

ОДНОРАНГОВАЯ АППРОКСИМАЦИЯ ПОЛОЖИТЕЛЬНЫХ МАТРИЦ С ИСПОЛЬЗОВАНИЕМ МЕТОДОВ ТРОПИЧЕСКОЙ МАТЕМАТИКИ¹

Кривулин Н. К., профессор кафедры статистического моделирования
СПбГУ, nkk@math.spbu.ru;

Романова Е. Ю., студент кафедры статистического моделирования
СПбГУ, romanova.ej@gmail.com

Аннотация

В статье рассматриваются проблемы аппроксимации матриц матрицами единичного ранга. Задача аппроксимации формулируется как задача минимизации $\log$ -чебышевского расстояния, которая затем сводится к задаче оптимизации, имеющей компактное представление в терминах тропической математики. Приводятся необходимые определения и результаты из области тропической математики, на основе которых дается решение исходной задачи аппроксимации.

Введение

К задаче аппроксимации матриц сводится значительное число прикладных задач из разных областей. Многие вычислительные задачи требуют решения системы линейных алгебраических уравнений. Например, задачи вычислительной гидродинамики, теории электрических цепей, уравнения балансов и сохранения в механике. Методы решения систем линейных уравнений принято разделять на итерационные и прямые. Прямые методы обычно основываются на LU -разложении и требуют больших затрат памяти и временных ресурсов. Применение техники малоранговой аппроксимации к множителям LU -разложения, изложенное в работе [1], значительно повышает эффективность этих методов. Схожий подход может быть применен и к решению задачи итерационными методами. Например, в [2] описано использование приближения LDL^T -разложения, полученного на основе малоранговой аппроксимации, в качестве предобуславливателя. Потребность в аппроксимации возникает и при обработке массивов

¹Работа выполнена при финансовой поддержке РГНФ, проект №16-02-00059.

данных. Матрицы, заполненные результатами какого-либо физического эксперимента, биологическими наблюдениями или оценками пользователей, могут иметь пропуски или значения, с которыми сложно работать. Аппроксимация матрицами из выбранного множества матриц дает возможность работать с данными в удобной и корректной с математической точки зрения форме.

Понижение ранга матрицы при помощи аппроксимации существенно упрощает ее структуру и позволяет сократить объем памяти, требующийся для её хранения. Логично выделять аппроксимацию матрицами единичного ранга, так как они устроены наиболее просто. Некоторые методы одноранговой аппроксимации описаны, например, в работах [3], [4].

Задача аппроксимации матрицы $\mathbf{A} \in \mathbb{R}^{n \times n}$ матрицами $\mathbf{X} \in \mathcal{S} \subset \mathbb{R}^{n \times n}$ формулируется как задача оптимизации

$$\min_{\mathbf{X} \in \mathcal{S}} d(\mathbf{A}, \mathbf{X}),$$

где d — функция расстояния на множестве матриц, измеряющая величину ошибки аппроксимации.

Подходы к решению задачи аппроксимации могут варьироваться в зависимости от нюансов исходной задачи и особенностей матрицы. Различия между подходами во многом определяются выбором функции расстояния.

Распространенным решением проблемы является применение к аппроксимации матриц разновидностей метода наименьших квадратов, в основе которого лежит минимизация евклидова расстояния. Вариант применения описан, например, в работе [5]. Метод надежен, но требует больших затрат вычислительных ресурсов, что делает его малоприменимым для решения задач больших размерностей или задач, в которых проблема экономии ресурсов является первостепенной. В [6] освещается использование расстояния Минковского (l_p) и расстояния Чебышева, которое рассматривается как предел расстояния Минковского при $p \rightarrow \infty$. В частности, в этой работе доказывается существование приближения Чебышева с рангом r для любой матрицы $\mathbf{A}$ с большим рангом. Но использование функции расстояния Минковского при $p > 2$ еще более трудоемко, чем евклидовой функции расстояния.

В работе [7] проблема чебышевской аппроксимации сформулирована в виде задачи линейного программирования, к решению которой могут применяться соответствующие методы, например, симплекс-метод. Для аппроксимации положительных матриц иногда целесообразнее перейти к оценке погрешности в логарифмической шкале. Задача ми-

нимизации log-чебышевского расстояния может быть сведена к задаче конического программирования второго порядка, как в работе [8], и решена, например, барьерным методом [7].

Далее в статье предлагается метод аппроксимации положительных матриц матрицами единичного ранга путем минимизации log-чебышевского расстояния между матрицами. Будет показано, что задача минимизации log-чебышевского расстояния может быть приведена к задаче, записанной в компактной форме в терминах идемпотентного полуполя $\mathbb{R}_{\max, \times}$, которое часто называют max-алгеброй. Затем для нахождения решения будут использованы результаты из области тропической математики.

Log-чебышевская одноранговая аппроксимация

Чебышевская аппроксимация положительной матрицы $\mathbf{A} = (a_{ij})$ при помощи положительной матрицы $\mathbf{X} = (x_{ij})$ в логарифмической шкале использует функцию расстояния

$$d(\mathbf{A}, \mathbf{X}) = \max_{i,j} |\log a_{ij} - \log x_{ij}|,$$

где логарифм берется по основанию больше единицы.

Справедливо следующее утверждение

Утверждение 1. Пусть $\mathbf{A}, \mathbf{X}$ — положительные матрицы. Минимизация по $\mathbf{X}$ величины $d(\mathbf{A}, \mathbf{X})$ эквивалентна минимизации

$$d'(\mathbf{A}, \mathbf{X}) = \max_{i,j} \max(a_{ij}x_{ij}^{-1}, x_{ij}a_{ij}^{-1}).$$

Следовательно, задача log-чебышевской аппроксимации может быть сведена к задаче

$$\min_{\mathbf{X}} d'(\mathbf{A}, \mathbf{X}). \quad (1)$$

В силу того, что любая матрица $\mathbf{X}$ ранга 1 имеет представление $\mathbf{X} = \mathbf{s}\mathbf{t}^T$, где векторы $\mathbf{s} = (s_i)$ и $\mathbf{t} = (t_j)$ не содержат нулевых элементов, целевую функцию задачи (1) можно записать в виде

$$d'(\mathbf{A}, \mathbf{X}) = d'(\mathbf{A}, \mathbf{s}\mathbf{t}^T) = \max_{i,j} \max(s_i^{-1}a_{ij}t_j^{-1}, s_ia_{ij}^{-1}t_j).$$

Таким образом, задача одноранговой аппроксимации сводится к задаче

$$\min_{\mathbf{s}, \mathbf{t}} \max_{i,j} \max(s_i^{-1}a_{ij}t_j^{-1}, s_ia_{ij}^{-1}t_j). \quad (2)$$

Элементы тропической математики

Приведем основные определения, обозначения и предварительные результаты тропической математики [9], на которые будем опираться в дальнейшем.

Идемпотентное полуполе

Идемпотентным полуполем называется алгебраическая система $(\mathbb{X}, \oplus, \otimes, 0, 1)$, где $\mathbb{X}$ — непустое множество, которое замкнуто относительно операций сложения $\oplus$ и умножения $\otimes$ и включает их нейтральные элементы 0 и 1 . Сложение является идемпотентным, то есть удовлетворяет условию $x \oplus x = x$ для всех $x \in \mathbb{X}$. Выполняется свойство дистрибутивности умножения относительно сложения и для каждого $x \neq 0$ существует обратный по умножению элемент x^{-1} такой, что $x^{-1} \otimes x = 1$.

Например, в вещественном полуполе $\mathbb{R}_{\max, \times} = (\mathbb{R}_+, \max, \times, 0, 1)$, где $\mathbb{R}_+$ — множество неотрицательных вещественных чисел, операция сложения определена как взятие максимума двух чисел и имеет нейтральный элемент 0 , а умножение $\otimes$ определено как арифметическое умножение с нейтральным элементом 1 . Понятия обратного элемента и степени имеют обычный смысл.

Матрицы и векторы

Множество всех матриц, которые имеют m строк и n столбцов с элементами из $\mathbb{X}$, обозначается через $\mathbb{X}^{m \times n}$. Матрица, все элементы которой равны 0 , называется нулевой и обозначается $\mathbf{0}$. Квадратная матрица, диагональные элементы которой равны числу 1 , а недиагональные — числу 0 , называется единичной и обозначается $\mathbf{I}$. Матрица называется неразложимой, если перестановкой строк вместе с такой же перестановкой столбцов ее нельзя привести к блочно-треугольному виду. Сложение и умножение двух матриц подходящего размера и умножение матрицы на число выполняются по стандартным правилам с заменой обычных арифметических операций на операции $\oplus$ и $\otimes$.

Для любой ненулевой матрицы $\mathbf{A} = (a_{ij}) \in \mathbb{X}^{m \times n}$ определена мультипликативно сопряженная матрица $\mathbf{A}^- = (a_{ij}^-) \in \mathbb{X}^{n \times m}$ с элементами $a_{ij}^- = a_{ji}^{-1}$, если $a_{ji} \neq 0$, и $a_{ij}^- = 0$ — в противном случае.

След матрицы $\mathbf{A} = (a_{ij}) \in \mathbb{X}^{n \times n}$ вычисляется по формуле

$$\text{tr} \mathbf{A} = a_{11} \oplus \cdots \oplus a_{nn}.$$

Для любой матрицы $\mathbf{A} \in \mathbb{X}^{n \times n}$ введем в рассмотрение матрицу $\mathbf{A}^* = \mathbf{I} \oplus \mathbf{A} \oplus \cdots \oplus \mathbf{A}^{n-1}$.

Множество всех векторов-столбцов размера n с элементами из $\mathbb{X}$ обозначается $\mathbb{X}^n$. Вектор, все элементы которого равны 0 , называется нулевым. Вектор называется регулярным, если он не имеет нулевых компонент. Для любого ненулевого вектора $\mathbf{x} = (x_i) \in \mathbb{X}^n$ определен вектор-строка $\mathbf{x}^- = (x_i^-)$, где $x_i^- = x_i^{-1}$, если $x_i \neq 0$, и $x_i^- = 0$ — иначе.

Собственное число и вектор матрицы

Число $\lambda \in \mathbb{X}$ и ненулевой вектор $\mathbf{x} \in \mathbb{X}^n$ называются собственным значением и собственным вектором матрицы $\mathbf{A} \in \mathbb{X}^{n \times n}$, если они удовлетворяют равенству

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x}.$$

Любая матрица $\mathbf{A}$ порядка n имеет собственное число, которое вычисляется по формуле

$$\lambda = \bigoplus_{m=1}^n \text{tr}^{1/m}(\mathbf{A}^m).$$

Если у матрицы $\mathbf{A}$ есть другие собственные числа, то они по величине не превосходят числа λ , которое называется спектральным радиусом матрицы.

Задача тропической оптимизации и ее решение

Предположим, что задана матрица $\mathbf{A} \in \mathbb{X}^{n \times n}$ и требуется решить задачу минимизации

$$\min_{\mathbf{x}, \mathbf{y}} \mathbf{x}^- \mathbf{A} \mathbf{y} \oplus \mathbf{y}^- \mathbf{A}^- \mathbf{x}, \quad (3)$$

где минимум берется по всем регулярным векторам $\mathbf{x}, \mathbf{y} \in \mathbb{X}^n$.

В работе [9] получен следующий результат

Лемма 1. Пусть $\mathbf{A} \in \mathbb{X}^{n \times n}$ — неразложимая матрица, μ — спектральный радиус матрицы $\mathbf{A}\mathbf{A}^-$. Тогда минимум в задаче (3) равен $\mu^{1/2}$ и достигается тогда, когда $\mathbf{x}$ и $\mathbf{y} = \mu^{-1/2} \mathbf{A}^- \mathbf{x}$ — собственные векторы матриц $\mathbf{A}\mathbf{A}^-$ и $\mathbf{A}^- \mathbf{A}$, соответствующие μ .

Следующая теорема дает полное решение задачи (3).

Теорема 1. Пусть $\mathbf{A} \in \mathbb{X}^{n \times n}$, μ — спектральный радиус матрицы $\mathbf{A}\mathbf{A}^-$. Тогда минимум в задаче (3) равен $\mu^{1/2}$ и достигается тогда и только тогда, когда

$$\begin{aligned}\mathbf{x} &= (\mu^{-1}\mathbf{A}\mathbf{A}^-)^*\mathbf{v} \oplus \mu^{-1/2}\mathbf{A}(\mu^{-1}\mathbf{A}^-\mathbf{A})^*\mathbf{w}, \\ \mathbf{y} &= \mu^{-1/2}\mathbf{A}^-(\mu^{-1}\mathbf{A}\mathbf{A}^-)^*\mathbf{v} \oplus (\mu^{-1}\mathbf{A}^-\mathbf{A})^*\mathbf{w},\end{aligned}$$

где $\mathbf{v}, \mathbf{w}$ — произвольные регулярные векторы размера n .

В частности, минимум достигается, когда $\mathbf{x}$ и $\mathbf{y} = \mu^{-1/2}\mathbf{A}^-\mathbf{x}$ — собственные векторы матриц $\mathbf{A}\mathbf{A}^-$ и $\mathbf{A}^-\mathbf{A}$, соответствующие μ .

Решение задачи аппроксимации

Рассмотрим задачу одноранговой аппроксимации (2). При замене арифметических операций на тропические, получим задачу

$$\min_{i,j} \bigoplus_{i,j} (s_i^{-1}a_{ij}t_j^{-1} \oplus s_ia_{ij}^{-1}t_j). \quad (4)$$

Целевую функцию задачи (4) можно записать в виде

$$\bigoplus_{i,j} (s_i^{-1}a_{ij}t_j^{-1} \oplus s_ia_{ij}^{-1}t_j) = \mathbf{s}^-\mathbf{A}(\mathbf{t}^-)^T \oplus \mathbf{t}^T\mathbf{A}^-\mathbf{s}.$$

Таким образом, задача (4) принимает вид

$$\min_{\mathbf{s}, \mathbf{t}} \mathbf{s}^-\mathbf{A}(\mathbf{t}^-)^T \oplus \mathbf{t}^T\mathbf{A}^-\mathbf{s}. \quad (5)$$

Положив в задаче (5) $\mathbf{s} = \mathbf{x}$, $\mathbf{t} = (\mathbf{y}^-)^T$, получим задачу тропической оптимизации в форме (3). Применение к ней теоремы 1 дает решение в виде следующего утверждения

Утверждение 2. Пусть $\mathbf{A} \in \mathbb{X}^{n \times n}$, μ — спектральный радиус матрицы $\mathbf{A}\mathbf{A}^-$. Тогда минимальная погрешность аппроксимации матрицы $\mathbf{A}$ матрицами единичного ранга равна $\mu^{1/2}$ и достигается на матрицах вида $\mathbf{s}\mathbf{t}^T$, где

$$\begin{aligned}\mathbf{s} &= (\mu^{-1}\mathbf{A}\mathbf{A}^-)^*\mathbf{v} \oplus \mu^{-1/2}\mathbf{A}(\mu^{-1}\mathbf{A}^-\mathbf{A})^*\mathbf{w}, \\ \mathbf{t}^T &= (\mu^{-1/2}\mathbf{A}^-(\mu^{-1}\mathbf{A}\mathbf{A}^-)^*\mathbf{v} \oplus (\mu^{-1}\mathbf{A}^-\mathbf{A})^*\mathbf{w})^-, \end{aligned}$$

и v, w — произвольные регулярные векторы размера n .

В частности, минимальная погрешность достигается, когда s — собственный вектор матрицы AA^+ , соответствующий μ , а $t^T = \mu^{1/2}(A^+s)^+$.

Литература

- [1] Соловьев С. Решение разреженных систем линейных уравнений методом Гаусса с использованием техники аппроксимации матрицами малого ранга // Выч. мет. программирование. 2014. Т. 15, № 3. С. 441–460.
- [2] Воронин К., Соловьев С. Решение уравнения Гельмгольца с использованием метода малоранговой аппроксимации в качестве предобуславливателя // Выч.мет. программирование. 2015. Т. 16. С. 268–280.
- [3] Luss R., Teboulle M. Conditional gradient algorithms for rank-one matrix approximations with a sparsity constraint // SIAM Review. 2013. Vol. 55, no. 1. P. 65–98.
- [4] Ispany M., Michaletzky G., Reiczigel J. Approximation of non-negative integer-valued matrices with application to Hungarian mortality data // MTNS 2010 Conference. Budapest: 2010. July.
- [5] Саати Т. Принятие решений. Метод анализа иерархий. Москва: Радио и связь, 1993. С. 217–219.
- [6] Zietak K. The Chebyshev approximation of a rectangular matrix by matrices of smaller rank as the limit of l_p -approximation // J. Comput. Appl. Math. 1984. Vol. 11. P. 297–305.
- [7] Boyd S., Vandenberghe L. Convex optimization. Cambridge University Press, 2004.
- [8] Lobo M., Vandenberghe L., Boyd S. Applications of second-order cone programming // Linear Algebra Appl. 1998. Vol. 284. P. 193–228.
- [9] Кривулин Н. Методы идемпотентной алгебры в задачах моделирования и анализа сложных систем. СПб: С.-Петерб. ун-т, 2009. С. 107–108.