентрация горючего в реакторе — положительная ограниченная величина, поэтому на функцию $u(r)$ необходимо наложить ограничение

$$0\leqslant u(r)\leqslant u_{\max}. \tag{10.78}$$

Теперь можно сформулировать задачу оптимизации. Она состоит в отыскании такого распределения горючего в реакторе, т. е. такой функции $u(r)$, удовлетворяющей условию (10.78), чтобы при наличии нетривиального решения уравнения (10.76) реактор имел минимально возможный размер, т. е. функционал

$$I(u)=\int_0^R \mathrm{d}r=R(u) \tag{10.79}$$

имел минимальное значение. Формально это задача на «быстродействие», поскольку функционал имеет вид $I=T$, где $t\equiv r$, $T=R$. Уравнение (10.75), описывающее распределение нейтронов, не является линейным, так как содержит произведение $\alpha^2 n \equiv u(r)\,n(r)$. Для нелинейных систем в общем случае нет универсального метода решения задачи оптимального быстродействия, подобного методу, изложенному в § 10.4. Однако рассматриваемая задача имеет простой вид и ее можно решить с помощью тех же рассуждений, которые привели к общему алгоритму в линейном случае. При этом «управляющая» (в данном случае нет процесса, развивающегося во времени, поэтому этот термин можно использовать только условно) функция $u(r)$, как и в линейном случае, является кусочно-постоянной со значениями 0 или $u_{\max}$.

Введем замену $rn=x$, тогда (10.76) принимает вид

$$\frac{d^2 x(r)}{dr^2}+u(r)\,x(r)=0. \tag{10.80}$$

Полагая $x=x_1$ и $\dfrac{dx}{dr}=x_2$, получаем систему уравнений

$$\frac{dx_1}{dr}=x_2,\ \frac{dx_2}{dr}=-ux_1. \tag{10.81}$$

Граничные условия для этой системы имеют вид

$$x_1(R)=0,\ x_1(0)=0. \tag{10.82}$$

Первое из этих условий следует из (10.77), а второе — из условия ограниченности $u(r)$ при $r=0$.

Таким образом, получена стандартная задача оптимального управления по быстродействию для уравнений (10.81) с граничными условиями (10.82). Составим функцию, участвующую в равенстве (10.64):

$$\hat{H}=\langle\psi(t),\ Bu(t)\rangle=x_2\psi_1-ux_1\psi_2, \tag{10.83}$$

где ψ_1 и ψ_2 удовлетворяют соответственно уравнениям

$$\frac{d\psi_1}{dr}=u\psi_2,\quad \frac{d\psi_2}{dr}=-\psi_1. \tag{10.84}$$

Граничные условия (10.82) относятся только к одной фазовой переменной x_1, поэтому задача имеет подвижные концы (x_2 в начальной и конечной точке произвольно). Условия трансверсальности в этом случае определяют граничные условия для функции $\psi_2(r)$

$$\psi_2(R)=0,\ \psi_2(0)=0. \tag{10.85}$$

Из уравнений (10.84) следует, что функция ψ_2 удовлетворяет уравнению

$$\frac{d^2\psi_2}{dr^2}=-u\psi_2.$$

Такому же уравнению удовлетворяет и функция $x_1(r)$. Учитывая, что и функции $x_1(r)$ и $\psi_2(r)$ имеют нулевые граничные условия, устанавливаем, что они отличаются только множителем $\psi_2(r) = Cx_1(r)$, где $C=\text{const}$. При этом функция $\hat{H}$ принимает вид

$$\hat{H} = x_2\psi_1 - Cux_1^2.$$

Поскольку величина Cx_1^2 положительна при всех $r \in [0, R]$, для того чтобы функция $\hat{H}$ принимала максимальные значения при каждом r, необходимо положить $u(r) \equiv 0$, если $C>0$, и $u(r) \equiv u_{\max}$, если $C<0$. Случай $u(r) \equiv 0$ приводит к тривиальному решению $x(r) \equiv 0$, которое означает отсутствие распада ядер урана в реакторе. Рассмотрим случай $u(r) \equiv u_{\max}$, $0 \leqslant r \leqslant R$. Из (10.81) найдем

$$x_1(r) = C_1 \sin\sqrt{u_{\max}}\,r, \quad C_1 > 0.$$

Постоянная C_1 не может быть конкретизирована в рамках рассмотренной простейшей стационарной модели реактора. Уточнить величину C_1 можно с помощью нестационарной модели, описывающей процесс нарастания плотности нейтронов в реакторе от начального нулевого значения до значения, соответствующего стационарному режиму.

Величину R определяют из граничного условия $x_1(R) = 0$. Чтобы это условие было выполнено, необходимо положить

$$R = \pi/\sqrt{u_{\max}}. \tag{10.86}$$

Величина R и является минимально возможной.

Полученный результат (для обеспечения минимальности размеров реактора необходимо взять во всем объеме реактора максимально возможную плотность горючего) тривиален с физической точки зрения.

2. **Задача на минимум критических размеров ядерного реактора при заданной общей загрузке топлива в реакторе.** По физическому смыслу эта задача аналогична предыдущей. Необходимо найти вид распределения горючего в реакторе, при котором в реакторе идет ядерная реакция $(n(r) \equiv 0)$ и радиус реактора минимален. Основное отличие состоит в том, что общее количество топлива, загружаемого в реактор, считают заданным. Этот факт можно выразить с помощью условия

$$\int_0^R u(r)\,\mathrm{d}r = I_0 = \text{const}. \tag{10.87}$$

Кроме того, для придания задаче большей общности заменим ограничения (10.78) на следующие:

$$u_{\min} \leqslant u(r) \leqslant u_{\max}. \tag{10.88}$$

Формально рассматриваемая задача по-прежнему является задачей на быстродействие с функционалом (10.79), ограничениями (10.88) на значения управляющей функции и интегральным ограничением

(10.87). Чтобы задача имела физический смысл, необходимо наложить ограничения на возможное значение величины I_0 в (10.87). В предыдущей задаче общее количество топлива в реакторе определялось величиной

$$I = \int_0^{\pi/\sqrt{u_{\max}}} u_{\max}\, \mathrm{d}r = \pi \sqrt{u_{\max}}. \tag{10.89}$$

Очевидно, ограничение (10.87) связано с необходимостью (для увеличения экономической эффективности работы реактора) уменьшить общее количество используемого топлива по сравнению со случаем, когда $u(r) \equiv u_{\max}$. Поэтому оптимизационная задача с ограничением (10.87) на величину общей загрузки реактора имеет смысл только если I_0 меньше, чем общая загрузка реактора в случае $u(r) \equiv u_{\max}$, т. е. если

$$I_0 < \pi \sqrt{u_{\max}}. \tag{10.90}$$

Будем решать задачу с помощью принципа максимума. Чтобы привести ее к стандартному виду, введем дополнительную фазовую переменную $x_3(r)$ с помощью уравнения

$$\frac{\mathrm{d}x_3}{\mathrm{d}r} = u. \tag{10.91}$$

Тогда ограничение (10.87) можно учесть, вводя граничные условия для $x_3(r)$ в виде

$$x_3(0) = 0, \ x_3(R) = I_0. \tag{10.92}$$

Фазовые переменные $x_1(r)$ и $x_2(r)$ те же, что и в первой задаче. Объединяя уравнения (10.81) и (10.91), получаем, что уравнения, описывающие распределение нейтронов в реакторе, имеют вид

$$\frac{\mathrm{d}x_1}{\mathrm{d}r} = x_2, \ \frac{\mathrm{d}x_2}{\mathrm{d}r} = u x_1, \ \frac{\mathrm{d}x_3}{\mathrm{d}r} = u. \tag{10.93}$$

В стандартной постановке задача оптимизации ядерного реактора выглядит следующим образом: требуется найти функцию $u(r)$, $0 \leqslant r \leqslant R$, удовлетворяющую условию (10.88), такую, что соответствующая траектория $x(r) = (x_1(r), x_2(r), x_3(r))$ в фазовом пространстве R^3 проходит при $r=0$ через прямую, определяемую условиями

$$x_1(0) = 0, \ x_3(0) = 0, \tag{10.94}$$

а при $r = R$ — через прямую, которая определяется условиями

$$x_1(R) = 0, \ x_3(R) = I_0. \tag{10.95}$$

При этом функционал

$$I(u) = \int_0^R \mathrm{d}r = R(u)$$

должен принять минимально возможное (среди таких функций $u(r)$) значение.

Это задача оптимального быстродействия с подвижными концами, так как на обоих концах фазовой траектории системы координата x_2 является произвольной. Поскольку второе из уравнений (10.93) нелинейно, в данном случае, вообще говоря, не справедливы теоремы, установленные для оптимальных быстродействий (см. § 10.4), однако метод решения оптимизационной задачи тот же, что и в линейном случае.

Чтобы применить принцип максимума, введем функцию $\hat{H}$:

$$\hat{H}=\psi_1 x_2-u\psi_2 x_1+u\psi_3, \tag{10.96}$$

где функции ψ_1, ψ_2, ψ_3 удовлетворяют уравнениям

$$\frac{d\psi_1}{dr}=u\psi_2,\ \frac{d\psi_2}{dr}=-\psi_1,\ \frac{d\psi_3}{dr}=0. \tag{10.97}$$

Из условий трансверсальности вытекают граничные условия для функций ψ_2:

$$\psi_2(0)=0,\ \psi_2(R)=0. \tag{10.98}$$

Для нахождения оптимального управления $u(r)$ преобразуем $\hat{H}$ к более удобному виду. Сравнивая уравнения для x_1, x_2 и ψ_1, ψ_2 и граничные условия для x_1 и ψ_2, получаем, что имеет место одно из соотношений

$$\psi_2(r)=-x_1(r),\ \psi_1(r)=x_2(r) \tag{10.99}$$

или

$$\psi_2(r)=x_1(r),\ \psi_1(r)=-x_2(r). \tag{10.100}$$

В соответствии с этим из (10.96) имеем

$$\hat{H}=\psi_1 x_2+u(-\psi_2 x_1+\psi_3)=\psi_1 x_2+u\left[\pm x_1^2+\psi_3\right], \tag{10.101}$$

где знак плюс соответствует случаю (10.99), а знак минус — случаю (10.100). Из (10.101) нетрудно получить, что $\hat{H}$ будет при каждом r максимальным, если $u(r)=u_{\max}$ при $\varphi(x_1, \psi_3)=\pm x_1^2+\psi_3>0$ и $u(r)=u_{\min}$ при $\varphi(x_1, \psi_3)>0$. Функцию $\varphi(x_1, \psi_3)$ называют функцией переключения, ее нули определяют точки, в которых $u(r)$ испытывает скачок от $u_{\max}$ к $u_{\min}$ или от $u_{\min}$ к $u_{\max}$. Функция $u(r)$ является кусочно-постоянной. Чтобы выяснить более конкретный вид этой функции, необходимо исследовать функцию переключения $\varphi(x_1, \psi_3)$.

Количество точек, где функция φ изменяет знак, зависит от знака x_1 и от значения постоянной ψ_3 (то, что $\psi_3=\mathrm{const}$, следует из третьего уравнения (10.97)). Учитывая, что график функций $x_1(r)$ имеет синусоидальный вид и $x_1(0)=x(R)=0$, получаем, что либо на интервале $0<r<R$ вообще нет точек, где функция $\varphi(x_1, \psi_3)$ изменяет знак, либо имеются две такие точки.

В первом случае получим такое же решение, как и в первой задаче: $u(r)\equiv u_{\max}$, $0\leqslant r\leqslant R$. Однако в настоящей задаче такое $u(r)$

не является оптимальным, поскольку не удовлетворено граничное условие $x_3(R)=I_0$. Действительно, из (10.89) и (10.91) следует, что при $u(r)\equiv u_{\max}$ имеем $x_3(R)=\pi\sqrt{u_{\max}}>I_0$.

Таким образом, функция переключения обязательно должна иметь две точки перемены знака. При этом возможны следующие два варианта:

а) $\psi_3<0$ и выполнено соотношение (10.99), т. е. в выражении (10.101) перед x_1^2 стоит знак $+$. Учитывая граничные условия (10.94), получаем $\varphi(x_1, \psi_3)<0$ при $0<r<r_1$, далее $\varphi>0$ при $r_1<r<r_2$ и снова $\varphi<0$ при $r_2<r<R$. Здесь r_1, r_2 — точки внутри отрезка $[0, R]$. Из принципа максимума следует, что оптимальное управление $u(r)$ имеет следующий вид: $u(r)=u_{\min}$ при $0<r<r_1$, $u(r)=u_{\max}$ при $r_1<r<r_2$ и $u(r)=u_{\min}$ при $r_2<r<R$;

б) $\psi_3>0$ и выполнено соотношение (10.100), т. е. в выражении (10.101) перед x_1^2 стоит знак «—». Здесь $\varphi>0$ при $0<r<r_1$, далее $\varphi<0$ при $r_1<r<r_2$ и снова $\varphi>0$ при $r_2<r<R$ (величины r_1 и r_2 имеют другие значения по сравнению с первым случаем). Из принципа максимума следует: $u(r)=u_{\max}$ при $0<r<r_1$, $u(r)=u_{\min}$ при $r_1<r<r_2$ и $u(r)=u_{\max}$ при $r_2<r<R$.

Остается определить точки r_1 и r_2, в которых функция изменяет знак, и установить, когда каждый из вариантов является истинно оптимальным, т. е. позволяет получить наименьшее значение радиуса реактора R.

Для определения r_1 и r_2 необходимо решить уравнения (10.93) для фазовых переменных $x_1(r)$ и $x_2(r)$. Чтобы учесть сразу случаи а) и б), положим: $u(r)=u_1$ при $0<r<r_1$, $u(r)=u_2$ при $r_1<r<r_2$ и $u=u_1$ при $r_2<r<R$. В случае а) имеем $u_1=u_{\min}$, $u_2=u_{\max}$; в случае б) получаем $u_1=u_{\max}$, $u_2=u_{\min}$. Для простоты положим $|\psi_3|=1$. Тогда из уравнения

$$\frac{d^2x_1}{dr^2}=ux_1$$

и граничных условий $x_1(0)=0$, $x_1(r_1)=1$, $x_1(r_2)=1$, $x_1(R)=0$ можно найти вид $x_1(r)$ в каждой области реактора:

$$x_1(r)=\frac{\sin u_1^{1/2}r}{\sin u_1^{1/2}r_1} \text{ при } 0<r<r_1; \tag{10.102}$$

$$x_1(r)=\frac{\cos u_2^{1/2}(r_2-r)}{\cos u_2^{1/2}(r_2-r_1)} \text{ при } r_1<r<r_2; \tag{10.103}$$

$$x_1(r)=\frac{\sin n_1^{1/2}(R-r)}{\sin u_1^{1/2}(R-r_2)} \text{ при } r_2<r<R. \tag{10.104}$$

Поскольку из (10.93) следует, что фазовая переменная $x_2(r)$ является производной по r от $x_1(r)$, получаем:

$$x_2(r)=\frac{u_1^{1/2}\cos u_1^{1/2}r}{\sin u_1^{1/2}r_1} \text{ при } 0<r<r_1; \tag{10.105}$$

$$x_2(r)=\frac{u_2^{1/2}\sin u_2^{1/2}(r_2-r)}{\cos u_2^{1/2}(r_2-r_1)} \text{ при } r_1<r<r_2; \qquad (10.106)$$

$$x_2(r)=-\frac{u_1^{1/2}\cos u_1^{1/2}(R-r)}{\sin u_1^{1/2}(R-r_2)} \text{ при } r_2<r<R. \qquad (10.107)$$

Теперь величины r_1, r_2 и R можно определить из трех алгебраических уравнений: двух уравнений, определяющих непрерывность функции $x_2(r)$ в точках r_1 и r_2 (непрерывность функции $x_1(r)$ следует из непрерывности функции $x_2(r)$, а непрерывность функции $x_3(r)$ следует из третьего уравнения (10.93)), и соотношения (10.87). Из условия непрерывности $x_2(r)$ в точках r_1 и r_2 получаем

$$u_1^{1/2}\operatorname{ctg} u_1^{1/2}r_1=u_2^{1/2}\operatorname{tg} u_2^{1/2}(r_2-r_1); \qquad (10.108)$$

$$0=-u_1^{1/2}\operatorname{ctg} u_1^{1/2}(R-r_2). \qquad (10.109)$$

Соотношение (10.87) для рассматриваемого трехфазного распределения ядерного горючего реактора имеет вид

$$I_0=u_1r_1+u_2(r_2-r_1)+u_1(R-r_2). \qquad (10.110)$$

Из (10.109) следует, что

$$R-r_2=\frac{\pi}{2u_1^{1/2}}. \qquad (10.111)$$

Принимая во внимание (10.111), из соотношения (10.110) имеем

$$r_2-r_1=\frac{I_0-u_1r_1-\frac{1}{2}\pi u_1^{1/2}}{u_2}. \qquad (10.112)$$

Подставляя r_2-r_1 в правую часть соотношения (10.108), получаем уравнение для определения r_1:

$$u_1^{1/2}\operatorname{ctg} u_1^{1/2}r_1=u_2^{1/2}\operatorname{tg} u_2^{1/2}\,\frac{I_0-u_1r_1-\frac{1}{2}\pi u_1^{1/2}}{u_2}. \qquad (10.113)$$

Из этого уравнения можно найти r_1 (с помощью численных методов). Затем из (10.112) получаем r_2, а из (10.111) — оптимальное значение величины R. Условий, из которых можно было бы определить, какой из вариантов является истинно оптимальным, нет. Поэтому в каждом конкретном случае необходимо непосредственно вычислять R как в случае а), так и в случае б) и сравнивать полученные значения.

Рассмотрим конкретный пример. Пусть $u_{\min}=\frac{1}{4}u_{\max}$ и $I_0=\frac{1}{2}\pi u_{\max}^{1/2}$. Вэтом случае нет необходимости вычислять значение r_1.

Воспользуемся для варианта б) формулой (10.112):

$$r_2=\frac{I_0-u_1 r\gamma-\frac{1}{2}\pi u_1^{1/2}}{u_2}+r_1=\frac{\frac{1}{2}\pi u_{\max}^{1/2}-u_{\max}r_1-\frac{1}{2}\pi u_{\max}^{1/2}}{\frac{1}{4}u_{\max}}+$$

$$+r_1=-3r_1.$$

Очевидно, что полученный результат не имеет смысла, и можно утверждать, что оптимальное решение дает вариант а). Используя для этого варианта формулу (10.112), имеем

$$r_2=\frac{\frac{1}{2}\pi u_{\max}^{1/2}-\frac{1}{4}u_{\max}r_1-\frac{1}{4}\pi u_{\max}^{1/2}}{u_{\max}}+r_1=\frac{3}{4}r_1+\frac{1}{4}\pi u_{\max}^{-1/2}.$$

Из формулы (10.111) следует, что

$$R=\frac{\pi}{2u_1^{1/2}}+r_2=\frac{\pi}{u_{\max}^{1/2}}+r_2=\frac{3\pi}{4u_{\max}^{1/2}}+\frac{3}{4}r_1,$$

где r_1 находится из уравнения (10.113), которое принимает вид

$$\frac{1}{2}\operatorname{ctg}\left(\frac{1}{2}u_{\max}^{1/2}r_1\right)=\operatorname{tg}\left(\frac{1}{4}\pi-\frac{1}{4}u_{\max}^{1/2}r_1\right).$$

3. Задача об оптимальном распределении температуры вдоль проходной печи, предназначенной для нагрева металла. Металл непрерывной лентой входит в печь и движется с постоянной скоростью v. На выходе из печи металл имеет постоянную температуру θ_0. В общей постановке задача оптимизации состоит в выборе такого распределения температуры $T(x)$, $0\leqslant x\leqslant L$, стенок печи, чтобы на выходе металл был нагрет до заданной температуры θ_L и при этом количество топлива, затраченного на нагрев, было минимальным.

Прежде всего дадим строгую математическую постановку задачи оптимизации. Для этого выведем уравнение, описывающее процесс нагрева металла внутри печи, и установим вид функционала, который задает расход топлива. Распределение температуры $T(x)$ стенок печи будем считать стационарным. Входная температура металла θ_0 также неизменна во времени, поэтому распределение температуры θ металла по длине печи будет стационарным: $\theta=\theta(x)$. Рассмотрим тепловой баланс малого элемента $\mathrm{d}V$ объема металла, движущегося со скоростью v внутри печи. Согласно закону сохранения энергии, $\mathrm{d}Q=\mathrm{d}u$, где $\mathrm{d}Q$ — теплота, переданная от стенок печи к элементу металла за время $\mathrm{d}t$, $\mathrm{d}u$ — приращение внутренней энергии элемента металла. Эти приращения можно выразить следующим образом:

$$\mathrm{d}Q=[T(x)-\theta(x)]\,k\,\mathrm{d}t\,\mathrm{d}m,$$

$$\mathrm{d}u=c\,\mathrm{d}\theta(x)\,\mathrm{d}m,$$

где k — коэффициент теплопередачи от стенки к металлу; c — теплоемкость металла; $\mathrm{d}m$ — масса рассматриваемого элемента объема металла; $\mathrm{d}\theta(x)$ — приращение температуры рассматриваемого элемента. Приравнивая правые части этих соотношений и используя равенство $\mathrm{d}t=\frac{\mathrm{d}x}{v}$, получаем

$$\mathrm{d}\theta\,\mathrm{d}m=(T-\theta)\,k\,\frac{\mathrm{d}x}{v}\,\mathrm{d}\,m,$$

откуда

$$h\,\frac{\mathrm{d}\theta}{\mathrm{d}x}=T-\theta. \tag{10.114}$$

Здесь $h=cv/k$ (h имеет размерность длины). Полученное уравнение описывает распределение $\theta(x)$ температуры металла в печи, $0\leqslant x\leqslant L$.

Для получения функционала, выражающего расход топлива, используем формулу для приращения энтальпии I (теплосодержания единицы массы металла):

$$\mathrm{d}\,I(x)=c\,\mathrm{d}\theta(x)=\eta(x)R\,\mathrm{d}\,G(x), \tag{10.115}$$

где $I(x)$ — энтальпия единицы массы металла в точке x; R — калорийность расходуемого топлива; $\eta(x)$ — коэффициент полезного использования топлива; $G(x)$ — расход топлива на всей длине печи от точки входа до точки x. Коэффициент полезного использования топлива η зависит от температуры $T(x)$, при которой идет нагрев, и определяется по формуле

$$\eta(x)=\chi\left[1-\frac{T(x)}{T^*}\right], \tag{10.116}$$

где χ — температурный коэффициент полезного использования топлива; T^* — характерная температура, имеющая смысл максимально возможной для данного топлива температуры нагрева ($\eta(x)=0$ при $T(x)=T^*$, т. е. использование топлива приводит к нулевому тепловому эффекту). Таким образом, температура $T(x)$ должна удовлетворять следующему условию:

$$0\leqslant T(x)\leqslant T^*. \tag{10.117}$$

Подставляя (10.116) в (10.115) и интегрируя по x от 0 до L, находим, что общий расход топлива в печи G можно вычислить следующим образом:

$$G=\int_0^L\frac{c\,\mathrm{d}\,\theta(x)}{R\chi\left[1-\frac{T(x)}{T^*}\right]}.$$

Используя уравнение (10.114), преобразуем эту формулу к виду

$$G=\int_0^L\frac{\beta\,[T(x)-\theta(x)]}{T^*-T(x)}\,\mathrm{d}\,x, \tag{10.118}$$

где $\beta=T^*c/(\chi R)$.

Теперь перейдем к строгой постановке задачи оптимизации. Необходимо найти функцию $T(x)$, удовлетворяющую условию (10.117) и такую, чтобы соответствующая фазовая траектория $\theta(x)$, $0 \leqslant x \leqslant L$, определяемая из уравнения (10.114), с граничным условием

$$\theta(x) \mid_{x=0} = \theta_0, \tag{10.119}$$

удовлетворяла также условию

$$\theta(x) \mid_{x=L} = \theta_L. \tag{10.120}$$

При этом необходимо, чтобы функционал (10.118) принял наименьшее возможное значение. Решим эту задачу с помощью принципа максимума.

Запишем уравнение (10.114) в следующем виде:

$$\frac{d\theta}{dx} = \frac{1}{h}(T - \theta). \tag{10.121}$$

Это фазовое уравнение системы. В соответствии с формулой (2.3) имеем

$$H = \psi_0 \beta \frac{T - \theta}{T^* - T} + \psi \frac{1}{h}(T - \theta). \tag{10.122}$$

Согласно принципу максимума, $\psi_0 \leqslant 0$, а функция ψ определяется из уравнения

$$\frac{d\psi}{dx} = -\frac{\partial H}{\partial \theta} = \frac{\psi_0 \beta}{T^* - T} + \frac{\psi}{h}.$$

Если $\psi_0 = 0$, то, как отмечалось в § 10.2, имеет место особая задача оптимального управления. В данной задаче случай $\psi_0 = 0$ не имеет физического смысла и его из дальнейшего рассмотрения можно исключить. Полагая $\psi_0 = -1$, получаем

$$H = \beta \frac{\theta - T}{T^* - T} + \psi \frac{1}{h}(T - \theta), \tag{10.123}$$

где функция $\psi(x)$ удовлетворяет уравнению

$$\frac{d\psi}{dx} = -\frac{\beta}{T^* - T} + \frac{\psi}{h}. \tag{10.124}$$

Теперь найдем функцию $T(x)$, которая при каждом x максимизирует функцию H. Это — сложная задача, поскольку учитывается условие (10.117). Однако можно внести некоторые естественные с физической точки зрения упрощающие предположения. В действительности предельная температура T^*, при которой тепловой эффект использования топлива сводится к нулю, велика. Поэтому ограничение $T(x) \leqslant T^*$ лишнее, так как все реальные режимы работы печи ему удовлетворяют. Ограничение $T(x) \geqslant 0$ также лишено содержания, так как случай $T(x) < 0$ в данной задаче не имеет физического смысла (предполагаем, конечно, что $\theta_0 < \theta_L$). Таким образом, можно считать, что при каждом x значение $T(x)$, реализующее максимум функции H, находится внутри отрезка $[0, T^*]$. В этом случае величину $T(x)$, максимизирующую функцию H, можно найти, приравнивая нулю частную производную от H по T:

$$\frac{\partial H}{\partial T}=\beta\frac{-T^*+\theta}{(T^*-T)^2}+\frac{\psi}{h}=0.$$

Решая это уравнение относительно T, получаем

$$T=T^*\pm\sqrt{-\frac{h\beta(-T^*+\theta)}{\psi}}.$$

Согласно сделанному выше предположению, $T\leqslant T^*$, поэтому

$$T=T^*-\sqrt{\frac{h\beta(T^*-\theta)}{\psi}}. \tag{10.125}$$

Проверим, что T реализует именно максимум (а не минимум) функции H. Вычислим вторую производную от H по T:

$$\frac{\partial^2 H}{\partial T^2}=\frac{2\beta(-T^*+\theta)}{(T^*-T)^3}.$$

Так как $T^*\geqslant T$ и $T^*\geqslant\theta$, то $\partial^2H/\partial T^2<0$, т. е. функция (10.125) действительно максимизирует функцию H.

С помощью (10.125) можно записать замкнутую систему уравнений для определения функций $\theta(x)$ и $\psi(x)$, соответствующих оптимальному режиму. Подставив (10.125) в (10.121) и (10.124), получим систему уравнений:

$$\frac{d\theta}{dx}=\frac{1}{h}(T^*-\theta)-\sqrt{\frac{\beta(T^*-\theta)}{h\psi}}, \tag{10.126}$$

$$\frac{d\psi}{dx}=-\sqrt{\frac{\beta\psi}{h(T^*-\theta)}}+\frac{\psi}{h}. \tag{10.127}$$

Граничными для этой системы являются условия (10.119), (10.120). Уравнения (10.126) и (10.127) нелинейные. Нелинейные уравнения редко удается решить аналитически, однако в данном случае это возможно, так как уравнение для функции $\theta(x)$ можно свести к линейному. Умножим уравнение (10.126) на ψ, а уравнение (10.127) — на $T^*-\theta$ и вычтем почленно второе уравнение из первого. Получим

$$\psi\frac{d\theta}{dx}-(T^*-\theta)\frac{d\psi}{dx}=0,$$

или

$$\frac{d}{dx}[\psi(T^*-\theta)]=0.$$

Отсюда непосредственно имеем

$$\psi(T^*-\theta)=\gamma=\text{const}. \tag{10.128}$$

Из выражения (10.125) для функции T следует, что функции ψ и $T^*-\theta$ имеют одинаковый знак во всех точках $x\in[0, L]$ (иначе подкоренное выражение будет отрицательным). Поэтому в (10.128) $\gamma>0$.

С помощью функции (10.128) уравнение (10.126) можно свести к линейному. Подставив $\psi=\gamma/(T^*-\theta)$, получим

$$\frac{d\theta}{dx}=\frac{1}{h}(T^*-\theta)-(T^*-\theta)\sqrt{\beta/(h\gamma)},$$

или

$$\frac{d\theta}{dx}=T^*\left(\frac{1}{h}-\sqrt{\beta/(h\gamma)}\right)-\theta\left(\frac{1}{h}-\sqrt{\beta/(h\gamma)}\right).$$

Запишем решение линейного уравнения:

$$\theta(x)=T^*+C_1 e^{-\left(\frac{1}{h}-C_2\sqrt{\beta/h}\right)x}, \tag{10.129}$$

где $C_2=1/\sqrt{\gamma}$. Постоянные C_1 и C_2 определим из граничных условий (10.119) и (10.120), из которых следуют уравнения

$$T^*+C_1=\theta_0,$$

$$T^*+C_1 e^{-\left(\frac{1}{h}-C_2\sqrt{\beta/h}\right)L}=\theta_L.$$

Выразим из первого уравнения $C_1=\theta_0-T^*$ и подставим это выражение во второе уравнение:

$$(\theta_0-T^*)e^{-\left(\frac{1}{h}-C_2\sqrt{\beta/h}\right)L}=\theta_L-T^*.$$

Отсюда находим

$$C_2=\sqrt{\frac{1}{h\beta}}+\frac{1}{L}\sqrt{\frac{h}{\beta}}\ln\frac{\theta_L-T^*}{\theta_0-T^*}. \tag{10.130}$$

Подставив выражение (10.130) и $C_1=\theta-T^*$ в (10.129), получим

$$\theta(x)=T^*+(\theta_0-T^*)e^{\frac{x}{L}\ln\frac{\theta_L-T^*}{\theta_0-T^*}}=T^*+(\theta_0-T^*)\left(\frac{\theta_L-T^*}{\theta_0-T^*}\right)^{x/L}. \tag{10.131}$$

Выражение для оптимального распределения $T(x)$ можно получить с помощью уравнения (10.114). Подставляя в это уравнение выражение (10.131) для функции $\theta(x)$, находим

$$T(x)=h\frac{d\theta(x)}{dx}+\theta(x)=\frac{h}{L}(\theta_0'-T^*)\ln\frac{\theta_L-T^*}{\theta_0-T^*}\left(\frac{\theta_L-T^*}{\theta_0-T^*}\right)^{x/L}+$$

$$+T^*+(\theta_0-T^*)\left(\frac{\theta_L-T^*}{\theta_0-T^*}\right)^{x/L}=$$

$$=T^*+(\theta_0-T^*)\left(\frac{h}{L}\ln\frac{\theta_L-T^*}{\theta_0-T^*}+1\right)\left(\frac{\theta_L-T^*}{\theta_0-T^*}\right)^{x/L}. \tag{10.132}$$

Из выражений (10.131) и (10.132) следует, что функции $\theta(x)$ и $T(x)$ удовлетворяют условиям $0 \leqslant \theta(x) \leqslant T^*$ и $0 \leqslant T(x) \leqslant T^*$.

В заключение определим минимальное значение функционала (10.118), соответствующее оптимальному распределению (10.132) температуры вдоль печи, т. е. минимальный расход топлива, обеспечивающий нагрев металла от температуры θ_0 до температуры θ_L. Подставляя выражения (10.131), (10.132) в (10.118), вычислим

$$G_{\min} = \int_0^t \beta \frac{(\theta_0 - T^*) \frac{h}{L} \ln \Omega^{-1} \Omega^{-x/L}}{(\theta_0 - T^*) \left(\frac{h}{L} \Omega^{-1} + 1 \right) \Omega^{-x/L}} \, dx = - \frac{\beta h \ln \Omega^{-1}}{\frac{h}{L} \ln \Omega^{-1} + 1} =$$

$$= \frac{\beta h \ln \Omega}{1 - \frac{h}{L} \ln \Omega},$$

где $\Omega = (\theta_0 - T^*)/(\theta_L - T^*)$.

Подчеркнем, что данную задачу удалось решить аналитически лишь потому, что ограничение (10.117) оказалось лишним.

С другими задачами оптимизации процессов нагрева металла в различных печах можно ознакомиться в кн.: *А. Г. Бутковский, С. А. Малый, Ю. Н. Андреев.* Оптимальное управление нагревом металла.— М.: Металлургия, 1972.

Глава 11

СИНТЕЗ ОПТИМАЛЬНЫХ УПРАВЛЕНИЙ

Формализация повышенных требований к решению задачи оптимального управления, учитывающих дальнейшую техническую реализацию, приводит к постановке задачи синтеза оптимальных управлений и к понятию оптимального синтезирующего управления. В настоящее время не существует эффективного метода решения задач синтеза из-за их чрезвычайной сложности. В простых случаях это решение возможно на основе принципа максимума или метода динамического программирования, излагаемого в данной главе.

Для изучения материала главы необходимо предварительно ознакомиться с основными понятиями теории оптимального управления, изложенными в предыдущей главе.

§ 11.1. ПОНЯТИЕ СИНТЕЗИРУЮЩЕГО УПРАВЛЕНИЯ

1. Программные управления и функция Беллмана. Рассмотрим задачу оптимального управления неавтономной системой (в векторной форме)

$$\frac{dy}{dt} = f(t, y, u), \quad t_0 \leqslant t \leqslant T, \tag{11.1}$$

где начальный t_0 и конечный T моменты времени фиксированы* с дополнительным условием

$$y(t) \in S(t) \subset \mathbf{R}^n, \ t_0 \leqslant t \leqslant T, \tag{11.2}$$

которое называют *фазовым ограничением*. Будем считать, что при $t=t_0$ множество $S(t)$ состоит из одной точки, т. е.

$$S(t_0) = \{y^{(0)}\}.$$

Таким образом, задача управления заключается в переводе системы из фазовой точки $y^{(0)}$ на множество $S(T)$, причем траектория перехода в промежуточные моменты времени $t_0 < t < T$ должна удовлетворять включению $y(t) \in S(t) \subset \mathbf{R}^n$.

Допустимыми будем считать (покомпонентно) кусочно-непрерывные управления u, для которых выполнено условие

$$u(t) \in U \subset \mathbf{R}^m, \ t_0 \leqslant t \leqslant T.$$

Как и в предыдущей главе, предполагаем, что компоненты вектор-функции f из (11.1) непрерывны по совокупности переменных, причем каждое допустимое управление однозначно определяет (покомпонентно) кусочно-гладкое решение $y=y(t)$, $t_0 \leqslant t \leqslant T$, системы (11.1) с начальным условием $y(t_0)=y^{(0)}$.

В качестве критерия оптимальности, подлежащего минимизации, рассмотрим смешанный функционал

$$I(u) = \int_{t_0}^{T} f_0(t, \ y(t), \ u(t)) \mathrm{d}t + \Phi(y(T)), \tag{11.3}$$

содержащий интегральный и терминальный члены. Функции f_0 и Φ считаем непрерывными соответственно на $\mathbf{R}^{n+m+1}$ и $\mathbf{R}^n$.

Сформулированную задачу оптимального управления будем называть в дальнейшем *исходной задачей*. Она состоит в отыскании допустимого управления, минимизирующего функционал (11.3) на множестве тех допустимых управлений, у которых соответствующие им (согласно (11.1)) траектории $y=y(t)$ удовлетворяют фазовому ограничению (11.2). Как и в предыдущей главе, здесь целевой функционал минимизируется не на всем множестве допустимых управлений, а только на определенном его подмножестве. В частном случае, когда $S(t)=\mathbf{R}^n$ при всех $t \in (t_0, T)$, исходная задача превращается в задачу оптимального управления с подвижным правым концом (см. гл. 10).

На основе исходной задачи введем в рассмотрение (t, y)-задачу, в которой начальным моментом является некоторое число $t \in [t_0, T)$, а начальным состоянием — некоторый вектор $y=y(t) \in S(t)$:

$$\frac{\mathrm{d}y(\tau)}{\mathrm{d}\tau} = f(\tau, \ y(\tau), \ u(\tau)), \ t \leqslant \tau \leqslant T,$$

---

* Вводимые в этом разделе понятия сохраняют свою силу и в случае, когда T не фиксировано, но лежит в пределах $t_0 < T \leqslant t_1$ при некотором $t_1 > t_0$.

$$y(t)=y;\ y(\tau)\in S(\tau), \qquad t\leqslant\tau\leqslant T,$$

$$u(\tau)\in U,\ t\leqslant\tau\leqslant T,$$

с функционалом

$$I(u\mid t,\ y)=\int_t^T f_0(\tau,\ y(\tau),\ u(\tau))\,d\tau+\Phi(y(T)),$$

который следует минимизировать. Здесь во избежание путаницы для обозначения текущего времени использована буква τ. Для каждой (t, y)-задачи считаем выполненными все предположения, аналогичные соответствующим предположениям исходной задачи. Нетрудно видеть, что $(t_0, y^{(0)})$-задача точно совпадает с исходной задачей оптимального управления, сформулированный выше. При $t\in(t_0, T)$ и $y\in S(t)$ введенная (t, y)-задача аналогична исходной, но сформулирована применительно к некоторым начальным данным t, y, являющимся как бы «промежуточными» для исходной задачи.

Будем считать, что *для каждой указанной выше пары точек (t, y) введенная (t, y)-задача обладает оптимальным управлением* $u^*(\tau)$ (точнее было бы писать $u^*(\tau|t, y)$). Такое управление называют *оптимальным программным управлением*, поскольку функция u^* задана на всем отрезке $[t, T]$ и заранее известны («запрограммированы») значения управления во все моменты времени $t\leqslant\tau\leqslant T$. Одновременно «запрограммированными» оказываются все промежуточные состояния, через которые проходит система во время процесса управления.

В силу сделанного выше предположения о существовании оптимального программного управления каждой паре точек (t, y), где $t\in[t_0, T]$, $y\in S(t)$, можно поставить в соответствие определенное число, а именно минимальное значение функционала I на оптимальном программном управлении, т. е. $I(u^*|t, y)$. Тем самым на множестве указанных пар определена следующая функция:

$$V(t,\ y)=\min_u I(u\mid t,\ y).$$

Эту функцию называют *функцией Беллмана*. При $t=T$ и $y\in S(T)$, согласно определению функционала I в виде равенства (11.3), естественно положить $V(T, y)=\Phi(y)$.

Очевидно, $V(t_0, y^{(0)})$ — оптимальное значение функционала исходной задачи: $V(t_0, y^{(0)})=I(u^*)$.

2. Задача синтеза оптимальных управлений. Если найдены оптимальное (программное) управление u^* и соответствующая ему оптимальная траектория y^* для исходной задачи оптимального управления из п. 1, то эту задачу с математической точки зрения можно считать решенной. Однако это решение обладает существенным с точки зрения практики недостатком: оно не учитывает возможность появления внешних непредсказуемых возмущений, способных изменить «запрограммированный» ход управляемого процесса. Например, если в некоторый момент времени $\bar{t}\in[t_0, T)$ система под действием внешнего возмущения скачкообразно перешла из состояния $y^*(\bar{t})$ в некоторое другое состояние $\bar{y}\in S(\bar{t})$, $\bar{y}\neq y^*(\bar{t})$, то управление $u^*(t)$, $\bar{t}\leqslant t\leqslant T$ уже, вообще говоря, не является оптимальным для пары $(\bar{t}, \bar{y})$, а значит, оно не будет оптимальным и при $t_0\leqslant t\leqslant T$. Более того, может случиться и так, что нарушится фазовое ограничение при $\bar{t}<t\leqslant T$.

Значительно более удобным с практической точки зрения является такая управляющая функция, которая указывает значение оптимального управления для каждой пары (t, y) вида $t\in[t_0, T]$, $y\in S(t)$. Это уже функция как времени, так и состояния, т. е. $u=u(t, y)$. С ее помощью, зная текущий момент времени и текущее состояние системы, можно вычислить соответствующее оптимальное управление. В таком случае техническая реализация процесса управления осуществляется без участия человека по схеме с *обратной связью* (рис. 11.1). В момент времени t измеряется и подается на ЭВМ информация о текущем состоянии системы $y(t)$ и по формуле $u(t)=u(t, y(t))$ вычисляется оптимальное значение управления. Затем полученное значение $u(t)$ реализуется исполнительным механизмом управляемой системы. Таким образом, при любом внешнем возмущении будет автоматически устанавливаться оптимальный переходный режим.

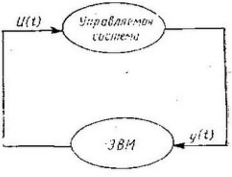

Рис. 11.1

Функцию $u=u(t, y)$, определенную для всех пар (t, y) вида $t\in[t_0, T]$, $y\in S(t)$, со значениями из множества U, будем называть *оптимальной синтезирующей функцией (оптимальным синтезирующим управлением)* для исходной задачи, если для каждой пары (t, y), $t\in[t_0, T)$, $y\in S(t)$, выполняются следующие требования:

а) система дифференциальных уравнений

$$\frac{dy}{d\tau}=f(\tau, y, u(\tau, y)),\ t\leqslant\tau\leqslant T,$$

имеет единственное (покомпонентно) кусочно-гладкое решение, которое обозначим через $\hat{y}(\tau)$, удовлетворяющее начальному условию $\hat{y}(t)=y$ и включению $\hat{y}(\tau)\in S(\tau)$, $t\leqslant\tau\leqslant T$;

б) функция вида $\hat{u}(\tau)=u(\tau, \hat{y}(\tau))$ (покомпонентно) кусочно-непрерывна на отрезке $[t, T]$;

в) функция $\hat{u}(\tau)$ является оптимальным программным управлением для (t, y)-задачи.

Первые два требования обеспечивают корректность определения синтезирующей функции. Требование в) гарантирует, что для каждой пары (t, y) значение синтезирующей функции совпадает с соответствующим значением оптимального управления.

При условии, что во время процесса управления не произойдет никаких «незапрограммированных» изменений состояния системы, оптимальное синтезирующее управление $\hat{u}$, как видно из приведенного выше определения, совпадает с некоторым оптимальным программным управлением u^*: $u^*(t, \hat{y}(t)) = u^*(t)$ для всех $t \in [t_0, T)$ Если же в некоторый момент $\bar{t} \in [t_0, T)$ под действием внешнего возмущения система перейдет из состояния $\hat{y}(\bar{t})$ в другое состояние $\bar{y}$, то использование оптимальной синтезирующей функции приведет к тому, что дальнейшее поведение системы определяет также оптимальное программное управление, но уже относительно новой пары $(\bar{t}, \bar{y})$. Тем самым после изменения состояния системы произойдет перестройка управления на оптимальный режим.

На практике измерение текущего состояния системы, вычисление значения оптимального управления и его реализацию невозможно осуществить в один и тот же момент времени t. Как правило, корректировка управления происходит с некоторым запаздыванием. Однако для широкого класса практических задач указанные действия выполняются за короткий промежуток времени, и поэтому рассматриваемая здесь математическая модель довольно точно описывает реальные процессы.

Решение задачи синтеза состоит в нахождении решения семейства (t, y)-задач при $t \in [t_0, T)$, $y \in S(t)$. В результате будет получено оптимальное программное управление $u = u(\tau | t, y)$, как функция текущего момента времени τ и зависящая от параметров t, y. Оптимальную синтезирующую функцию получают из найденного управления, отождествляя текущее время τ с начальным моментом t, т. е.

$$\hat{u} = \hat{u}(t, y) = u(\tau \mid t, y) \mid_{\tau = t}.$$

Если при этом функция $\hat{u}(t, y)$ оказывается корректно определенной в смысле требований а) и б) из определения оптимальной синтезирующей функции, то задачу синтеза можно считать решенной. Однако реализовать эту схему на практике удается лишь в редких случаях. В настоящее время нет общих методов решения задач синтеза. В некоторых случаях оптимальное синтезирующее управление удается получить с помощью принципа максимума или метода динамического программирования.

3. Стационарная синтезирующая функция. Об оптимальной синтезирующей функции говорят, что она является *стационарной*, если ее значения не зависят от текущего времени, т. е. для каждого состояния y верно равенство $u(t, y) = u(t', y)$ при всех возможных различных моментах времени t и t'.

Ожидать, что для сформулированной в п. 1 исходной задачи оптимального управления существует именно стационарная синтезирующая функция, не приходится, поскольку при разных t и t' введенные выше (t, y)- и (t', y)-задачи (при одном и том же состоянии y) различны, так как функции f, f_0, а также множество $S(t)$ явно зависят от t.

Однако даже если исходная система автономна и, кроме того $f_0 = f_0(y, u)$ и $S(t) = S$ при $t_0 < t < T$, то оптимальная синтезирующая функция, как правило, не является стационарной. Это объясняется тем, что, поскольку конечный момент времени T фиксирован, (t, y)- и (t', y)-задачи имеют различные временные промежутки управления ($T-t \neq T-t'$), а значит, и различные оптимальные программные управления, что приводит к неравенству $u(t, y) \neq \neq u(t', y)$.

§ 11.2. ДИНАМИЧЕСКОЕ ПРОГРАММИРОВАНИЕ

1. Уравнение Беллмана. Вновь рассмотрим сформулированную в п. 1 § 11.1 исходную задачу оптимального управления, заключающуюся в минимизации функционала (11.3) на подмножестве тех допустимых управлений, которым (согласно (11.1)) отвечают траектории, удовлетворяющие фазовому ограничению (11.2). При этом конечный момент времени T может быть фиксированным или нефиксированным (см. подстрочное замечание в п. 1 § 11.1). Наряду с исходной задачей в дальнейшем рассмотрении участвуют (t, y)-задача и функция Беллмана $V(t, y)$. Все предположения, сделанные в п. 1 § 11.1 относительно исходной и (t, y)-задач, здесь также считаем выполненными.

Обозначим через $\Gamma \subset \mathbf{R}^{n+1}$ область определения функции Беллмана:

$$\Gamma = \{(t, y) \in \mathbf{R}^{n+1} \mid t_0 \leqslant t \leqslant T,\ y \in S(t)\}.$$

Теорема 11.1. *Пусть (t, y) — произвольная внутренняя точка множества Γ, в которой функция Беллмана предполагается дифференцируемой. Предположим также, что сложная функция $V(t, y(t))$ дифференцируема в точке t для каждой траектории $y(t)$ системы* (11.1), *удовлетворяющей фазовому ограничению и проходящей в момент t через точку y. Тогда для всех векторов $u \in U$ функция $V(t, y)$ удовлетворяет неравенству,*

$$V_t(t, y) + \langle V_y(t, y), f(t, y, u)\rangle + f_0(t, y, u) \geqslant 0, \qquad (11.4)$$

где $V_t = \partial V/\partial t$, $V_y = (\partial V/\partial y_1, \partial V/\partial y_2, \ldots, \partial V/\partial y_n)^T$.

Если $u^*(\tau)$ — оптимальное управление в (t, y)-задаче, то функция $V(t, y)$ необходимо удовлетворяет дифференциальному уравнению в частных производных

$$\min_{u \in U} \{V_t(t, y) + \langle V_y(t, y), f(t, y, u)\rangle + f_0(t, y, u)\} = 0. \qquad (11.5)$$

При этом для каждого оптимального управления $u^*(\tau)$ минимум в левой части (11.5) *достигается на правостороннем пределе $u = u^*_+$ вида $u^*_+ = \lim\limits_{\tau \to t+0} u^*(\tau)$.*

□ Рассмотрим произвольную точку (t, y) из внутренности множества Γ и произвольный вектор $u \in U$. Допустимое управление $\bar{u}$ определим формулой

$$\bar{u}(\tau) = \bar{u}_{\Delta t}(\tau) = \begin{cases} u, & t \leqslant \tau \leqslant t + \Delta t, \\ \hat{u}(\tau), & t + \Delta t < \tau \leqslant T. \end{cases}$$

и введем соответствующую траекторию $\bar{y}(\tau)$, начинающуюся в момент времени t в точке y и удовлетворяющую фазовому ограничению $\bar{y}(\tau) \in S(\tau)$, $t \leqslant \tau \leqslant T$. Здесь $\hat{u}(\tau)$ — оптимальное программное управление в $(t+\Delta t, \bar{y}(t+\Delta t))$-задаче, а положительное число Δt мало настолько, что $(t+\Delta t, \bar{y}(t+\Delta t)) \in \Gamma$ (такое Δt существует, поскольку (t, y) — внутренняя точка множества Γ). Таким образом, в течение достаточно малого временно́го промежутка Δt система под действием постоянного управления $u(\tau)=u$ переводится по траектории $\bar{y}(\tau)$, $t \leqslant \tau \leqslant t+\Delta t$, в некоторую точку $\bar{y}(t+\Delta t)$, а затем при $\tau > t+\Delta t$ применяется оптимальное программное управление $\hat{u}(\tau)$, которое согласно сделанному ранее предположению, существует.

Вычислим значение функционала $I(u|t, y)$ на управлении $\bar{u}$:

$$I(\bar{u} \mid t, y) = \int_t^T f_0(\tau, \bar{y}(\tau), \bar{u}(\tau))\, d\tau + \Phi(\bar{y}(T)) =$$

$$= \int_t^{t+\Delta t} f_0(\tau, \bar{y}(\tau), \bar{u}(\tau))\, d\tau + \int_{t+\Delta t}^T f_0(\tau, \bar{y}(\tau), \bar{u}(\tau))\, d\tau + \Phi(\bar{y}(T)).$$

По определению функции Беллмана, последние два слагаемых представляют собой $V(t+\Delta t, \bar{y}(t+\Delta t))$ и, кроме того, $V(t, \bar{y}(t)) = V(t, y) \leqslant I(\bar{u}|t, y)$. Поэтому можно записать неравенство

$$V(t, \bar{y}(t)) \leqslant \int_t^{t+\Delta t} f_0(\tau, \bar{y}(\tau), \bar{u}(\tau))\, d\tau + V(t+\Delta t, \bar{y}(t+\Delta t)),$$

или после деления на положительное Δt

$$\frac{V(t+\Delta t, \bar{y}(t+\Delta t)) - V(t, \bar{y}(t))}{\Delta t} + \frac{1}{\Delta t} \int_t^{t+\Delta t} f_0(\tau, \bar{y}(\tau), \bar{u}(\tau))\, d\tau \geqslant 0.$$

При $\Delta t \to +0$ первое слагаемое определяет производную $dV(\tau, \bar{y}(\tau))/d\tau|_{\tau=t}$, а второе слагаемое, согласно первой интегральной теореме о среднем [16], дает $f_0(t, y, u)$, т. е.

$$\left.\frac{dV(\tau, \bar{y}(\tau))}{d\tau}\right|_{\tau=t} + f_0(t, y, u) \geqslant 0. \tag{11.6}$$

Учитывая, что функция $\bar{y}(\tau)$ удовлетворяет системе (11.1), запишем выражение для производной более подробно:

$$\left.\frac{dV(\tau, \bar{y}(\tau))}{d\tau}\right|_{\tau=t} = \frac{dV(t, \bar{y}(t))}{dt} = \frac{\partial V(t, y)}{\partial t} +$$

$$+ \sum_{i=1}^n \frac{\partial V(t, \bar{y})}{\partial y_i} \frac{dy_i(t)}{dt} = V_t(t, y) + \langle V_y(t, y), f(t, y, u)\rangle.$$

Подставляя найденное выражение в (11.6), получаем требуемое неравенство (11.4).

Докажем вторую часть теоремы. Пусть u^* — произвольное оптимальное программное управление в (t, y)-задаче, а y^* — соответствующая оптимальная траектория. Возьмем достаточно малое положительное Δt и вычислим

$$I(u^* \mid t, y) = \int_t^T f_0(\tau, y^*(\tau), u^*(\tau))\, d\tau + \Phi(y^*(T)) =$$

$$= \int_t^{t+\Delta t} f_0(\tau, y^*(\tau), u^*(\tau))\, d\tau + \int_{t+\Delta t}^T f_0(\tau, y^*(\tau), u^*(\tau))\, d\tau +$$

$$+ \Phi(y^*(T)). \tag{11.7}$$

Для рассматриваемой исходной задачи справедлив **принцип оптимальности**, утверждающий, что любой «кусок» оптимального управления (оптимальной траектории) также является оптимальным управлением (оптимальной траекторией). Для задач более частного вида, чем исходная, принцип оптимальности сформулирован и доказан в § 10.1. Проверка выполнения принципа оптимальности для исходной задачи по сути не отличается от доказательства, приведенного в § 10.1.

Согласно принципу оптимальности, часть оптимальной траектории y^*, начинающаяся в момент времени $t+\Delta t$ в точке $y^*(t+\Delta t) \in \in \mathbf{R}^n$, также является оптимальной, а значит, последние два слагаемых в (11.7) представляет собой $V(t+\Delta t, y^*(t+\Delta t))$. Поэтому, учитывая равенство $I(u^* \mid t, y) = V(t, y^*(t))$, из (11.7) получаем

$$V(t, y^*(t)) = \int_t^{t+\Delta t} f_0(\tau, y^*(\tau), u^*(\tau))\, d\tau + V(t+\Delta t, y^*(t+\Delta t)).$$

Разделим обе части равенства на Δt:

$$\frac{V(t+\Delta t, y^*(t+\Delta t)) - V(t, y^*(t))}{\Delta t} +$$

$$+ \frac{1}{\Delta t} \int_t^{t+\Delta t} f_0(\tau, y^*(\tau), u^*(\tau))\, d\tau = 0.$$

В пределе при $\Delta t \to +0$ имеем

$$\frac{dV(t, y^*(t))}{dt} + f_0(t, y^*(t), u_+^*) = 0,$$

где $u_+^* = \lim\limits_{\tau \to t+0} u^*(\tau)$. Учитывая, что $y^*(t) = y$, запишем

$$V_t(t, y) + \langle V_y(t, y), f(t, y, u_+^*) \rangle + f_0(t, y, u_+^*) = 0.$$

Это равенство вместе с доказанным неравенством (11.4) означают, что выполнено (11.5). ■

Уравнение в частных производных (11.5) называют *уравнением динамического программирования* или *уравнением Беллмана.*

Эффективных методов решения таких уравнений в настоящее время не существует.

В теореме 11.1 сформулировано необходимое условие оптимальности, и его можно использовать для отыскания решения исходной задачи оптимального управления. Более того, как будет показано в следующем пункте, с помощью уравнения Беллмана в некоторых случаях удается построить оптимальное синтезирующее управление.

2. Достаточные условия оптимального синтеза. Теорема 11.2. *Пусть некоторая гладкая функция $V(t, y)$, определенная на множестве $\Gamma\subset\mathbf{R}^{n+1}$, представляет собой решение уравнения Беллмана* (11.5) *с граничным условием*

$$V(T, y)=\Phi(y) \text{ для } y\in S(T). \tag{11.8}$$

Предположим, что минимум в (11.5) *реализуется на функции $u(t, y)$ вида $u(t, y)\in U$ для всех $(t, y)\in\Gamma$, причем эта функция удовлетворяет требованиям а) и б) из определения оптимальной синтезирующей функции. Тогда $u(t, y)$ — оптимальная синтезирующая функция для исходной задачи, а $V(t, y)$ — функция Беллмана.*

Рассмотрим произвольную точку $(t, y)\in\Gamma$. Согласно требованию а) из определения оптимальной синтезирующей функции, система дифференциальных уравнений

$$\frac{dy}{d\tau}=f(\tau, y, u(\tau, y)),\ t\leqslant\tau\leqslant T,$$

имеет единственное кусочно-гладкое решение $\hat{y}(\tau)$, которое удовлетворяет следующим условиям: $\hat{y}(t)=y$; $\hat{y}(\tau)\in S(\tau)$, $t\leqslant\tau\leqslant T$. В силу требования б) того же определения функция вида $\hat{u}(\tau)=u(\tau, \hat{y}(\tau))$ кусочно-непрерывная на отрезке $[t, T]$, причем $\hat{u}(\tau)\in U$ при $t\leqslant\tau\leqslant T$. Для доказательства теоремы достаточно убедиться в выполнении требования в).

Введем обозначение

$$R(t, y, u)=V_t(t, y)+\langle V_y(t, y), f(t, y, u)\rangle+f_0(t, y, u). \tag{11.9}$$

Выберем произвольное допустимое в (t, y)-задаче управление $u(\tau)$ с соответствующей ему траекторией $y(\tau)$. Так как функция $V(\tau, y(\tau))$ является композицией гладкой и кусочно-гладкой функций, то она представляет собой кусочно-гладкую на $[t, T]$ функцию. Поэтому на основании (11.9) и (11.1) имеем

$$\frac{dV(\tau, y(\tau))}{d\tau}=R(\tau, y(\tau), u(\tau))-f_0(\tau, y(\tau), u(\tau))$$

для всех точек τ отрезка $[t, T]$, за исключением, быть может, некоторого их конечного числа. Отсюда, интегрируя по τ в пределах от t до T и учитывая (11.8), получаем

$$\Phi(y(T))-V(t, y(t))=\int_t^T R(\tau, y(\tau), u(\tau))\,d\tau-$$

$$-\int_t^T f_0(\tau, y(\tau), u(\tau))\,d\tau,$$

что эквивалентно равенству

$$I(u \mid t, y)=\int_t^T R(\tau, y(\tau), u(\tau))\,d\tau+V(t, y). \tag{11.10}$$

(Это представление целевого функционала (t, y)-задачи будет использовано ниже.)

В терминах обозначения (11.9) уравнение Беллмана (11.5) можно записать в виде

$$\min_{u\in V} R(\tau, y, u)=0.$$

Так как функция $u(t, y)$, по условию, реализует минимум, то для всех векторов $u\in U$ и $y\in S(\tau)$, $t\leqslant\tau\leqslant T$, имеем

$$R(\tau, y, u(\tau, y))=0=\min_{u\in U} R(\tau, y, u)\leqslant R(\tau, y, u),\ t\leqslant\tau\leqslant T. \tag{11.11}$$

Для произвольного допустимого в (t, y)-задаче управления $u(\tau)$ и соответствующей ему траектории $y(\tau)$, удовлетворяющей фазовому ограничению, выполняется включение $u(\tau)\in U$ при $t\leqslant\tau\leqslant T$. Поэтому из (11.11) следует, что

$$R(\tau, \hat y(\tau), u(\tau, \hat y(\tau)))=R(\tau, \hat y(\tau), \hat u(\tau))=0\leqslant$$
$$\leqslant R(\tau, y(\tau), u(\tau)),\ t\leqslant\tau\leqslant T, \tag{11.12}$$

где $\hat u$ и $\hat y$ — функции, введенные в начале доказательства. Вспоминая, что $\hat y(t)=y(t)=y$, на основании (11.10) и (11.12) можно записать неравенство

$$I(u \mid t, y)-I(\hat u \mid t, y)=\int_t^T R(\tau, y(\tau), u(\tau))\,d\tau-$$
$$-\int_t^T R(\tau, \hat y(\tau), \hat u(\tau))\,d\tau=\int_t^T R(\tau, y(\tau), u(\tau))\,d\tau\geqslant 0, \tag{11.13}$$

которое в силу произвольности управления u означает оптимальность управления $\hat u$ в (t, y)-задаче. Таким образом, $u(t, y)$ — оптимальная синтезирующая функция.

Проверим, что функция $V(t, y)$ действительно является функцией Беллмана для исходной задачи. В самом деле, из (11.13) следует равенство

$$I(u \mid t, y)-I(\hat u \mid t, y)=\int^T R(\tau, y(\tau), u(\tau))\,d\tau,$$

которое вместе с (11.10) дает

$$V(t, y)=I(\hat u \mid t, y)=\min_u I(u \mid t, u).$$

Последнее и означает, что $V(t, y)$ — функция Беллмана. ■

В соответствии с теоремой 11.2 решение задачи синтеза оптимального управления при определенных условиях сводится к решению уравнения Беллмана с граничным условием (11.8). Одним из существенных недостатков, связанных с применением теоремы 11.2, заключается в априорном требовании дифференцируемости функции Беллмана. Еще не решая уравнения Беллмана, нужно быть уверенным в дифференцируемости его решения. Это требование слишком «жесткое», поскольку даже в несложных линейных задачах оптимального управления функция Беллмана, как правило, не является всюду дифференцируемой [5]. Далее, при решении уравнения Беллмана следует отыскивать не только функцию $V(t, y)$, но и функцию $u(t, y)$, реализующую минимум в (11.5), причем функция $u(t, y)$ должна удовлетворять требованиям а) и б) из определения оптимальной синтезирующей функции. Минимум в (11.5) может достигаться не на одном, а сразу на нескольких векторах множества U, поэтому построение требуемой функции $u(t, y)$ нередко оказывается крайне сложной задачей.

Несмотря на указанные трудности, существуют задачи, для которых удается синтезировать оптимальное управление с помощью метода динамического программирования. Одним из наиболее интересных с этой точки зрения является класс линейных задач оптимального управления с квадратичным целевым функционалом.

3. Оптимальный синтез в линейных системах с квадратичным критерием качества. Пусть управляемая система линейная и в векторной форме имеет вид

$$\frac{\mathrm{d}\, y}{\mathrm{d}\, t}=A(t)y+B(t)u,\ t_0\leqslant t\leqslant T,$$

где $\mathrm{d}y/\mathrm{d}t=(\mathrm{d}y_1/\mathrm{d}t,\ \mathrm{d}y_2/\mathrm{d}t,\ \ldots,\ \mathrm{d}y_n/\mathrm{d}t)^T$, $A(t)$ — матрица размера $n\times n$; $B(t)$ — матрица размера $n\times m$, причем компоненты этих матриц являются непрерывными функциями. Начальный t_0 и конечный T моменты времени считаем заданными. Задано также начальное состояние $y^{(0)}$, а фазовые ограничения при $t>t_0$ отсутствуют, т. е. $S(t)=\mathbf{R}^n$, $t_0<t\leqslant T$. Согласно терминологии, принятой в предыдущей главе, это задача оптимального управления линейной системой со свободным правым концом и фиксированным временем управления. Целевой функционал имеет следующий квадратичный вид:

$$I(u)=\int_{t_0}^{T}[\langle y(t),\ P(t)y(t)\rangle+\langle u(t),\ Q(t)u(t)\rangle]\mathrm{d}t+$$
$$+\langle y(T),\ Ry(T)\rangle.$$

Здесь матрицы $P(t)$, $Q(t)$ и R размера $n\times n$ считаются симметричными, причем компоненты матриц $P(t)$ и $Q(t)$ составлены из непрерывных функций. Кроме того, предполагаем, что матрицы $P(t)$ и R неотрицательно определены, а $Q(t)$ — положительно-определенная матрица при всех $t\in[t_0, T]$. Допустимыми управлениями являются кусочно-непрерывные на $[t_0, T]$ функции $u(t)$, принимающие произвольные значения из $\mathbf{R}^m$ (т. е. $U=\mathbf{R}^m$).

Сформулированная задача представляет собой частный случай исходной задачи оптимального управления, и поэтому для построения оптимальной синтезирующей функции можно использовать теорему 11.2. Предположим, что функция Беллмана в рассматриваемой задаче имеет вид

$$V(t, y)=\langle y, C(t)y\rangle, \tag{11.14}$$

где $C(t)=(c_{ij}(t))$ — симметричная матрица размера $n\times n$ с гладкими элементами, причем $C(T)=R$.

Вычислим частные производные функции вида (11.14). Имеем

$$V_t(t, y)=\left\langle y, \frac{\mathrm{d}C(t)}{\mathrm{d}t}y\right\rangle,$$

где $\mathrm{d}C(t)/\mathrm{d}t$ — матрица размера $n\times n$, составленная из соответствующих производных $\mathrm{d}c_{ij}(t)/\mathrm{d}t$. Далее находим

$$V_y(t, y)=C(t)y+C^T(t)y.$$

Функция (11.14) удовлетворяет уравнению Беллмана, поэтому в данном случае равенство

$$\min_{u\in\mathbf{R}^m}\left\{\left\langle y, \frac{dC(t)}{dt}y\right\rangle+\langle C(t)y+C^T(t)y, A(t)y+B(t)u\rangle+\right.$$
$$\left.+\langle y, P(t)y\rangle+\langle u, Q(t)u\rangle\right\}=0 \tag{11.15}$$

выполняется для любого $t_0<t\leqslant T$ и всех $y\in\mathbf{R}^n$. При каждом t выражение в фигурных скобках является квадратичной функцией относительно u с положительно-определенной матрицей $Q(t)$. Следовательно, на $\mathbf{R}^m$ минимум этой квадратичной функции существует и достигается в единственной точке, которую можно найти, приравнивая градиент квадратичной функции нулю, т. е. из условия

$$B^T(t)[C(t)y+C^T(t)y]+2Q(t)u=0_n.$$

Отсюда, учитывая симметричность матрицы $C(t)$, находим точку минимума:

$$u=-Q^{-1}(t)B^T(t)C(t)y. \tag{11.16}$$

Теперь найдем матрицу $C(t)$. Подставим выражение (11.16) в уравнение (11.15). В результате получим уравнение для определения матрицы $C(t)$:

$$\left\langle y, \frac{dC(t)}{dt}y\right\rangle+\langle C(t)y+C^T(t)y, A(t)y-$$
$$-B(t)Q^{-1}(t)B(t)C(t)y\rangle+\langle y, P(t)y\rangle+$$
$$+\langle Q^{-1}(t)B(t)C(t)y, B(t)C(t)y\rangle=0.$$

Преобразуем это уравнение, опуская для краткости аргумент t. Используя равенства.

$$\langle Cy,\ Ay\rangle = \langle y,\ C^T Ay\rangle = \langle y,\ CAy\rangle,$$
$$\langle C^T y,\ Ay\rangle = \langle A^T C^T y,\ y\rangle = \langle y,\ A^T Cy\rangle,$$
$$\langle Cy,\ BQ^{-1}B^T Cy\rangle = \langle y,\ CBQ^{-1}B^T y\rangle,$$
$$\langle Q^{-1}B^T Cy,\ B^T Cy\rangle = \langle y,\ C^T B(Q^{-1})^T B^T Cy\rangle = \langle y,\ CBQ^{-1}B^T Cy\rangle,$$

имеем

$$\left\langle y,\ \frac{dC}{dt}y\right\rangle + \langle y,\ CAy\rangle + \langle y,\ A^T Cy\rangle - \langle y,\ CBQ^{-1}B^T Cy\rangle +$$
$$+\langle y,\ Py\rangle = 0.$$

Это равенство заведомо выполняется для всех векторов $y \in \mathbf{R}^n$, если выполнено условие

$$\frac{dC}{dt} + CA + A^T C - CBQ^{-1}B^T C + P = \mathbf{0}, \tag{11.17}$$

где в правой части равенства находится нулевая матрица размера $n \times n$. Полученное дифференциальное уравнение относительно C называют *матричным уравнением Риккати*. Это уравнение представляет собой систему n^2 нелинейных дифференциальных уравнений первого порядка типа Риккати. Оказывается (см. кн.: *Флеминг У., Ришел Р.* Оптимальное управление детерминированными и стохастическими системами. — М.: Мир, 1978), при сделанных в начале пункта предположениях, матричное уравнение Риккати (11.17) всегда имеет решение $C(t)$, определенное для $-\infty < t \leqslant T$ и удовлетворяющее условию $C(T) = R$. Следовательно, поиск решения уравнения Беллмана в виде квадратичной функции (11.14) имеет полное основание и можно положить

$$u(t,\ y) = -Q^{-1}(t)B^T(t)C(t)y, \tag{11.18}$$

где $C(t)$ — решение уравнения (11.17) с условием $C(T) = R$. Функция (11.18) удовлетворяет требованиям а) и б) определения оптимальной синтезирующей функции, причем, как следует из ее построения, она реализует минимум в уравнении Беллмана. По теореме 11.2, это искомая оптимальная синтезирующая функция рассматриваемой задачи.

Если матрицы A, B, P и Q не зависят от t, т. е. являются стационарными, то решением задачи синтеза также является стационарная оптимальная синтезирующая функция, линейная по фазовым переменным.

§ 11.3. ПРИКЛАДНЫЕ ЗАДАЧИ

1. Управление движением материальной точки. Пусть управляемая система описывается уравнением

$$\frac{d^2y}{dt^2}=u, \tag{11.19}$$

где $u(t)$ — управляющая функция, на которую наложено ограничение

$$|u(t)|\leqslant 1. \tag{11.20}$$

Запишем фазовые уравнения системы (11.19) в следующем виде:

$$\frac{dy_1}{dt}=y_2,\ \frac{dy_2}{dt}=u. \tag{11.21}$$

Пусть требуется за минимальное время перевести систему из некоторого начального состояния $y^{(0)}$ в начало координат фазового пространства (т. е. конечным положением системы является точка $y=(0, 0)$).

Найдем программное управление, соответствующее фиксированному начальному условию. Воспользуемся принципом максимума. Функция $\hat{H}$ в данном случае (см. § 10.4) имеет вид

$$\hat{H}=\langle\psi(t),\ Bu(t)\rangle=\psi_1y_2+\psi_2u, \tag{11.22}$$

где $\psi_1(t)$ и $\psi_2(t)$ — вспомогательные функции, определяемые из системы уравнений

$$\frac{d\psi_1}{dt}=-\frac{\partial\hat{H}}{\partial y_1},\ \frac{d\psi_2}{dt}=-\frac{\partial\hat{H}}{\partial y_2},$$

т. е.

$$\frac{d\psi_1}{dt}=0,\ \frac{d\psi_2}{dt}=-\psi_1.$$

Общее решение этой системы таково:

$$\psi_1=C_1,\ \psi_2=C_2-C_1t,$$

где C_1, C_2 — постоянные. Функция $\hat{H}$ принимает вид

$$\hat{H}=C_1y_2+(c_2-C_1t)u.$$

Максимум по переменной u достигается при

$$u(t)=\operatorname{sign}(C_2-C_1t). \tag{11.23}$$

Функция C_2-C_1t на отрезке управления $[0, T]$ может менять знак не более одного раза, поэтому $u(t)$ — кусочно-постоянная функция, принимающая значения $+1$ и -1 и имеющая не более двух интервалов постоянства. При фиксированном начальном фазовом положении $y^{(0)}=(y_{01}, y_{02})$ точку, в которой функция $u(t)$ изменяет знак (точку переключения), можно определить, подставляя (11.23) в систему (11.21) и решая эту систему совместно с начальными и конечными условиями $y_1(0)=y_{01}$, $y_2(0)=y_{02}$, $y_1(T)=0$, $y_2(T)=0$. В результате получаем программное управление $u(t)$. Определим вид фазовых траекторий системы при возможных значениях управляющей функции.

Пусть на некотором отрезке времени $u(t)\equiv1$; тогда, решая систему (11.21), получаем, что фазовая траектория системы имеет вид

$$y_2(t)=t+D_1,\ y_1(t)=\frac{t^2}{2}+D_1t+D_2=\frac{1}{2}(t+D_1)^2+$$
$$+\left(D_2-\frac{1}{2}D_1^2\right).$$

Исключая время, запишем фазовую траекторию в переменных y_1, y_2:

$$y_1=\frac{1}{2}y_2^2+D \tag{11.24}$$

(здесь D_1, D_2, $D=D_2-\frac{1}{2}D_1^2$ — постоянные). Из (11.24) следует, что в фазовом пространстве R^2 фазовая траектория системы на промежутке времени, где $u(t)=1$, представляет собой дугу параболы. Семейство парабол, определяемых уравнением (11.24) при различных значениях константы D изображено на рис. 11.2.

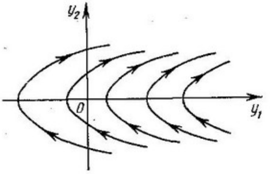

Рис. 11.2

На тех интервалах времени, где $u(t)=-1$, решение системы (11.21) имеет вид

$$y_2(t)=-t+E_1,\ y_1(t)=-\frac{t^2}{2}+E_1t+E_2=-\frac{1}{2}(-t+E_1)^2+$$
$$+\left(E_2+\frac{1}{2}E_1^2\right).$$

В переменных y_1, y_2 фазовая траектория записывается следующим образом:

$$y_1=-\frac{1}{2}y_2^2+E, \tag{11.25}$$

где $E=E_2+\frac{1}{2}E_1^2$. Семейство парабол, по которым система движется при $u(t)\equiv-1$, изображено на рис. 11.3. Отметим, что по параболам (11.24) система движется снизу вверх, так как при $u(t)\equiv$

$\equiv 1$ выполнено условие $\frac{dy_2}{dt} = u = 1 > 0$, т. е. с ростом t координата y_2 увеличивается. По параболам (11.25) система движется сверху вниз, так как $\frac{dy_2}{dt} = -1 < 0$.

Пусть задано начальное фазовое положение $y^{(0)}$. Если точка $y^{(0)} = (y_{01}, y_{02})$ лежит на параболе вида (11.24) или вида (11.25), проходящей через начало координат, то управление в этом случае имеет вид $u(t) \equiv 1$ или $u(t) \equiv -1$ на всем интервале управления

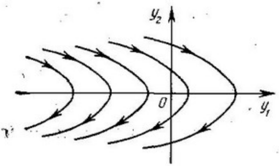

Рис. 11.3

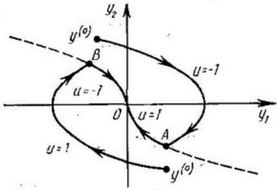

Рис. 11.4

[0, T], а фазовая траектория представлеят собой часть параболы от точки $y^{(0)}$ до начала координат.

Если точка $y^{(0)}$ находится выше оси Oy_1 и правее параболы, проходящей через начало координат, то фазовая траектория состоит из двух частей. Сначала система движется вниз по параболе, проходящей через точку $y^{(0)}$ и соответствующей $u(t) \equiv -1$. Достигнув точки пересечения данной параболы с параболой, соответствующей $u(t) \equiv 1$ и проходящей через начало координат, система далее движется по этой параболе (рис. 11.4). В точке A значение функции $u(t)$ меняется с -1 на $+1$.

Если $y^{(0)}$ находится ниже оси Oy_1 и левее параболы, проходящей через начало координат, то система сначала движется вверх по параболе, проходящей через точку $y^{(0)}$, а затем вниз по параболе, которая проходит через начало координат (рис. 11.4). Точка B, в которой происходит переход от одной параболы к другой, является точкой переключения управления со значения $u \equiv 1$ на $u \equiv -1$.

Случаи, когда точка $y^{(0)}$ лежит выше оси Oy_1 и левее параболы, проходящей через начало координат, и когда $y^{(0)}$ лежит ниже оси Oy_1 и правее параболы, проходящей через начало координат, исследуют аналогично.

Семейство фазовых траекторий, по которым система переходит из произвольной начальной фазовой точки в начало координат, изображено на рис. 11.5. На этом рисунке AOB — линия переключения управления. Если система находится в фазовой точке выше этой линии, то управление имеет значение $u \equiv -1$, если ниже, то значение $u\equiv 1$. Часть AO линии переключения представляет собой дугу параболы $y_1=y_2^2$, а часть OB — дугу параболы $y_1=-\frac{1}{2}y_2^2$.

При произвольном начальном состоянии $y^{(0)}$ система сначала движется по параболе, проходящей через начальную точку, а затем (после достижения линии переключения) по линии переключения в начало координат. В случае если $y^{(0)}$ не лежит на линии переключения, управляющая функция $u(t)$ обязательно имеет одну точку переключения с -1 на $+1$ или с $+1$ на -1. На рис. 11.5 изображено графическое решение задачи синтеза оптимальных по быстродействию управлений. Для практической реализации найденного оптимального управления необходимо ввести в систему реле, изменяющее значение управляющего параметра в зависимости от фазового положения системы. Область, расположенную на рис. 11.5 выше линии переключения, можно задать с помощью условия $y_1+\frac{1}{2}y_2^2\operatorname{sign}y_2>0$, область ниже линии переключения — с помощью условия $y_1+\frac{1}{2}y_2^2\operatorname{sign}y_2<0$. Таким образом, управляющее устройство должно измерять величину $y_1+\frac{1}{2}y_2^2\operatorname{sign}y_2$ придавать управляющему параметру значения $u\equiv -1$, (если эта величина положительна) и $u\equiv 1$ (если эта величина отрицательна). Если выполнено условие $y_1+\frac{1}{2}y_2^2\operatorname{sign}y_2=0$, то управляющее устройство должно придавать управляющему параметру значение $u\equiv -1$ при $y_2>0$ и значение $u\equiv 1$ при $y_2<0$.

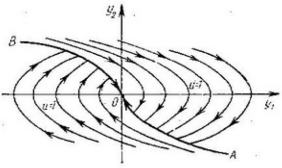

Рис. 11.5

2. Синтез оптимального по быстродействию управления электродвигателем при заданном значении выполняемой работы. Уравнение электродвигателя постоянного тока в безразмерных единицах имеет вид (см. пример 3 из § 9.7)

$$\frac{d\nu(\tau)}{d\tau}=i(\tau)-\mu_0, \tag{11.26}$$

где $\nu(\tau)$ — угловая скорость вращения якоря, $i(\tau)$ — сила тока в якоре, $\mu_0<1$ — момент сопротивления нагрузки, τ — время. Работу A, выполняемую двигателем на интервале времени $[0, T]$, можно вычислить (в относительных единицах) по формуле

$$A=\int_0^T \mu\nu(\tau)\,d\tau. \tag{11.27}$$

Ток $i(\tau)$ в якоре выбирают в качестве управляющего параметра. Требуется подобрать такую функцию $i(\tau)$, $0 \leqslant \tau \leqslant T$, чтобы за минимальное время T электродвигатель выполнил заданную работу A_0, т. е. функционал (11.27) при оптимальном управлении $i(\tau)$ должен иметь заданное значение A_0. Потребуем, чтобы угловая скорость вращения якоря в начальный и конечный моменты времени была равна нулю: $\nu(0)=\nu(T)=0$. На управляющее воздействие $i(\tau)$ наложим естественное ограничение

$$|i(\tau)| \leqslant I. \tag{11.28}$$

Так как $i(\tau)=\dfrac{I(\tau)}{I_н(\tau)}$, где I и $I_н$ — соответственно текущее и номинальное значения тока в якоре, условие (11.28) означает, что ток в якоре не должен превышать номинального значения.

Наконец, предположим, что нагрузка может подключаться к электродвигателю в любой из моментов времени $t \in [0, T]$. В этом случае значение момента сопротивления нагрузки в уравнении (11.26) необходимо считать функцией времени: $\mu(\tau)=\mu_0\chi(\tau_1)$, где τ_1 — момент подключения нагрузки. Фактически $\mu(\tau)$ (как и $i(\tau)$) — управляющее воздействие, поскольку от выбора момента подключения нагрузки τ_1 зависит время выполнения заданной работы. При этом кажущийся наиболее естественным режим, в котором нагрузка подключена к электродвигателю с самого начала ($\tau_1=0$), на самом деле не является оптимальным. Как будет показано, минимум времени, затрачиваемого на выполнения заданной работы A_0, соответствует режиму, в котором $\tau_1>0$. В этом режиме сначала разгоняют якорь электродвигателя на холостом ходу, а уже затем подключают нагрузку.

Чтобы учесть условие выполнения электродвигателем заданной работы, введем новую переменную $y(\tau)$ с помощью уравнения $\dfrac{dy(\tau)}{d\tau}=\mu(\tau)\nu(\tau)$ и граничного условия $y(T)=A_0$. Величина $y(\tau)$ определяет работу, совершенную электродвигателем на интервале времени $[0, \tau]$. Таким образом, система фазовых уравнений, описывающих электродвигатель, имеет вид

$$\frac{d\nu}{d\tau}=i-\mu, \tag{11.29}$$

$$\frac{dy}{d\tau}=\mu\nu. \tag{11.30}$$

Здесь $\nu(\tau)$ и $y(\tau)$ — фазовые координаты процесса, $i(\tau)$, $\mu(\tau)$ — управляющие воздействия. На первое из них наложено ограничение (11.8), а второе должно иметь вид $\mu(\tau)=\mu_0\chi(\tau_1)$. Задача оптимального управления электродвигателя состоит в выборе таких функций $i(\tau)$ и $\mu(\tau)$, чтобы система за минимальное время из начального состояния с фазовыми координатами

$$\nu(0)=0, \quad y(0)=0 \tag{11.31}$$

перешла в состояние с фазовыми координатами

$$\nu(T)=0, \quad y(T)=A_0. \tag{11.32}$$

Запишем функцию $\widehat{H}$ (см. § 10.4) для сформулированной задачи оптимального быстродействия:

$$\widehat{H}=\psi_1(i-\mu)+\psi_2\mu\nu=\psi_1 i+(\psi_2\nu-\psi_1)\mu. \tag{11.33}$$

Функции $\psi_1(\tau)$ и $\psi_2(\tau)$ удовлетворяют уравнениям

$$\frac{d\psi_1}{d\tau}=-\frac{\partial\widehat{H}}{\partial\nu}=-\psi_2\mu, \quad \frac{d\psi_2}{d\tau}=-\frac{\partial\widehat{H}}{\partial y}=0.$$

Решение этих уравнений имеет вид $\psi_2(\tau)\equiv\psi_2(0)=\mathrm{const}$, $\psi_1(\tau)=$ $=\psi_1(0)-\psi_2(0)\mu_0\chi(\tau_1)\tau$. Поскольку при $\tau=0$ выполнено $\nu(0)=0$, $\mu(0)=0$, можно записать $\widehat{H}=\psi_1(0)i(0)$. В соответствии с принципом максимума на оптимальной траектории $\widehat{H}=\mathrm{const}\geqslant 0$. Учитывая, что $i(0)>0$ (так как для совершения работы необходимо привести якорь двигателя во вращение), получаем, что $\psi_1(0)\geqslant 0$. Из выражения для $\psi_1(\tau)$ следует, что функция ψ_1 либо не имеет ни одной точки перемены знака (если $\psi_2(0)<0$), либо имеет одну такую точку (если $\psi_2(0)>0$). Соответственно функция $i(\tau)$ либо не имеет ни одного переключения, т. е. $i(\tau)\equiv 1$, либо имеет одну точку переключения на интервале управления. В первом из этих случаев мы не получили решения поставленной задачи, поскольку решение $\nu(\tau)$ фазового уравнения (11.29) при $i\equiv 1$ является линейно возрастающей функцией (так как $1-\mu_0>0$) и условие $\nu(T)=0$ не может быть выполнено ни при каком T. Таким образом, в оптимальном режиме функция $i(\tau)$ имеет одну точку переключения $\tau=\tau_2$, в которой она изменяется с $+1$ на -1. Из физических соображений следует, что нагрузка должна подключаться к электродвигателю во время его разгона, а не во время торможения, т. е. должно быть выполнено условие $0<\tau_1<\tau_2<T$. Графики функций $i(\tau)$, $\mu(\tau)$, $\nu(\tau)$ и $y(\tau)$ в оптимальном режиме изображены на рис. 11.6. Чтобы найти программное управление $i(\tau)$, $\mu(\tau)$, осталось определить значения величин τ_1 и τ_2 (точек переключений управлений). Это можно сделать, подставляя общее решение системы уравнений (11.29), (11.30) (содержащее четыре неизвестных величины: τ_1, τ_2 и две произвольные постоянные) в граничные условия (11.31), (11.32).

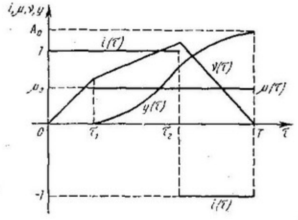

Рис. 11.6

Рассмотрим более сложную задачу синтеза оптимальных по быстродействию управлений. Исследуем вид фазовых траекторий системы при различных возможных значениях управляющих воздействий $i(\tau)$, $\mu(\tau)$. Как было установлено, на интервале управления необходимо рассмотреть три участка. На участке $[0, \tau_1]$ выполнено $i(\tau)\equiv 1$, $\mu(\tau)\equiv 0$, на участке $[\tau_1, \tau_2]$ имеем $i(\tau)\equiv 1$, $\mu(\tau)\equiv\mu_0$, на участке $[\tau_2, T]$ справедливо $i(\tau)\equiv -1$, $\mu(\tau)\equiv\mu_0$. Установим вид фазовых траекторий на плоскости $yO\nu$. На последнем участке фазовые уравнения (11.29), (11.30) имеют вид

$$\frac{d\nu}{d\tau}=-(1+\mu_0),\quad \frac{dy}{d\tau}=\mu_0\nu.$$

Разделив второе из этих уравнений на первое, получим

$$\frac{dy}{d\nu}=-\frac{\mu_0}{1+\mu_0}\nu.$$

Далее находим вид фазовых траекторий в координатах (y, ν):

$$y=C_1-\frac{\mu_0}{2(1+\mu_0)}\nu^2. \tag{11.34}$$

По параболам этого семейства фазовая точка системы движется при $i\equiv -1$, $\mu\equiv\mu_0$. В конечную точку $(A_0, 0)$ на плоскости $yO\nu$ фа-

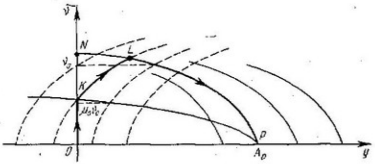

Рис. 11.7

зовая точка системы попадает по траектории вида (11.34), соответствующей $C_1=A_0$ (жирная линия на рис. 11.7).

При $i\equiv 1$, $\mu\equiv\mu_0$ фазовые уравнения соответственно принимают вид

$$\frac{d\nu}{d\tau}=1-\mu_0,\quad \frac{dy}{d\tau}=\mu_0\nu.$$

Разделив второе уравнение на первое, получим

$$\frac{dy}{d\nu}=\frac{\mu_0}{1-\mu_0}\nu.$$

Фазовые траектории в этом случае таковы:

$$y=C_2+\frac{\mu_0}{2(1-\mu_0)}\nu^2. \tag{11.35}$$

Напомним, что $\mu_0<1$ (определение μ_0 приведено в § 9.7), поэтому на параболах вида (11.35) координата y возрастает с ростом ν. На рис. 11.7 траектории из семейства (11.35) изображены пунктирными линиями.

При $i \equiv 1$, $\mu \equiv 0$ (начальный участок управления) фазовые уравнения имеют вид

$$\frac{d\nu}{d\tau} = 1, \quad \frac{dy}{d\tau} = 0.$$

Отсюда, учитывая начальные условия (11.31), получаем $\nu(\tau) = \tau$, $y(\tau) = 0$. В плоскости $y0\nu$ траектория системы на участке $0 \leqslant \tau \leqslant \tau_1$ представляет собой часть оси 0ν.

Таким образом, оптимальная траектория системы имеет следующий вид. Сначала (до момента τ_1) фазовая точка системы движется по отрезку OK оси $O\nu$ (рис. 11.7) под действием управления $\tau \equiv 1$, $\mu \equiv 0$. В момент τ_1 происходит переключение значения функции $\mu(\tau)$ и фазовая точка системы движется под действием управления $i \equiv 1$, $\mu = \mu_0$ по отрезку KL параболы семейства (11.35). В момент τ_2 происходит переключение значения функции $i(\tau)$ и фазовая точка системы движется под действием управления $i \equiv -1$, $\mu = \mu_0$ по параболе

$$y = A_0 - \frac{\mu_0}{2(1+\mu_0)} \nu^2 \tag{11.36}$$

и достигает точки $(A_0, 0)$ за минимальное время T. Отрезок NP параболы (11.36) представляет собой линию переключения управляющего воздействия $i(\tau)$. Таким образом, для завершения решения задачи синтеза осталось найти аналитическое выражение для линии переключения управляющего воздействия $\mu(\tau)$.

В соответствии с теоремой 10.1 на оптимальной траектории выполнено следующее условие: $H = \text{const}$. При $\tau = 0$ из (11.33) получаем

$$H = \psi_1(0), \tag{11.37}$$

так как $i(0) = 1$, $\mu(0) = 0$. В точке $\tau = \tau_2$ происходит переключение функции $i(\tau)$, т. е. $\psi_1(\tau_2) = 0$, поэтому имеем (учитывая, что $\psi_2(\tau) \equiv \psi_2(0) = \text{const}$) $H(\tau_2) = \psi_2(0)\nu(\tau_2)\mu(\tau_2)$. Учитывая (11.37), получаем

$$\psi_2(0)\nu(\tau_2)\mu(\tau_2) = \psi_1(0). \tag{11.38}$$

Поскольку в точке $\tau = \tau_1$ происходит переключение функции $\mu(\tau)$, выполнено равенство $\psi_2(0)\nu(\tau_1) - \psi_1(\tau_1) = 0$. Так как $\mu(\tau) \equiv \equiv 0$ на участке $[0, \tau_1]$, то $\psi_1(\tau) \equiv \psi_1(0)$ на этом участке и из предыдущего соотношения следует, что $\psi_1(0) = \psi_2(0)\nu(\tau_1)$. Подставляя выражение для $\psi_1(0)$ в (11.38), найдем

$$\mu(\tau_2)\nu(\tau_2) = \nu(\tau_1). \tag{11.39}$$

Отсюда следует, что линия переключения управляющего воздействия $\mu(\tau)$ состоит из тех точек (y, ν) на параболах (11.35), для которых выполнено условие $\nu = \mu(\tau_2)\nu(\tau_2) = \mu_0\nu(\tau_2)$ (рис. 11.7). Очевидно, что эта линия представляет собой параболу, переходящую через точку (A_0, 0), т. е. ее уравнение имеет вид $y = A_0 - D\nu^2$. Для нахождения коэффициента D рассмотрим некоторую параболу семейства (11.35). Предположим она пересекается с параболой (11.36) в точке с координатой ν_0. Тогда с искомой параболой $y = A_0 - D\nu^2$ она пересекается в точке с координатой $\mu_0\nu_0$ (рис. 11.7). Запишем условия пересечения в этих двух точках:

$$C_2 + \frac{\mu_0}{2(1-\mu_0)}\nu_0^2 = A_0 - \frac{\mu_0}{2(1+\mu_0)}\nu^2,$$

$$C_2 + \frac{\mu_0}{2(1-\mu_0)}(\mu_0\nu_0)^2 = A_0 - D(\mu_0\nu_0)^2.$$

Почленно вычитая второе уравнение из первого, получаем соотношение

$$\frac{\mu_0}{2(1-\mu_0)}[\nu_0^2 - (\mu_0\nu_0)^2] = D(\mu_0\nu_0)^2 - \frac{\mu_0\nu_0^2}{2(1+\mu_0)},$$

из которого находим

$$D = \frac{2+\mu_0}{2\mu_0(1+\mu_0)}.$$

Линия переключения управляющего воздействия $\mu(\tau)$ в плоскости $yO\nu$ задается уравнением

$$y = A_0 - \frac{2+\mu_0}{2\mu_0(1+\mu_0)}\nu^2. \qquad (11.40)$$

Итак, *оптимальная* фазовая траектория рассматриваемой задачи с заданными начальными и конечными условиями (11.31), (11.32) построена. Фактически осуществлен синтез оптимальных по быстродействию управлений для задачи с произвольным начальным значением скорости вращения якоря и произвольным значением выполняемой работы. Пусть, например, в начальный момент времени скорость вращения якоря отличная от нуля и равна ν_1 и требуется построить управление, обеспечивающее выполнение за минимальное время работы A, причем $A > A_0$. Чтобы свести эту задачу к рассмотренной ранее, достаточно формально сместить точку отсчета выполняемой электродвигателем работы и положить, что в момент времени $\tau = 0$ «выполненная» работа равна $A_0 - A$ (т. е. отрицательна). В этом случае начальному состоянию управляемой системы соответствует точка ($A_0 - A$, ν_1) в фазовой плоскости $yO\nu$. Оптимальная фазовая траектория состоит по-прежнему из трех участков. Сначала фазовая точка системы движется по отрезку прямой, параллельной оси $O\nu$, под действием управления $i \equiv 1$, $\mu \equiv 0$. В тот момент, когда точка достигает линии переключения (11.40), происходит изменение управления: $i \equiv 1$, $\mu \equiv \mu_0$. Дальнейшее движение происходит по отрезку параболы из семейства (11.35). При достижении линии (11.36) происходит переключение управляющего воздействия $i(\tau)$ и управление принимает следующий вид: $i\equiv-1$, $\mu\equiv\mu_0$. По отрезку параболы (11.36) фазовая точка системы переходит в конечную точку $(A_0, 0)$, которая в данном случае соответствует моменту выполнения работы A.

Таким образом, окончательно решение задачи синтеза может быть записано в виде (рис. 11.8)

$$(i(\tau),\ \mu(\tau))=\begin{cases}(1,\ 0) \text{ ниже линии } PM,\\ (1,\ \mu_0) \text{ между линиями } PM \text{ и } PN \text{ и на линии } PM,\\ (-1,\ \mu_0) \text{ на линии } PN.\end{cases}$$

Отметим, что постановка задачи оптимального быстродействия с произвольными значениями $\nu(0)$ и A имеет смысл только если начальная фазовая точка системы лежит ниже линии PN. Если начальная точка находится выше этой линии, задача теряет смысл, так как невозможно подобрать допустимое управление, которое переводило бы систему в точку $(A_0,\ 0)$. В этом случае электродвигатель имеет в начальный момент времени слишком большую скорость и совершаемая им до полной остановки работа будет обязательно больше заданного значения A.

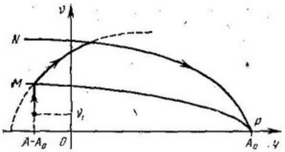

Рис. 11.8

3. Синтез оптимальных по быстродействию управлений ядерным реактором. В § 10.5 были решены задачи оптимального распределения горючего в ядерном реакторе в случае стационарного режима работы реактора. Рассмотрим теперь более сложную задачу управления нестационарными режимами работы ядерного реактора.

Составим математическую модель нестационарного процесса диффузии нейтронов в реакторе. Будем считать, что активная зона реактора занимает пространственную область Ω, границу которой обозначим через S. Стационарный процесс диффузии в реакторе можно описать уравнением

$$D_0\nabla N_0(M)+\frac{k_{\text{кр}}-1}{T_0}N_0(M)=0,\quad M\in\Omega \tag{11.41}$$

с граничным условием

$$N_0(M)|_{M\in S}=0. \tag{11.42}$$

Здесь D_0 — эффективный коэффициент диффузии, T_0 — время жизни теплового нейтрона, $k_{\text{кр}}$ — критический коэффициент размножения нейтронов, зависящий от D_0, T_0 и геометрических размеров реактора.

Если реальный коэффициент размножения нейтронов в реакторе k_0 отличен от $k_{\text{кр}}$, плотность нейтронов в каждой точке внутри реактора изменяется. Математическая модель этого процесса может учитывать или не учитывать запаздывающие нейтроны, которые излучаются из осколков ядер не в момент распада, а с некоторым запаздыванием по времени. В первом случае уравнение нестационарного процесса диффузии нейтронов можно записать в виде

$$\frac{\partial N}{\partial t}=D_0\Delta N+\frac{k_0-1}{T_0}N, \qquad (11.43)$$

где k_0 — коэффициент размножения, не учитывающий запаздывающие нейтроны; $N(t, M)$ — плотность нейтронов в точке M в момент времени t. Оба члена в правой части уравнения имеют простой физический смысл. Первое слагаемое определяет изменение плотности нейтронов, вызванное их движением через поверхность, которая ограничивает произвольный бесконечно малый объем внутри реактора; второе слагаемое определяет изменение плотности нейтронов вследствие рождения их при распаде ядер внутри указанного объема. В том случае, когда учитываются и запаздывающие нейтроны, необходимо в правую часть уравнения (11.43) добавить слагаемое, описывающее изменение плотности запаздывающих нейтронов внутри рассматриваемого бесконечно малого объема. При этом все запаздывающие нейтроны разбивают на m группы (как правило $m\leqslant 6$) и вводят коэффициенты размножения k_i для каждой из этих групп. Общий коэффициент размножения является суммой коэффициентов размножения мгновенных и запаздывающих нейтронов:

$$k=k_0+\sum_{i=1}^{n}k_i.$$

Учитывая запаздывающие нейтроны каждой группы, получим [11] уравнение

$$\frac{\partial N}{\partial t}=D_0\Delta N+\frac{k_0-1}{T_0}N+\sum_{i=1}^{m}\frac{k_i\lambda_i}{T_0}\int_0^\infty N(M,\,t-\tau)\,\mathrm{e}^{-\lambda_i\tau}\mathrm{d}\tau, \qquad (11.44)$$

где λ_i — величина, обратная времени жизни запаздывающих нейтронов m-й группы.

В дальнейшем будем предполагать, что плотность нейтронов в каждый момент времени по всему объему реактора одинакова, т. е. $N(M,t)\equiv N(t)$. Кроме того, будем учитывать только одну группу запаздывающих нейтронов ($m=1$). Уравнение (11.44) принимает вид

$$\frac{\mathrm{d}N}{\mathrm{d}t}=\frac{k_0-1}{T_0}N+\frac{k_1\lambda}{T_0}\int_0^\infty N(t-\tau)\,\mathrm{e}^{-\lambda\tau}\mathrm{d}\tau. \qquad (11.45)$$

(Здесь $\lambda=\lambda_1$.) Это интегродифференциальное уравнение можно свести к системе двух обыкновенных дифференциальных уравнений. Введем функцию $r(t)$ по формуле

$$r(t)=\frac{k_1}{T_0}\int\limits_0^\infty N(t-\tau)\,e^{-\lambda\tau}d\tau. \qquad (11.46)$$

Тогда вместо уравнения (11.45) можно записать следующую систему уравнений:

$$\frac{dN}{dt}=\frac{k_0-1}{T_0}N+\lambda r, \qquad (11.47)$$

$$\frac{dr}{dt}=\frac{k_1}{T_0}N-\lambda r. \qquad (11.48)$$

Будем считать, что до момента времени $t=0$ в реакторе протекал стационарный процесс диффузии нейтронов, причем значение N_0 плотности постоянно. Поэтому одно из начальных условий для системы (11.47), (11.48) имеет вид

$$N(t)|_{t=0}=N_0=\text{const}>0. \qquad (11.49)$$

Второе начальное условие получим из уравнения (11.48). В стационарном режиме $\frac{dr}{dt}=0$, $r=r_0=\text{const}$. Следовательно, из (11.48) получаем $\frac{k_1}{T_0}N_0=\lambda r_0$. Отсюда имеем

$$r(t)\,|_{t=0}=r_0=\frac{k_1}{T_0}N_0. \qquad (11.50)$$

В стационарном режиме общий коэффициент размножения равен критическому коэффициенту размножения, т. е. при $t<0$ выполнено равенство $k=k_0+k_1=k_{\text{кр}}$. Предположим теперь, что начиная с момента времени $t=0$ происходит произвольное изменение во времени коэффициента размножения: $k(t)=k_{\text{кр}}+\delta k(t)$. Поскольку $k(t)=k_0(t)+k_1(t)$, можно записать

$$k_0(t)=k_{\text{кр}}-k_1(t)+\delta k(t). \qquad (11.51)$$

Коэффициенты размножения k и k_1 считаем пропорциональными в каждый момент времени: $k_1(t)=\beta k(t)$, причем коэффициент пропорциональности β имеет смысл доли запаздывающих нейтронов в их общем потоке. Подставляя

$$k_1(t)=\beta[k_{\text{кр}}+\delta k(t)] \qquad (11.52)$$

в выражение (11.51), получаем

$$k_0(t)=[k_{\text{кр}}+\delta k(t)](1-\beta). \qquad (11.53)$$

Если рассматривается только одна группа запаздывающих нейтронов, $k_{\text{кр}}=1$. Действительно, полагая в (11.47), (11.48), что производные равны нулю, и складывая эти уравнения, находим для стационарного режима

$$\frac{k_0-k_1-1}{T_0}N_0=0.$$

В стационарном режиме $k_0+k_1=k_{\text{кр}}$, отсюда $k_{\text{кр}}=1$. Учитывая этот факт и подставляя (11.52) и (11.53) в уравнения (11.47), (11.48), получаем систему уравнений, описывающих нестационарный процесс диффузии нейтронов в реакторе:

$$\frac{dN(t)}{dt}=-\frac{\beta}{T_0}N(t)+\frac{1-\beta}{T_0}\delta k(t)N(t)+\lambda r(t); \tag{11.54}$$

$$\frac{dr(t)}{dt}=\frac{\beta}{T_0}N(t)+\frac{\beta}{T_0}\delta k(t)N(t)-\lambda r(t). \tag{11.55}$$

Для этих уравнений начальными являются условия (11.49), (11.50).

Функцию $r(t)$ в уравнениях (11.54), (11.55) надо понимать по-другому, чем в определении (11.46). Теперь величина k_1 является переменной во времени, поэтому необходимо положить

$$r(t)=\frac{1}{T_0}\int_0^\infty k_1(t-\tau)N(t-\tau)e^{-\lambda\tau}d\tau.$$

Перейдем к постановке задачи оптимального управления, а именно задачи о наиболее быстром переводе реактора из одного стационарного режима работы в другой. Пусть до момента $t=0$ в реакторе протекал стационарный процесс диффузии с постоянным значением $N_0>0$ плотности нейтронов в реакторе. При этом $k=k_{\text{кр}}=1$. Начиная с момента $t=0$ коэффициент размножения начинает изменяться и в реакторе протекает нестационарный процесс, описываемый уравнениями (11.54), (11.55) с начальными условиями (11.44), (11.50). Задача оптимального быстродействия состоит в выборе такой функции $\delta k(t)$, чтобы за минимальное время T плотность нейтронов в реакторе достигла заданного значения

$$N(t)|_{t=T}=N_1. \tag{11.56}$$

Затем начиная с момента $t=T$ необходимо обеспечить выполнение условия $k=k_{\text{кр}}=1$, т. е. чтобы в реакторе опять протекал стационарный процесс диффузии с постоянной плотностью нейтронов N_1. Условие, накладываемое на функцию $r(t)$ в момент времени $t=T$ (и во все моменты $t>T$), аналогично условию (11.50) и имеет вид

$$r(t)|_{t=T}=\frac{k_1}{T_0\lambda}N_1. \tag{11.57}$$

Значение коэффициента размножения нейтронов изменяется в результате изменения расположения поглотителя внутри реактора, поэтому варьировать величину $k(t)$ можно только в пределах, определяемых конструкцией реактора. В связи с этим необходимо наложить ограничения на функцию $\delta k(t)$

$$|\delta k(t)|\leqslant\delta k_{\max}. \tag{11.58}$$

Вообще говоря, необходимо наложить ограничение и на функцию $N(t)$, получающуюся при решении системы уравнений (11.54), (11.55), так как, согласно физическому смыслу, $N(t) \geqslant 0$. Однако можно строго доказать [11], что на самом деле эта система при любых $N(t)$ и $N_0 > 0$ имеет неотрицательное решение $N(t)$, т. е. вводить специальное ограничение на $N(t)$ не нужно.

Таким образом, задача оптимального быстродействия состоит в выборе такой кусочно-непрерывной функции $\delta k(t)$, удовлетворяющей условию (11.58), чтобы фазовая траектория системы $N(t)$, $r(t)$, проходящая в момент времени $t=0$ через точку $\left(N_0, \frac{k_1}{T_0\lambda} N_0\right)$, за минимальное время достигла точки $\left(N_1, \frac{k_1}{T_0\lambda} N_1\right)$. Это задача оптимального управления с закрепленными концами.

Функция $H(N, r, \psi_1, \psi_2, \delta k)$ для системы уравнений (11.54), (11.55) имеет вид

$$H=\psi_1\left(-\frac{\beta}{T_0} N+\frac{1-\beta}{T_0}\delta k N+\lambda r\right)+\psi_2\left(\frac{\beta}{T_0} N+\frac{\beta}{T_0}\delta k N-\lambda r\right)=$$

$$=(\psi_1-\psi_2)\left(-\frac{\beta}{T_0} N+\lambda r\right)+\left(\frac{1-\beta}{T_0}\psi_1 N+\frac{\beta}{T_0}\psi_2 N\right)\delta k.$$

Функции $\psi_1(t)$ и $\psi_2(t)$ удовлетворяют уравнениям $\frac{d\psi_1}{dt}=-\frac{\partial H}{\partial N}$, $\frac{\partial\psi_2}{dt}=-\frac{\partial H}{\partial r}$, т. е.

$$\frac{d\psi_1}{dt}=\left(\frac{\beta}{T_0}-\frac{1-\beta}{T_0}\delta k\right)\psi_1-\frac{\beta}{T_0}(1+\delta k)\psi_2,\ \frac{d\psi_2}{dt}=-\lambda\psi_1+\lambda\psi_2. \quad (11.59)$$

В соответствии с принципом максимума оптимальное управление $\delta k(t)$ следует выбирать так, чтобы функция H при каждом t имела максимум, т. е. чтобы достигала максимума функция $N\left(\frac{1-\beta}{T_0}\psi_1+\frac{\beta}{T_0}\psi_2\right)\delta k$. Функция $N(t)$ при всех δk удовлетворяет условию $N \geqslant 0$, поэтому максимум реализуется при

$$\delta k(t)=\delta k_{\max}\operatorname{sign}\left(\frac{1-\beta}{T_0}\psi_1+\frac{\beta}{T_0}\psi_2\right). \quad (11.60)$$

Воспользовавшись общей схемой решения задачи синтеза оптимальных по быстродействию управлений, установим вид фазовых траекторий системы, соответствующих случаям $\delta k(t)\equiv\delta k_{\max}$ и $\delta k(t)\equiv-\delta k_{\max}$. При $\delta k\equiv\delta k_{\max}$ уравнения (11.54), (11.55) принимают соответственно вид

$$\frac{dN}{dt}=-\frac{\beta}{T_0} N+\frac{1-\beta}{T_0}\delta k_{\max} N+\lambda r, \quad (11.61)$$

$$\frac{dr}{dt}=\frac{\beta}{T_0} N+\frac{\beta}{T_0}\delta k_{\max} N-\lambda r. \quad (11.62)$$

Найдем общее решение этой системы. Определяя из первого уравнения $r(t)$ и подставляя полученное выражение во второе, получаем уравнение вторго порядка относительно $N(t)$:

$$\frac{d^2N}{dt^2}+\left(\lambda+\frac{\beta}{T_0}-\frac{1-\beta}{T_0}\,\delta k_{\max}\right)\frac{dN}{dt}-\frac{\delta k_{\max}}{T_0}N=0. \qquad (11.63)$$

Запишем характеристическое уравнение:

$$\gamma^2+\left(\lambda+\frac{\beta}{T_0}-\frac{1-\beta}{T_0}\,\delta k_{\max}\right)\gamma-\frac{\lambda\delta k_{\max}}{T_0}=0. \qquad (11.64)$$

Результат решения системы (11.61), (11.62) зависит от того, какие корни, вещественные или комплексные, имеет уравнение (11.64). Для того чтобы дальнейшее исследование было более конкретным, предположим, что параметры рассматриваемого реактора имеют следующие значения: $\lambda\approx 0{,}1$; $\beta\approx 6{,}5\cdot 10^{-3}$, $T_0\approx 10^{-3}$, $\delta k_{\max}<\beta\ll 1$. В этом случае уравнение (11.64) имеет два вещественных корня: γ_1 и γ_2, причем $\gamma_1>0$, $\gamma_2<0$.

Общее решение уравнения (11.63) представляет собой линейную комбинацию экспонент $e^{\gamma_1 t}$ и $e^{\gamma_2 t}$. Решение системы (11.61), (11.62) с начальными условиями (11.49), (11.50) имеет вид

$$N(t)=C_1e^{\gamma_1 t}+C_2e^{\gamma_2 t},\quad r(t)=C_3e^{\gamma_1 t}+C_4e^{\gamma_2 t}. \qquad (11.65)$$

Постоянные C_1, C_2, C_3, C_4 можно выразить через коэффициенты уравнений и величины γ_1, γ_2, N_0. Однако при решении задачи синтеза необходимо перейти от параметрического задания фазовой траектории к явному выражению в координатах N, r. Осуществить такой переход, используя выражения для $N(t)$ и $r(t)$, сложно, поэтому получим уравнение фазовой траектории в координатах N, r с помощью уравнений (11.61), (11.62).

Преобразуем эти условия, вводя новые функции

$$y_1=(\lambda+\gamma_1)N+\lambda r,\quad y_2=(\lambda+\gamma_2)N+\lambda r. \qquad (11.66)$$

Вычислим производные $\frac{dy_1}{dt}$ и $\frac{dy_2}{dt}$, используя уравнения (11.61) и (11.62). Имеем

$$\frac{dy_1}{dt}=(\lambda+\gamma_1)\frac{dN}{dt}+\lambda\frac{dr}{dt}=\frac{\lambda}{T_0}\,\delta k_{\max}N-\frac{\gamma_1\beta}{T_0}N+$$
$$+\frac{\gamma_1}{T_0}\,\delta k_{\max}N-\frac{\gamma_1}{T_0}\,\delta k_{\max}N+\gamma_1\lambda r, \qquad (11.67)$$

$$\frac{dy_2}{dt}=(\lambda+\gamma_2)\frac{dN}{dt}+\lambda\frac{dr}{dt}=\frac{\lambda}{T_0}\,\delta k_{\max}N-\frac{\gamma_2\beta}{T_0}N+$$
$$+\frac{\gamma_2}{T_0}\,\delta k_{\max}N-\frac{\gamma_2\beta}{T_0}\,\delta k_{\max}N+\gamma_2\lambda r. \qquad (11.68)$$

С помощью уравнения (11.64) сумму второго, третьего и четвертого слагаемых в (11.67), (11.68) можно записать следующим образом:

$$-\frac{\gamma_{1,2}\beta}{T_0}N+\frac{\gamma_{1,2}}{T_0}\,\delta k_{\max}N-\frac{\gamma_{1,2}\beta}{T_0}\,\delta k_{\max}N=\gamma_{1,2}^2N+\lambda\gamma_{1,2}N-\frac{\lambda}{T_0}\,\delta k_{\max}N.$$

Тогда соотношения (11.67), (11.68) принимают соответственно вид

$$\frac{dy_1}{dt}=\gamma_1^2N+\lambda\gamma_1N+\gamma_1\lambda r=\gamma_1y_1, \tag{11.69}$$

$$\frac{dy_2}{dt}=\gamma_2^2N+\lambda\gamma_2N+\gamma_2\lambda r=\gamma_2y_2. \tag{11.70}$$

Почленно разделим первое из этих уравнений на второе: $\frac{dy_1}{dy_2}=\frac{\gamma_1y_1}{\gamma_2y_2}$. Полученное дифференциальное уравнение легко интегрируется после разделения переменных (при этом надо решать уравнение отдельно в каждой четверти плоскости y_1Oy_2). Общее решение можно записать в виде

$$(y_1)^{\gamma_2}=C(y_2)^{\gamma_1}, \tag{11.71}$$

где C — произвольная постоянная. Эта запись является условной, так как в случае, когда γ_1 и γ_2 — дробные числа, величины $(y_1)^{\gamma_1}$ и $(y_2)^{\gamma_2}$ могут не иметь смысла. При $y_1<0$, $y_2>0$ решение (11.71) имеет вид $y_1=C(y_2)^{\gamma_1/\gamma_2}$, при $y_1>0$, $y_2<0$ — $\left(\frac{y_1}{C}\right)^{\gamma_2/\gamma_1}=y_2$, при $y_1<0$, $y_2<0$ — $|y_1|^{\gamma_2}=C|y_2|^{\gamma_1}$

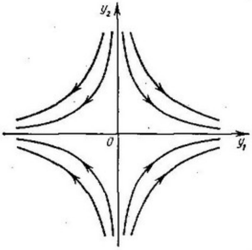

Рис. 11.9

При различных значениях константы C уравнение (11.71) описывает различные кривые из семейства фазовых траекторий, соответствующих случаю $\delta k\equiv\equiv\delta k_{\max}$. Это семейство в координатах y_1, y_2 изображено на рис. 11.9. С помощью уравнений (11.67), (11.70) нетрудно установить направление движения фазовой точки системы на каждой из кривой указанного семейства. При $y_1>0$, $y_2>0$ выполнены неравенства $\frac{dy_1}{dt}>0$, $\frac{dy_2}{dt}<0$ (так как $\gamma_1>0$, $\gamma_2<0$), т. е. с течением времени координата y увеличивается, а координата y_2 уменьшается. Поэтому фазовая точка движется в I четверти плоскости y_1Oy_2 слева направо. Аналогично рассматриваются случаи, когда фазовая точка системы находится в других четвертях плоскости. Направления движения фазовой точки указаны на рис. 11.9 стрелками.

Перейдем к системе координат N, r. Уравнения осей Oy_1 и Oy_2 в координатах N, r имеют вид

$$r=-\frac{\lambda+\gamma_2}{\lambda}N \quad (\text{ось } Oy_1), \tag{11.72}$$

$$r=-\frac{\lambda+\gamma_1}{\lambda}N \quad (\text{ось } Oy_2). \tag{11.73}$$

Покажем, что ось Oy_1 находится в I четверти плоскости NOr. Для этого достаточно доказать, что коэффициент $(\lambda+\gamma_2)/\lambda$ в уравнении (11.72) отрицателен, т. е. $\lambda+\gamma_2<0$. Поскольку γ_2 является меньшим корнем уравнения (11.64), можно записать

$$\gamma_2=-\frac{1}{2}\left[\lambda+\frac{\beta}{T_0}-\frac{1-\beta}{T_0}\delta k_{\max}+\right.$$

$$\left.+\sqrt{\left(\lambda+\frac{\beta}{T_0}-\frac{1-\beta}{T_0}\delta k_{\max}^2+\frac{\lambda^2\delta k_{\max}}{T_0^2}\right)}\right]<$$

$$<-\left(\lambda+\frac{\beta}{T_0}-\frac{1-\beta}{T_0}\delta k_{\max}\right).$$

Следовательно, $\lambda+\gamma_2<-\frac{\beta}{T_0}+\frac{1-\beta}{T_0}\delta k_{\max}$. Для рассматриваемого реактора на тепловых нейтронах выполнено неравенство $\delta k_{\max}<\beta$. Тогда из полученного неравенства следует, что $\lambda+\gamma_2<0$. Аналогично можно доказать, что ось Oy_2 находится во II четверти плоскости NOr.

Угол α, который образует ось Oy_1 и ось ON, в соответствии с (11.72) определяется соотношением $\operatorname{tg}\alpha=-\frac{\lambda+\gamma_2}{\lambda}$. Угол α_0, образованный ON и лучом OL, на котором лежат все начальные и конечные условия, в соответствии с (11.50) и (11.57) определяется соотношением $\operatorname{tg}\alpha_0=\frac{r_0}{N_0}=\frac{k_1}{T_0\lambda}$. При заданных параметрах реактора величины $\operatorname{tg}\alpha$ и $\operatorname{tg}\alpha_0$ удовлетворяют неравенству $\operatorname{tg}\alpha_0>\operatorname{tg}\alpha$, т. е. ось Oy_1 лежит ниже луча OL, поэтому каждая фазовая траектория семейства (11.71) пересекает луч OL (рис. 11.10).

Аналогично можно установить вид фазовых траекторий системы при $\delta k(t)\equiv-\delta k_{\max}$. В этом случае уравнение (11.64) имеет отрицательные корни $\gamma_2<\gamma_1<0$. Семейство фазовых траекторий (11.71) изображено на рис. 11.11. В I четверти плоскости $y'_1Oy'_2$ фазовые траектории представляют собой степенные функции $y'_2=C(y'_1)^{\gamma_2/\gamma_1}$, а в остальных четвертях получаются симметричным отражением относительно осей координат. Стрелками показаны направления движения фазовой точки системы с ростом t. На рис. 11.12 изображено то же семейство фазовых траекторий, но в плоскости NOr. Ось Oy_1' в этой плоскости лежит в I четверти, причем проходит выше луча OL, на котором находятся начальная (N_0, r_0) и конечная (N_1, r_1) точки задачи оптимального управления. Этот факт следует из того, что при заданных числовых параметрах реактора выполнено соотношение

$$\operatorname{tg}\alpha=-\frac{\lambda+\gamma_2}{\lambda}>\frac{\beta}{T_0\lambda}=\operatorname{tg}\alpha_0.$$

Таким образом, получены семейства фазовых траекторий, соот-

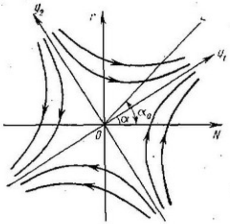

Рис. 11.10

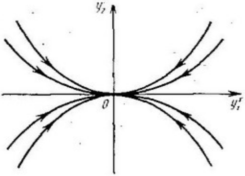

Рис. 11.11

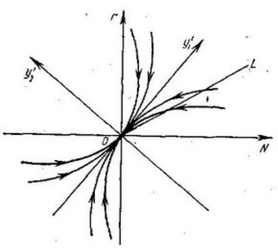

Рис. 11.12

ветствующих постоянным значениям δk_{max} и $-\delta k_{max}$ управляющего параметра $\delta k(t)$. Из рис. 11.10 и рис. 11.12 видно, что все фазовые траектории обоих рассматриваемых семейств, выходящие из луча OL, обязательно лежат внутри угла, который образуют оси Oy_1 и Oy'_1. Поэтому в дальнейшем будем рассматривать только траектории, лежащие внутри этого угла.

Для краткости обозначим начальную и конечную точки задачи оптимального управления через P_0 и P_1. Поскольку фазовые траектории каждого из семейств (при $\delta k \equiv \delta k_{max}$ и $\delta k \equiv -\delta k_{max}$) пересекают луч OL только один раз, управление $\delta k(t)$, не имеющее точек переключения и тождественно равное δk_{max} или $-\delta k_{max}$, не может являться решением задачи оптимального быстродействия. Функция $\delta k(t)$, имеющая одну точку переключения, может быть решением поставленной задачи. Таким решением могут быть и функции $\delta k(t)$ имеющие более одной точки переключения. Чтобы установить, сколько точек переключения имеет оптимальное управление, надо подробно исследовать вид функции $\varphi(t) = \frac{1-\beta}{T_0}\psi_1(t) + \frac{\beta}{T_0}\psi_2(t)$, которая входит в выражение (11.60). Используя уравнения (11.59), можно доказать [11], что функция $\varphi(t)$ имеет на интервале управления только одну точку перемены знака, поэтому оптимальное управление имеет только одну точку переключения.

Оптимальная фазовая траектория, по которой система за минимальное время переходит из точки P_0 в точку P_1, имеет следующий вид. Если точка P_0 лежит на луче OL ближе к началу координат, чем P_1, то сначала фазовая точка системы движется по фазовой

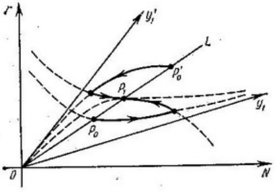

Рис. 11.13

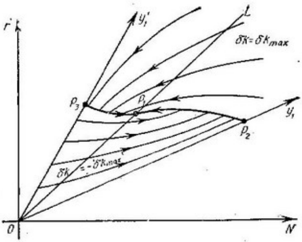

Рис. 11.14

траектории первого семейства до пересечения с той траекторией второго семейства, которая проходит через точку P_1 (рис. 11.13). Момент времени t_n, в который фазовая точка системы достигает указанной траектории второго семейства, является моментом переключения управления со значения $\delta k_{\max}$ на значение $-\delta k_{\max}$. При $t>t_n$ фазовая точка системы движется по траектории второго семейства и достигает P_1.

Если точка P'_0 лежит дальше от начала координат, чем P_1, то сначала фазовая точка системы движется по траектории из второго семейства под действием управления $\delta k \equiv -\delta k_{\max}$. В момент t_n, когда фазовая точка системы достигает траектории первого семейства, проходящей через точку P_1, происходит переключение управления со значения $-\delta k_{\max}$ на значение $\delta k_{\max}$. При $t>t_n$ фазовая точка системы движется по указанной траектории первого семейства и достигает P_1.

Дадим геометрическую интерпретацию решения задачи синтеза оптимальных по быстродействию управлений. Пусть зафиксирована конечная точка P_1 на луче OL, в которую должна попасть фазовая точка системы. Тогда при любом начальном положении на заключительном отрезке времени (после переключения управления) фазовая точка системы достигает P_1, двигаясь либо по линии P_2P_1 — части траектории из второго семейства, либо по линии P_3P_1 — части траектории второго семейства (рис. 11.14). Вся линия $P_3P_1P_2$ является линией переключения управления $\delta k(t)$. Во всех фазовых точках выше этой линии управление имеет значение $-\delta k_{\max}$, а во всех фазовых точках ниже этой линии — значение $\delta k_{\max}$. На части P_3P_1 линии переключения управление имеет значение $\delta k_{\max}$, а на части P_1P_2 — значение $-\delta k_{\max}$. Окончательно решение задачи синтеза запишем в следующем виде:

$$\delta k(t) = \begin{cases} \delta k_{\max} \text{ при } (N, r) \in Q_1 \text{ и при } (N, r) \in P_1P_3, \\ -\delta k_{\max} \text{ при } (N, r) \in Q_2 \text{ и при } (N, r) \in P_1P_2. \end{cases}$$

Здесь Q_2 — часть угла $y_1Oy'_1$ выше линии $P_3P_1P_2$, Q_1 — часть этого угла ниже этой линии.

ЗАКЛЮЧЕНИЕ

Процесс решения любой прикладной оптимизационной задачи состоит из двух этапов: сначала задачу формализуют, а затем с помощью соответствующего аналитического или численного метода решают. На этапе формализации строят математическую модель, т. е. конкретные технические, физические и экономические условия и требования воплощают в виде задачи оптимизации с конкретной целевой функцией и определенным допустимым множеством. Математическая модель должна быть адекватна исходной прикладной задаче, т. е. наиболее полно и точно соответствовать ей. Для этого необходимо учитывать все факторы, участвующие в задаче, что может привести к слишком сложной задаче оптимизации, на решение которой потребуется неоправданно много сил и средств. Поэтому этап формализации оказывается тесно связанным с последующим этапом. Нужно построить такую математическую модель, которая, с одной стороны, по возможности точно соответствует исходной прикладной задаче, а с другой — достаточно проста и за приемлемое время поддается решению доступными методами и средствами. В связи с этим, на этапе формализации исключительно важную роль играют интуиция, опыт и кругозор исследователя. В принципиальном отношении второй этап проще первого; в данной книге приведены методы, способствующие реализации именно второго этапа решения.

В книге представлен довольно широкий круг современных разделов теории оптимизации. Хотелось бы отметить те направления теории оптимизации, которые здесь не получили отражения. Вне поля зрения читателя осталась та часть теории, которая занимается исследованием экстремальных задач с учетом случайных факторов. Она носит название *стохастического программирования.* Особое место занимает *теория оптимального управления недетерминированными (стохастическими) системами.*

Если же говорить о детерминированных задачах, то в книге отражено большинство важнейших разделов современной теории оптимизации. Отсутствуют разделы, посвященные *дискретному программированию* и *теории оптимального управления системами с распределенными параметрами.* Предмет дискретного программирования составляют задачи минимизации и максимизации функции многих переменных на дискретном множестве, в частности, на множестве, состоящем из конечного числа элементов. Эти задачи имеют свою специфику, нередко носят ярко выраженный комбинаторный характер и требуют для решения специальных методов, существенно отличающихся, например, от методов линейного и нелинейного программирования. Теория оптимального управления системами с распределенными параметрами посвящена изучению задач оптимального управления объектами, поведение которых описывается дифференциальными уравнениями в частных производных [9]. Эта область теории оптимизации во многом еще находится на пути окончательного становления. Кроме того, здесь не рассмотрены весьма важные с практической точки зрения *методы регуляризации экстремальных задач*. Об этом разделе теории оптимизации можно получить представление из [9].

Необходимо отметить, что вторая часть книги, в которой рассматриваются задачи оптимизации в функциональных пространствах, включает, в основном, вопросы, связанные с необходимыми условиями оптимальности и почти не содержит описания соответствующих численных методов. В особенности это относится к главе по теории оптимального управления. Не следует думать, что такие методы вообще отсутствуют [12]. Описание этих методов чрезмерно увеличило и усложнило бы главы второй части в сравнении с главами первой части.

При написании книги авторы ставили перед собой две цели: ознакомить читателей с основами теории оптимизации и помочь им в применении этой теории на практике. Хочется верить, что эти цели хотя бы частично достигнуты.

ЛИТЕРАТУРА

1. *Алексеев В. М., Галеев Э. М., Тихомиров В. М.* Сборник задач по оптимизации. — М.: Наука, 1984.
2. *Алексеев В. М., Тихомиров В. М., Фомин С. В.* Оптимальное управление. — М.: Наука, 1979.
3. *Ашманов С. А.* Линейное программирование. — М.: Наука, 1981.
4. *Беллман Р.* Динамическое программирование. — М.: ИЛ, 1960.
5. *Болтянский В. Г.* Математические методы оптимального управления. — М.: Наука, 1969.
6. *Болтянский В. Г.* Оптимальное управление дискретными системами. — М.: Наука, 1973.
7. *Бояринов А. И., Кафаров В. В.* Методы оптимизации в химической технологии. — М.: Наука, 1969.
8. *Васильев Ф. П.* Численные методы решения экстремальных задач. — М.: Наука, 1980.
9. *Васильев Ф. П.* Методы решения экстремальных задач. — М.: Наука, 1981.
10. *Даффин Р., Питерсон Э., Зенер К.* Геометрическое программирование. — М.: Мир, 1972.
11. *Егоров А. И.* Оптимальное управление тепловыми и диффузионными процессами. — М.: Наука, 1978.
12. *Зубов В. И.* Лекции по теории управления. М.: Наука, 1975.
13. *Кини Р. Л., Райфа Х.* Принятие решений при многих критериях: предпочтения и замещения. — М.: Радио и связь, 1981.
14. *Колмогоров А. Н., Фомин С. В.* Элементы теории функций и функционального анализа. — М.: Наука, 1982.
15. *Коша А.* Вариационное исчисление. — М.: Высшая школа, 1983.
16. *Кудрявцев Л. Д.* Курс математического анализа. — М.: Высшая школа, 1981, т. 1—2.
17. *Моисеев Н. Н., Иванилов Ю. П., Столярова Е. М.* Методы оптимизации. — М.: Наука, 1978.
18. *Подиновский В. В., Ногин В. Д.* Парето-оптимальные решения многокритериальных задач. — М.: Наука, 1982.
19. *Поляк Б. Т.* Введение в оптимизацию. — М.: Наука, 1983.
20. *Понтрягин Л. С., Болтянский В. Г., Гамкрелидзе Р. В., Мищенко Е. Ф.* Математическая теория оптимальных процессов. — М.: Наука, 1984.
21. *Пшеничный Б. Н., Данилин Ю. М.* Численные методы в экстремальных задачах. — М.: Наука, 1975.

ОСНОВНЫЕ ОБОЗНАЧЕНИЯ

$\{x \in X \mid P(x)\}$ — совокупность элементов x множества X, для которых имеет место $P(x)$;
$\varnothing$ — пустое множество;
$x = (x_1, x_2, \ldots, x_n)$ — вектор-строка с компонентами $x_1, x_2, \ldots, x_n$;
$f: X \to Y$ — функция, отображающая X в Y;
$\mathbf{R}^n$ — n-мерное евклидово пространство;
$0_n = (0, 0, \ldots, 0)$ — нулевой вектор пространства $\mathbf{R}^n$;

$\langle x, y \rangle = \sum\limits_{i=1}^{n} x_i y_i$ — скалярное произведение векторов $x, y \in \mathbf{R}^n$;

$\| x \| = \sqrt{\langle x, x \rangle}$ — норма вектора $x \in \mathbf{R}^n$;
$U_\varepsilon(x)$ — ε-окрестность точки $x \in \mathbf{R}^n$;
$A^\top$ — транспонированная матрица A;

$\nabla f(x) = \left(\dfrac{\partial f(x)}{\partial x_1}, \dfrac{\partial f(x)}{\partial x_2}, \ldots, \dfrac{\partial f(x)}{\partial x_n}\right)^\top$ — градиент функции f, вычисленный в точке $x \in \mathbf{R}^n$;
$\nabla f(x)$ — градиент функционала f, вычисленный на элементе x;

$\nabla^2 f(x) = \left(\dfrac{\partial^2 f(x)}{\partial x_i \partial x_j}\right)_{n \times n}$ — гессиан функции f, вычисленный в точке $x \in \mathbf{R}$;

$C_0[a, b]$ — линейное пространство непрерывных функций на отрезке $[a, b]$;
$C_n[a, b]$ — линейное пространство n раз непрерывно дифференцируемых функций на отрезке $[a, b]$;
$C_n^m[a, b]$ — линейное пространство m-мерных вектор-функций, компоненты которых n раз непрерывно дифференцируемы на $[a, b]$;
$\|x\|_0$ — норма элемента $x \in C_0[a, b]$;
$\|x\|_1$ — норма элемента $x \in C_1[a, b]$;
$U_\varepsilon^0(x)$ — ε-окрестность нулевого порядка элемента $x \in C_0[a, b]$;
$U_\varepsilon^1(x)$ — ε-окрестность первого порядка элемента $x \in C_1[a, b]$.

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

ОГЛАВЛЕНИЕ