Использование тропической оптимизации для одноранговой аппроксимации положительных матриц с пропусками[1]

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

Предлагается полное прямое решение задачи аппроксимации положительных матриц с пропусками при помощи матриц единичного ранга. Задача аппроксимации формулируется как задача минимизации log-чебышевского расстояния между матрицами и решается при помощи результатов из области тропической математики. Приводится решение задачи аппроксимации для произвольной положительной матрицы с пропусками и для матрицы, у которой нет полностью пропущенных столбцов (строк).

Введение

Задача малоранговой аппроксимации возникает во многих областях прикладной математики [1]. Для того, чтобы перейти от исходных данных к данным, содержащим только основную информацию, часто используют аппроксимацию матрицами единичного ранга. При помощи одноранговой аппроксимации могут решаться, например, некоторые задачи восстановления изображений [2], статистики [3] и теории графов [4]. Иногда в таких задачах приходится иметь дело с матрицами, некоторые элементы которых неопределены (пропущены). Тогда решение задачи аппроксимации будет также включать заполнение пропусков в данных [5, 6].

В работах [7, 8, 9] задача аппроксимации положительной квадратной матрицы при помощи матрицы единичного ранга формулируется как задача минимизации log-чебышевского расстояния между матрицами. Задача аппроксимации приводится к задаче, записанной в терминах тропической (идемпотентной) математики, — области математики,

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

которая изучает теорию и приложения идемпотентных полуколец и полуполей. Подход к решению, в основе которого лежат методы и результаты тропической оптимизации, используется для получения полного решения в компактной векторной форме.

В настоящей статье решение задачи одноранговой аппроксимации, которое было получено в [9], обобщается на случай положительных прямоугольных матриц, некоторые значения которых могут быть неопределены (матрицы с пропусками). Задача log-чебышевской аппроксимации заменяется эквивалентной задачей тропической оптимизации, записанной в терминах max-алгебры — идемпотентного полуполя, которое в качестве сложения использует операцию взятия максимума, а в качестве умножения — обычное арифметическое умножение. Приводятся необходимые определения и результаты идемпотентной математики, которые затем используются для решения задачи тропической оптимизации при различных предположениях. При помощи полученных результатов строится решение рассматриваемой задачи аппроксимации в компактной векторной форме. Представлены решения задачи аппроксимации для произвольной положительной матрицы с пропусками и для положительной матрицы без полностью неопределенных столбцов (строк).

Элементы тропической математики

Приведем основные определения и результаты тропической математики [7, 10], которые будут применяться в дальнейшем. Для более подробного изучения моделей и методов тропической математики могут быть использованы работы [11, 12].

Идемпотентное полуполе

Пусть $\mathbb{X}$ — непустое числовое множество, замкнутое относительно операций сложения $\oplus$ и умножения $\otimes$. Сложение и умножение ассоциативны и коммутативны, и имеют в качестве нейтральных элементов нуль $\mathbb{0}$ и единицу $\mathbb{1}$ соответственно. Сложение является идемпотентным, то есть удовлетворяет условию $x \oplus x = x$ для любого элемента $x \in \mathbb{X}$. Умножение дистрибутивно относительно сложения и обратимо: для любого ненулевого элемента $x \in \mathbb{X}$ существует обратный по умножению элемент $x^{-1} \in \mathbb{X}$ такой, что $x \otimes x^{-1} = \mathbb{1}$. Далее знак умножения $\otimes$ для краткости опускается. Множество $\mathbb{X}$ вместе с определенными на нем операциями $\oplus$, $\otimes$ и их нейтральными элементами, образует алгебраическую систему, которую называют идемпотентным полуполем.

Возведение элемента $x \in \mathbb{X}$ в целую степень $p > 0$ определяется обычным образом. Дополнительно будем считать, что уравнение $x^p = a$ разрешимо для любого натурального числа p и для любого элемента $a \in \mathbb{X}$, что обеспечивает существование рациональных степеней.

Полуполе $\mathbb{R}_{\max,\times}$ определено на множестве неотрицательных вещественных чисел и в качестве сложения $\oplus$ имеет операцию взятия максимума, а в качестве умножения $\otimes$ использует обычное арифметическое умножение. Нейтральные элементы $\mathbb{0}$ и $\mathbb{1}$ совпадают с арифметическими 0 и 1. Понятия степени и обратного элемента имеют обычный смысл. Такое полуполе часто называют max-алгеброй.

Матрицы и векторы

Обозначим через $\mathbb{X}^{m\times n}$ множество матриц с элементами из $\mathbb{X}$, которые имеют m строк и n столбцов. Матрица, состоящая только из нулевых элементов, называется нулевой и обозначается $\mathbf{0}$. Матрица, диагональные элементы которой равны единице $\mathbb{1}$, а недиагональные — нулю $\mathbb{0}$, называется единичной и обозначается $\boldsymbol{I}$. В случае max-алгебры нулевая и единичная матрица имеют стандартный вид. Матрица без нулевых столбцов (строк) называется регулярной по столбцам (строкам).

Матричное сложение и умножение, а также умножение матрицы на число определяются по стандартным правилам с заменой арифметических операций на операции $\oplus$ и $\otimes$.

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

Рассмотрим квадратную матрицу $\boldsymbol{A} = (a_{ij}) \in \mathbb{X}^{n\times n}$. След матрицы $\boldsymbol{A}$ вычисляется по формуле $\operatorname{tr}\boldsymbol{A} = a_{11} \oplus \cdots \oplus a_{nn}$.

Спектральный радиус матрицы $\boldsymbol{A}$ определяется выражением $\lambda = \operatorname{tr}\boldsymbol{A} \oplus \cdots \oplus \operatorname{tr}^{1/n}(\boldsymbol{A}^n)$.

Матрица Клини для матрицы $\boldsymbol{A}$ имеет вид $\boldsymbol{A}^{*} = \boldsymbol{I} \oplus \cdots \oplus \boldsymbol{A}^{n-1}$.

Обозначим множество векторов-столбцов размера n через $\mathbb{X}^n$. Вектор, который не имеет нулевых компонент, называется регулярным. В max-алгебре регулярность вектора означает, что все его элементы положительны.

Мультипликативно сопряженным транспонированием ненулевого вектора $\boldsymbol{x} = (x_i)$ называется его преобразование в вектор-строку $\boldsymbol{x}^- = (x_i^-)$, где $x_i^- = x_i^{-1}$, если $x_i \neq \mathbb{0}$, и $x_i^- = \mathbb{0}$ — иначе.

Задача тропической оптимизации

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

$$\min_{\boldsymbol{x},\boldsymbol{y}} \quad \boldsymbol{x}^-\boldsymbol{A}\boldsymbol{y} \oplus \boldsymbol{y}^-\boldsymbol{A}^-\boldsymbol{x}. \tag{1}$$

Следующая теорема представляет собой обобщение результата, полученного для квадратных матриц в работе [9], на общий случай прямоугольных матриц.

Теорема 1. *Пусть $\boldsymbol{A} \in \mathbb{X}^{m\times n}$ — ненулевая матрица, μ — спектральный радиус матрицы $\boldsymbol{A}\boldsymbol{A}^-$. Тогда минимум в задаче* (1) *равен $\mu^{1/2}$, а все регулярные решения имеют вид*

$$\boldsymbol{x} = (\mu^{-1}\boldsymbol{A}\boldsymbol{A}^-)^*\boldsymbol{v} \oplus \mu^{-1/2}\boldsymbol{A}(\mu^{-1}\boldsymbol{A}^-\boldsymbol{A})^*\boldsymbol{w},$$
$$\boldsymbol{y} = \mu^{-1/2}\boldsymbol{A}^-(\mu^{-1}\boldsymbol{A}\boldsymbol{A}^-)^*\boldsymbol{v} \oplus (\mu^{-1}\boldsymbol{A}^-\boldsymbol{A})^*\boldsymbol{w}; \quad \boldsymbol{v} \in \mathbb{X}^m, \quad \boldsymbol{w} \in \mathbb{X}^n.$$

Для регулярной по столбцам матрицы решение записывается в следующей форме.

Теорема 2. *Пусть $\boldsymbol{A} \in \mathbb{X}^{m\times n}$ — регулярная по столбцам матрица, μ — спектральный радиус матрицы $\boldsymbol{A}\boldsymbol{A}^-$. Тогда минимум в задаче* (1) *равен $\mu^{1/2}$, а все регулярные решения имеют вид*

$$\boldsymbol{x} = (\mu^{-1}\boldsymbol{A}\boldsymbol{A}^-)^*\boldsymbol{u}, \quad \boldsymbol{u} \in \mathbb{X}^m;$$
$$\mu^{-1/2}\boldsymbol{A}^-\boldsymbol{x} \leq \boldsymbol{y} \leq \mu^{1/2}(\boldsymbol{x}^-\boldsymbol{A})^-.$$

Для матрицы, регулярной по строкам, может быть получен аналогичный результат.

Приложение к задаче аппроксимации

Предположим, что задана положительная матрица $\boldsymbol{A}$ размера $m\times n$, некоторые элементы которой пропущены (не определены). Пропущенные элементы положительной матрицы доопределим нулями. Задача приближения матрицы $\boldsymbol{A}$ формулируется как задача минимизации ошибки аппроксимации, которая вычисляется по всем известным (ненулевым) элементам матрицы. В качестве функции ошибки аппроксимации используется функция расстояния Чебышева в логарифмической шкале, где логарифм берется по основанию больше единицы.

Учитывая монотонность логарифма, задача log-чебышевской аппроксимации положительной матрицы $\boldsymbol{A} = (a_{ij})$ при помощи положительной матрицы $\boldsymbol{X} = (x_{ij})$ формулируется как задача минимизации функции

$$\max_{i,j:a_{ij}\neq 0} |\log a_{ij} - \log x_{ij}| = \log \max_{i,j:a_{ij}\neq 0} \max(a_{ij}x_{ij}^{-1}, a_{ij}^{-1}x_{ij}).$$

Так как минимумы логарифма и его аргумента в правой части этого равенства достигаются одновременно, рассматриваемая задача аппроксимации эквивалентна задаче минимизации аргумента логарифма. Далее заметим, что любая положительная матрица $\boldsymbol{X}$ может быть представлена в виде произведения $\boldsymbol{X} = \boldsymbol{s}\boldsymbol{t}^T$, где $\boldsymbol{s} = (s_i)$ и $\boldsymbol{t} = (t_j)$ — положительные векторы размеров m и n соответственно. Следовательно, задача аппроксимации принимает вид

$$\min_{\boldsymbol{s},\boldsymbol{t}} \quad \max_{i,j:a_{ij}\neq 0} \max(s_i^{-1}a_{ij}t_j^{-1}, s_i a_{ij}^{-1}t_j).$$

Записывая целевую функцию этой задачи в терминах max-алгебры, получим равенство

$$\bigoplus_{i,j:a_{ij}\neq 0} (s_i^{-1}a_{ij}t_j^{-1} \oplus s_i a_{ij}^{-1}t_j) = \boldsymbol{s}^-\boldsymbol{A}(\boldsymbol{t}^-)^T \oplus \boldsymbol{t}^T\boldsymbol{A}^-\boldsymbol{s},$$

которое выполняется в силу того, что при сопряженном транспонировании нулевые элементы матрицы $\boldsymbol{A}$ переходят в нулевые элементы матрицы $\boldsymbol{A}^-$, что обеспечивает равенство обоих частей.

Таким образом, задача одноранговой аппроксимации сводится к задаче нахождения матрицы $\boldsymbol{X} = \boldsymbol{s}\boldsymbol{t}^T$, где $\boldsymbol{s}$ и $\boldsymbol{t}$ — положительные векторы, решающие задачу

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

После замены $\boldsymbol{x} = \boldsymbol{s}$, $\boldsymbol{y} = (\boldsymbol{t}^T)^-$ последняя задача принимает форму (1) и для получения полного решения задачи аппроксимации остается применить результаты предыдущего раздела.

Использование теоремы 1 дает следующее решение задачи одноранговой аппроксимации.

Теорема 3. *Пусть* $\boldsymbol{A}$ *— положительная матрица с пропусками, которые доопределены нулями,* μ *— спектральный радиус матрицы* $\boldsymbol{A}\boldsymbol{A}^{-}$*. Тогда минимальная погрешность* log*-чебышевской аппроксимации составляет* $\log \mu^{1/2}$*, а все аппроксимирующие матрицы имеют вид* $\boldsymbol{s}\boldsymbol{t}^T$*, где*

$$\boldsymbol{s} = (\mu^{-1}\boldsymbol{A}\boldsymbol{A}^{-})^{*}\boldsymbol{v} \oplus \mu^{-1/2}\boldsymbol{A}(\mu^{-1}\boldsymbol{A}^{-}\boldsymbol{A})^{*}\boldsymbol{w},$$
$$\boldsymbol{t}^T = (\mu^{-1/2}\boldsymbol{A}^{-}(\mu^{-1}\boldsymbol{A}\boldsymbol{A}^{-})^{*}\boldsymbol{v} \oplus (\mu^{-1}\boldsymbol{A}^{-}\boldsymbol{A})^{*}\boldsymbol{w})^{-}; \quad \boldsymbol{v} \in \mathbb{X}^m, \quad \boldsymbol{w} \in \mathbb{X}^n.$$

Следующий результат опирается на теорему 2 и справедлив для матриц, у которых нет полностью пропущенных столбцов.

Теорема 4. *Пусть* $\boldsymbol{A}$ *— положительная матрица с пропусками без полностью неопределенных столбцов, пропуски в которой доопределены нулями,* μ *— спектральный радиус матрицы* $\boldsymbol{A}\boldsymbol{A}^{-}$*. Тогда минимальная погрешность* log*-чебышевской аппроксимации составляет* $\log \mu^{1/2}$*, а все аппроксимирующие матрицы имеют вид* $\boldsymbol{s}\boldsymbol{t}^T$*, где*

$$\boldsymbol{s} = (\mu^{-1}\boldsymbol{A}\boldsymbol{A}^{-})^{*}\boldsymbol{u}, \quad \boldsymbol{u} \in \mathbb{X}^m;$$
$$\mu^{-1/2}\boldsymbol{s}^{-}\boldsymbol{A} \leq \boldsymbol{t}^T \leq \mu^{1/2}(\boldsymbol{A}^{-}\boldsymbol{s})^{-}.$$

Для матриц без полностью пропущенных строк может быть получен аналогичный результат.

Литература

[1] Kumar N. K., Schneider J. Literature survey on low rank approximation of matrices // Linear Multilinear Algebra. 2017. Vol. 65, N 11. P. 2212–2244. DOI:10.1080/03081087.2016.1267104

[2] Gong K., Zhou J., Tohme M., Judenhofer M., Yang Y., Qi J. Sinogram blurring matrix estimation from point sources measurements with rank-one approximation // IEEE Trans. Medical Imaging. 2017. Vol. 36, N 10. P. 2179–2188. DOI:10.1109/TMI.2017.2711479

[3] Aissa-El-Bey A., Seghouane K. Sparse canonical correlation analysis based on rank-1 matrix approximation and its application for FMRI signals // 2016 IEEE Intern. Conf. on Acoustics, Speech and Signal Processing, Shanghai, China. IEEE, 2016. P. 4678–4682. DOI:10.1109/ICASSP.2016.7472564

[4] Belachew M.T., Gillis N. Solving the maximum clique problem with symmetric rank-one nonnegative matrix approximation // J. Optim. Theory Appl. 2017. Vol. 173, N 1. P. 279–296. DOI:10.1007/s10957-016-1043-6

[5] Shi Q., Lu H., Cheung Y. M. Rank-one matrix completion with automatic rank estimation via L1-norm regularization // IEEE Trans. Neural Netw. Learn. Syst. 2018. Vol. 29, N 10. P. 4744–4757. DOI:10.1109/TNNLS.2017.2766160

[6] Markovsky I., Usevich K. Structured low-rank approximation with missing data // SIAM J. Matrix Anal. Appl. 2013. Vol. 34, N 2. P. 814–830. DOI:10.1137/120883050

[7] Кривулин Н. К. Методы идемпотентной алгебры в задачах моделирования и анализа сложных систем. СПб: С.-Петерб. ун-т, 2009.

[8] Krivulin N. Using tropical optimization techniques to evaluate alternatives via pairwise comparisons // 2016 Proc. 7th SIAM Workshop on Combinatorial Scientific Computing / Eds A. H. Gebremedhin, E. G. Boman, B. Ucar. Philadelphia: SIAM, 2016. P. 62–72. DOI:10.1137/1.9781611974690.ch7

[9] Krivulin N. K., Romanova E. Yu. Rank-one approximation of positive matrices based on methods of tropical mathematics // Vestnik St. Petersburg Univ. Math. 2018. Vol. 51, N 2. P. 133–143. DOI:10.3103/S106345411802005X

[10] 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

[11] Butkovič P. Max-linear systems. London: Springer, 2010. (Springer Monographs in Mathematics) DOI:10.1007/978-1-84996-299-5

[12] McEneaney W. M. Max-plus methods for nonlinear control and estimation. Boston: Birkhäuser, 2006. (Systems and Control: Foundations and Applications) DOI:10.1007/0-8176-4453-9