Решение многомерной задачи тропической оптимизации с приложением к одноранговой аппроксимации положительных матриц[1]

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

Мартынкина Е.С., студентка кафедры статистического моделирования СПбГУ, katyamartynkina7@gmail.com

Аннотация

В работе предлагается полное решение многомерной задачи оптимизации, сформулированной в терминах тропической (идемпотентной) математики. Сначала приводятся основные определения и результаты идемпотентной алгебры, необходимые для построения решения. Затем для целевой функции находится достижимая нижняя граница, которая связана с вычислением спектрального радиуса матрицы. Получено множество всех решений задачи многомерной тропической оптимизации, для чего рассматриваемая задача сводится к уже известной задаче. Для некоторых частных случаев сформулированы следствия из полученного результата. В качестве приложения рассматривается задача одноранговой аппроксимации положительных матриц, возникающая, например, в области машинного обучения, технического зрения и в статистике.

Введение

Тропическая (идемпотентная) математика охватывает область, связанную с изучением свойств полуколец с идемпотентным сложением и их приложениями. Исследованиям в этой области посвящено ряд работ, включая [1–4]. Важным направлением развития этой области является разработка методов и алгоритмов решения задач оптимизации, сформулированных в терминах тропической математики.

[1]Работа выполнена при финансовой поддержке РФФИ, проект №18-010-00723.

Существует широкий класс задач оптимизации, в которых целевая функция и ограничения выражаются при помощи операций максимума и минимума, а также арифметических операций. К этому классу относятся, например, задачи сетевого планирования [5], минимаксные задачи размещения объектов в пространстве [6]. Во многих случаях решение подобных задач можно упростить путем их представления на языке тропической математики и использования ее результатов. В настоящей работе методы тропической оптимизации применяются к решению задачи аппроксимации положительных матриц матрицами единичного ранга.

Работа построена следующим образом. Сначала решается многомерная задача тропической оптимизации: находится минимум целевой функции, описывается множество решений, на которых этот минимум достигается. В качестве следствия из полученных результатов построены решения для некоторых частных случаев. Показано, как задача одноранговой аппроксимации положительных матриц может быть сведена к рассмотренным частным случаям.

Элементы тропической математики

В этом разделе приводятся основные понятия и результаты тропической (идемпотентной) математики [1–4]. Идемпотентное полуполе определяется как набор $\langle\mathbb{X},\oplus,\odot,\mathbb{0},\mathbb{1}\rangle$, где $\mathbb{X}$ обозначает непустое множество с заданными на нем ассоциативными и коммутативными операциями сложения $\oplus$ с нейтральным элементом $\mathbb{0}$ (ноль) и умножения $\odot$ с нейтральным элементом $\mathbb{1}$ (единица).

Сложение обладает свойством идемпотентности: для любого элемента $x\in\mathbb{X}$ выполняется равенство $x\oplus x=x$. Умножение дистрибутивно относительно сложения и обратимо: для любого $x\ne\mathbb{0}$ существует x^{-1} такой, что $x\odot x^{-1}=\mathbb{1}$.

Идемпотентность сложения задает на множестве $\mathbb{X}$ частичный порядок: $x\le y$ тогда и только тогда, когда $x\oplus y=y$. Отсюда следует, что неравенство $x\oplus y\le z$ эквивалентно двум неравенствам $x\le z$ и $y\le z$. Кроме того, операции сложения и умножения являются монотонными. Будем предполагать, что указанный частичный порядок продолжен до линейного на множестве $\mathbb{X}$. Ниже операция $\max$ понимается в смысле указанного линейного порядка.

Целые степени определяются обычным образом: $x^0=\mathbb{1}$, $x^p=x\odot x^{p-1}$, $x^{-p}=(x^{-1})^p$, $\mathbb{0}^p=\mathbb{0}$ для всех $x\ne\mathbb{0}$ и натуральных p. В дальнейшем будем считать полуполе алгебраически полным в том смысле, что для любого $a \in \mathbb{X}$ и натурального p существует единственное решение уравнения $x^p = a$, из чего вытекает что введенная операция возведения в степень может быть распространена на случай степеней с рациональным показателем. Далее символ $\odot$ для упрощения записи будем опускать.

Рассмотрим полуполе $\mathbb{R}_{\max,\times} = \langle \mathbb{R}_+ \cup \{0\}, \max, \times, 0, 1\rangle$. В нем сложение определено как операция $\max$, умножение определено как обычно, роль нуля $\mathbb{0}$ играет арифметический нуль, а единицы $\mathbb{1}$ — число 1. Обратный элемент и степень имеют обычный смысл. Порядок, индуцированный идемпотентным сложением, совпадает с обычным линейным порядком, заданным на $\mathbb{R}$.

Обозначим через $\mathbb{X}^{m\times n}$ множество матриц размера $m \times n$ над полуполем $\mathbb{X}$. Нулевая матрица имеет все элементы равными нулю и обозначается символом $\mathbf{0}$. Операции сложения $\oplus$ и умножения $\odot$ матриц выполняются по стандартным правилам с заменой обычных скалярных операций сложения и умножения на тропические. Свойства монотонности скалярных операций распространяются на операции над матрицами, для которых отношение $\leq$ вводится как покомпонентное.

Рассмотрим множество $\mathbb{X}^{n\times n}$ квадратных матриц порядка n. Как обычно, матрица является диагональной, если все ее недиагональные элементы равны нулю. Единичной называется диагональная матрица $\boldsymbol{I}$, у которой все элементы на диагонали равны единице. Матрица $\boldsymbol{A}^{-1}$ называется обратной по отношению к $\boldsymbol{A}$, если $\boldsymbol{A}^{-1}\boldsymbol{A} = \boldsymbol{A}\boldsymbol{A}^{-1} = \boldsymbol{I}$.

Мультипликативно сопряженным транспонированием ненулевой матрицы $\boldsymbol{A} \in \mathbb{X}^{m\times n}$ называется преобразование в матрицу $\boldsymbol{A}^- \in \mathbb{X}^{n\times m}$ с элементами $a_{ij}^- = a_{ji}^{-1}$, если $a_{ji} \neq \mathbb{0}$, и $a_{ij} = \mathbb{0}$ в противном случае. Если для матрицы $\boldsymbol{A}$ существует обратная, то $\boldsymbol{A}^- = \boldsymbol{A}^{-1}$.

Операция возведения в целую положительную степень матрицы вводится обычным путем. Для любой матрицы $\boldsymbol{A}$ и целого $p > 0$ имеем $\boldsymbol{A}^0 = \boldsymbol{I}$, $\boldsymbol{A}^p = \boldsymbol{A}^{p-1}\boldsymbol{A} = \boldsymbol{A}\boldsymbol{A}^{p-1}$, $\mathbf{0}^p = \mathbf{0}$.

Следом квадратной матрицы $\boldsymbol{A} = (a_{ij}) \in \mathbb{X}^{n\times n}$ называется число $\operatorname{tr}\boldsymbol{A} = a_{11} \oplus \cdots \oplus a_{nn}$.

Любая матрица $\boldsymbol{A}$ порядка n имеет спектральный радиус (максимальное собственное число)

$$\varrho(\boldsymbol{A}) = \bigoplus_{m=1}^{n} \operatorname{tr}^{1/m}(\boldsymbol{A}^m). \tag{1}$$

Для любой квадратной матрицы $\boldsymbol{A}$ порядка n определим матрицу

$$\boldsymbol{A}^* = \boldsymbol{I} \oplus \boldsymbol{A} \oplus \cdots \oplus \boldsymbol{A}^{n-1}. \tag{2}$$

Матрица, состоящая из одного столбца (строки), образует вектор-столбец (вектор-строку). Далее все векторы, если не указано иначе, считаются вектор-столбцами. Множество вектор-столбцов размерности n обозначим через $\mathbb{X}^n$. Нулевой вектор имеет все компоненты равными $\mathbb{0}$. Вектор называется регулярным, если у него нет нулевых компонент.

Пусть задана матрица $\boldsymbol{A} \in \mathbb{X}^{n \times n}$ и требуется найти все регулярные векторы $\boldsymbol{x} \in \mathbb{X}^n$, которые решают задачу

$$\min_{\boldsymbol{x}} \quad \boldsymbol{x}^{-} \boldsymbol{A} \boldsymbol{x}. \tag{3}$$

Решение этой задачи обеспечивается следующим утверждением, полное доказательство которого приводится в работе [7].

Лемма 1 (Об экстремальном свойстве спектрального радиуса)**.** *Пусть $\boldsymbol{A}$ — матрица со спектральным радиусом $\lambda > \mathbb{0}$. Тогда минимум в задаче* (3) *равен λ, а все регулярные решения задачи имеют вид*

$$\boldsymbol{x} = \left(\lambda^{-1} \boldsymbol{A}\right)^{*} \boldsymbol{u}, \quad \boldsymbol{u} \in \mathbb{X}^n.$$

Многомерная задача тропической оптимизации

Задачи тропической оптимизации обычно состоят в минимизации или максимизации некоторой целевой функции, заданной на векторах над идемпотентным полуполем [8,9]. В настоящей работе рассматривается следующая многомерная задача тропической оптимизации.

Предположим, что заданы матрицы $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{X}^{n \times n}$. Требуется найти регулярные векторы $\boldsymbol{x}$, $\boldsymbol{y} \in \mathbb{X}^n$, на которых достигается

$$\min_{\boldsymbol{x}, \boldsymbol{y}} \quad \boldsymbol{y}^{-} \boldsymbol{A} \boldsymbol{x} \boldsymbol{x}^{-} \boldsymbol{B} \boldsymbol{y}. \tag{4}$$

В работе получен следующий результат, который дает полное решение задачи (4).

Теорема 1. *Пусть $\boldsymbol{A}$ и $\boldsymbol{B}$ — матрицы порядка n, $\mu > \mathbb{0}$ — спектральный радиус матрицы $\boldsymbol{A}\boldsymbol{B}$. Тогда минимум в задаче* (4) *равен μ, а все регулярные решения имеют вид*

$$\begin{aligned} \boldsymbol{x} &= \alpha(\mu^{-1/2} \left(\mu^{-1} \boldsymbol{B}\boldsymbol{A}\right)^{*} \boldsymbol{B}\boldsymbol{v} \oplus (\mu^{-1} \boldsymbol{B}\boldsymbol{A})^{*} \boldsymbol{w}), \\ \boldsymbol{y} &= \beta((\mu^{-1} \boldsymbol{A}\boldsymbol{B})^{*} \boldsymbol{v} \oplus \mu^{-1/2} \left(\mu^{-1} \boldsymbol{A}\boldsymbol{B}\right)^{*} \boldsymbol{A}\boldsymbol{w}), \end{aligned}$$

при любых $\alpha > \mathbb{0}$, $\beta > \mathbb{0}$ и $\boldsymbol{v}, \boldsymbol{w} \in \mathbb{X}^n$.

Теперь сформулируем следствия из полученного решения задачи для частных случаев $\boldsymbol{B} = \boldsymbol{A}^{-}$ и $\boldsymbol{B} = \boldsymbol{A}$.

Предположим, что задана матрица $\boldsymbol{A} \in \mathbb{X}^{n\times n}$. Требуется найти регулярные векторы $\boldsymbol{x}$, $\boldsymbol{y} \in \mathbb{X}^{n}$, на которых достигается

$$\min_{\boldsymbol{x},\boldsymbol{y}} \quad \boldsymbol{y}^{-}\boldsymbol{A}\boldsymbol{x}\boldsymbol{x}^{-}\boldsymbol{A}^{-}\boldsymbol{y}. \tag{5}$$

Следствие 1. *Пусть $\boldsymbol{A}$ — матрица порядка n, $\mu > \mathbb{0}$ — спектральный радиус матрицы $\boldsymbol{A}\boldsymbol{A}^{-}$. Тогда минимум в задаче* (5) *равен μ, а все регулярные решения имеют вид*

$$\begin{aligned}
\boldsymbol{x} &= \alpha\big(\left(\mu^{-1}\boldsymbol{A}^{-}\boldsymbol{A}\right)^{*}\boldsymbol{v} \oplus \mu^{-1/2}\left(\mu^{-1}\boldsymbol{A}^{-}\boldsymbol{A}\right)^{*}\boldsymbol{A}^{-}\boldsymbol{w}\big),\\
\boldsymbol{y} &= \beta\big(\mu^{-1/2}\left(\mu^{-1}\boldsymbol{A}\boldsymbol{A}^{-}\right)^{*}\boldsymbol{A}\boldsymbol{v} \oplus (\mu^{-1}\boldsymbol{A}\boldsymbol{A}^{-})^{*}\boldsymbol{w}\big),
\end{aligned}$$

при любых $\alpha > \mathbb{0}$, $\beta > \mathbb{0}$ и $\boldsymbol{v}, \boldsymbol{w} \in \mathbb{X}^{n}$.

Пусть задана матрица $\boldsymbol{A} \in \mathbb{X}^{n\times n}$. Рассмотрим задачу нахождения регулярных векторов $\boldsymbol{x}$, $\boldsymbol{y} \in \mathbb{X}^{n}$, на которых достигается

$$\min_{\boldsymbol{x},\boldsymbol{y}} \quad \boldsymbol{y}^{-}\boldsymbol{A}\boldsymbol{x}\boldsymbol{x}^{-}\boldsymbol{A}\boldsymbol{y}. \tag{6}$$

Следствие 2. *Пусть $\boldsymbol{A}$ — матрица порядка n со спектральным радиусом $\lambda > \mathbb{0}$. Тогда минимум в задаче* (6) *равен λ^{2}, а все регулярные решения имеют вид*

$$\begin{aligned}
\boldsymbol{x} &= \alpha((\lambda^{-2}\boldsymbol{A}^{2})^{*}\boldsymbol{v} \oplus \lambda^{-1}\boldsymbol{A}(\lambda^{-2}\boldsymbol{A}^{2})^{*}\boldsymbol{w}),\\
\boldsymbol{y} &= \beta(\lambda^{-1}\boldsymbol{A}(\lambda^{-2}\boldsymbol{A}^{2})^{*}\boldsymbol{v} \oplus (\lambda^{-2}\boldsymbol{A}^{2})^{*}\boldsymbol{w}),
\end{aligned}$$

при любых $\alpha > \mathbb{0}$, $\beta > \mathbb{0}$ и $\boldsymbol{v}, \boldsymbol{w} \in \mathbb{X}^{n}$.

Одноранговая аппроксимация положительных матриц

Пусть имеется $(n \times n)$-матрица $\boldsymbol{A} = (a_{ij})$ с элементами $a_{ij} > 0$. Рассмотрим задачу аппроксимации матрицы $\boldsymbol{A}$ матрицей единичного ранга $\boldsymbol{y}\boldsymbol{x}^{-}$, где $\boldsymbol{x}^{-} = (x_1^{-1}, \ldots, x_n^{-1})$ и $\boldsymbol{y} = (y_1, \ldots, y_n)^{\mathrm{T}}$ — положительные n-векторы.

В работе в качестве функции расстояния берется максимальный разброс ошибки аппроксимации $\log a_{ij} - \log(y_i/x_j)$ по всем элементам матрицы $\boldsymbol{A}$ в логарифмической шкале.

Преобразуем величину максимального разброса, используя свойство монотонности логарифма по основанию больше единицы, а также тождество $\min(a, b) = -\max(-a, -b)$

$$\max_{1\leq i,j\leq n}(\log a_{ij} - \log(y_i/x_j)) - \min_{1\leq i,j\leq n}(\log a_{ij} - \log(y_i/x_j)) =$$
$$= \log\left(\max_{1\leq i,j\leq n} a_{ij}x_j/y_i \times \max_{1\leq k,l\leq n} a_{kl}^{-1}y_k/x_l\right).$$

В силу монотонности, минимизацию полученного логарифма можно заменить на минимизацию его аргумента. Тогда рассматриваемая задача аппроксимации принимает вид

$$\min_{x_j,y_i} \max_{1\leq i,j\leq n} a_{ij}x_j/y_i \times \max_{1\leq k,l\leq n} a_{kl}^{-1}y_k/x_l.$$

В терминах идемпотентного полуполя $\mathbb{R}_{\max,\times}$ задача аппроксимации записывается в виде

$$\min_{\boldsymbol{x},\boldsymbol{y}} \quad \boldsymbol{y}^{-}\boldsymbol{A}\boldsymbol{x}\boldsymbol{x}^{-}\boldsymbol{A}^{-}\boldsymbol{y}.$$

Полученная задача имеет вид задачи (5), решение которой дает следствие (1).

Предположим, что матрица $\boldsymbol{A}$ — обратно симметрическая. Учитывая, что тогда $\boldsymbol{A}^{-} = \boldsymbol{A}$, задача аппроксимации принимает вид

$$\min_{\boldsymbol{x},\boldsymbol{y}} \quad \boldsymbol{y}^{-}\boldsymbol{A}\boldsymbol{x}\boldsymbol{x}^{-}\boldsymbol{A}\boldsymbol{y},$$

а ее решение дает следствие (2).

Заключение

В работе изучена многомерная задача тропической оптимизации, в которой целевая функция задана при помощи двух матриц. Построено полное решение задачи тропической оптимизации для произвольных матриц в явном виде в замкнутой форме и рассмотрено два частных случая задачи. Полученные результаты применены для решения задач одноранговой аппроксимации положительных матриц.

Литература

[1] Baccelli F., Cohen G., Olsder G. J., Quadrat J.-P. Synchronization and Linearity: An Algebra for Discrete Event Systems. Wiley Series in Probability and Statistics. Chichester: Wiley, 1993. 514 p.

[2] Golan J. S. Semirings and Affine Equations over Them: Theory and Applications. Dordrecht: Springer, 2003. Vol. 556 of Mathematics and Its Applications. 256 p. DOI:10.1007/978-94-017-0383-3

[3] Маслов В. П., Колокольцов В. Н. Идемпотентный анализ и его применение в оптимальном управлении. М.: Физматлит, 1994. 144 с.

[4] Кривулин Н. К. Методы идемпотентной алгебры в задачах моделирования и анализа сложных систем. СПб.: Изд-во С.-Петерб. ун-та, 2009. 256 с.

[5] Krivulin N. Tropical optimization problems in project scheduling // Proc. 7th Multidisciplinary Intern. Conf. on Scheduling: Theory and Applications / eds Z. Hanzálek, G. Kendall, B. McCollum, P. Šůcha. MISTA, 2015. P. 492–506.

[6] Кривулин Н. К., Плотников П. В. Прямое решение минимаксной задачи размещения в прямоугольной области на плоскости с прямоугольной метрикой // Вестн. С.-Петербург. ун-та. 2018. Вып. 2, Т.14. С. 116-130

[7] Krivulin N. Extremal properties of tropical eigenvalues and solutions to tropical optimization problems // Linear Algebra Appl. 2015. Vol. 468. P. 211–232. DOI: 10.1016/j.laa.2014.06.044

[8] Krivulin N. A multidimensional tropical optimization problem with nonlinear objective function and linear constraints // Optimization. 2015. Vol. 64, N 5. P. 1107–1129. DOI: 10.1080/02331934.2013.840624

[9] Krivulin N. Tropical optimization problems // Advances in Economics and Optimization: Collected Scientific Studies Dedicated to the Memory of L. V. Kantorovich / eds L. A. Petrosyan, J. V. Romanovsky, D. W. kay Yeung. Economic Issues, Problems and Perspectives. New York: Nova Sci. Publ., 2014. P. 195-214