Using tropical optimization in rank-one approximation of non-negative matrices

N. Krivulin and E. Romanova

Introduction

Low rank approximation of matrices finds wide use in many areas [1], such as machine learning, statistics, and data compression. In many applications, approximation by matrices of unit rank is of interest to deal only with the basic information involved in the data under consideration [2, 3].

In [4, 5], the problem of rank-one approximation of positive square matrices is formulated as a problem of minimizing the log-Chebychev distance between matrices. The optimization problem is represented in terms of tropical (idempotent) mathematics, which deals with the theory and applications of idempotent semifields. A solution approach based on methods and results of tropical optimization is used to provide a complete direct solution given in compact vector form.

In this paper, we extend the above results to solve the rank-one approximation problem in the case of rectangular non-negative matrices. The problem is formulated in terms of max-algebra, which is a tropical semifield with the maximum in the role of addition, and with multiplication defined as usual. We start with necessary definitions and results of tropical mathematics, and then apply them to obtain complete solutions to a tropical optimization problem under different assumptions. The results obtained serve as the basis to derive a solution to the approximation problem in question in compact closed vector form. We offer the solution in different forms for the case of arbitrary non-negative matrices and for the case of matrices without zero columns.

This work was supported in part by the Russian Foundation for Basic Research, Grant No. 18-010-00723.

1. Tropical algebra

We begin with preliminary definitions and results of tropical mathematics from [4, 6], which are used in what follows.

1.1. Idempotent semifield

Suppose $\mathbb{X}$ is a nonempty set that is closed under addition $\oplus$ and multiplication $\otimes$. Addition and multiplication are associative and commutative, and have respective neutral elements, zero $\mathbb{0}$ and identity $\mathbb{1}$. Addition is idempotent, resulting in the equality $x \oplus x = x$ for all $x \in \mathbb{X}$. Multiplication distributes over addition, and is invertible in the sense that, for each nonzero element $x \in \mathbb{X}$, there exists inverse element x^{-1} such that $x \otimes x^{-1} = \mathbb{1}$. Together with the operations $\oplus$ and $\otimes$, and their neutral elements, the set $\mathbb{X}$ forms the algebraic system, which is usually called the idempotent semifield. In what follows, the multiplication sign $\otimes$ is dropped for simplicity.

The semifield $\mathbb{R}_{\max,\times}$ is defined on the set of non-negative real numbers, and equipped with the addition $\oplus$ defined as maximum, and the multiplication $\otimes$ defined as usual. The neutral elements $\mathbb{0}$ and $\mathbb{1}$ coincide with the arithmetic zero 0 and one 1. The power and inversion notations have the usual meaning. This semifield is often called max-algebra.

1.2. Matrix algebra

Let $\mathbb{X}^{m\times n}$ be the set of matrices over $\mathbb{X}$, with m rows and n columns. A matrix with all zero elements is the zero matrix denoted $\mathbf{0}$. A matrix with $\mathbb{1}$ on the diagonal and $\mathbb{0}$ elsewhere is identity matrix, which is denoted by $\boldsymbol{I}$. In the case of max-algebra, the zero and identity matrices have the usual form. Any matrix without zero columns is called column-regular.

Matrix addition and multiplication, and multiplication by scalars are defined as usual, except that the arithmetic operations are replaced by $\oplus$ and $\otimes$.

The multiplicative conjugate transpose of a nonzero matrix $\boldsymbol{A} = (a_{ij}) \in \mathbb{X}^{m\times n}$ is the matrix $\boldsymbol{A}^{-} = (a_{ij}^{-}) \in \mathbb{X}^{n\times m}$ with the elements $a_{ij}^{-} = a_{ji}^{-1}$ if $a_{ji} \neq \mathbb{0}$, and $a_{ij}^{-} = \mathbb{0}$ otherwise.

Consider a square matrix $\boldsymbol{A} = (a_{ij}) \in \mathbb{X}^{n\times n}$. The trace of the matrix $\boldsymbol{A}$ is calculated as $\operatorname{tr}\boldsymbol{A} = a_{11} \oplus \cdots \oplus a_{nn}$.

The spectral radius of $\boldsymbol{A}$ is the scalar $\lambda = \operatorname{tr}\boldsymbol{A} \oplus \cdots \oplus \operatorname{tr}^{1/n}(\boldsymbol{A}^{n})$.

The Kleene star operator for $\boldsymbol{A}$ is given by the formula $\boldsymbol{A}^{*} = \boldsymbol{I} \oplus \cdots \oplus \boldsymbol{A}^{n-1}$.

The set of column vectors of order n is denoted by $\mathbb{X}^{n}$. A vector that has no zero elements is called regular. In max-algebra, the regularity of a vector means that the vector is positive.

The multiplicative conjugate transpose of a nonzero column vector $\boldsymbol{x} = (x_i)$ is a row vector $\boldsymbol{x}^{-} = (x_i^{-})$, where $x_i^{-} = x_i^{-1}$ if $x_i \neq \mathbb{0}$, and $x_i^{-} = \mathbb{0}$ otherwise.

2. Tropical optimization problem

Given a rectangular matrix $\boldsymbol{A} \in \mathbb{X}^{m\times n}$, the problem is to find regular vectors $\boldsymbol{x} \in \mathbb{X}^{m}$ and $\boldsymbol{y} \in \mathbb{X}^{n}$ that achieve the minimum

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

The following theorem generalizes the result of the paper [5], and gives a complete solution to problem (1) in explicit form.

Theorem 1. *Let $\boldsymbol{A} \in \mathbb{X}^{m\times n}$ be a nonzero matrix, μ be the spectral radius of the matrix $\boldsymbol{A}\boldsymbol{A}^{-}$. Then, the minimum in problem* (1) *is equal to $\mu^{1/2}$, and all regular solutions are given by*

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

For column-regular matrices, the solution can be represented as follows.

Theorem 2. *Let $\boldsymbol{A} \in \mathbb{X}^{m\times n}$ be a column-regular matrix, μ be the spectral radius of $\boldsymbol{A}\boldsymbol{A}^{-}$. Then, the minimum in problem* (1) *is equal to $\mu^{1/2}$, and all regular solutions are given by*

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

3. Application to matrix approximation

The problem of approximating a non-negative rectangular matrix $\boldsymbol{A} = (a_{ij})$ by positive matrix $\boldsymbol{X} = (x_{ij})$ is formulated to minimize the Chebyshev distance in logarithmic scale, given by

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

Since the logarithm (on the base greater than one) is monotone increasing, the approximation problem is equivalent to minimizing the argument of the logarithm. Observing that any positive matrix $\boldsymbol{X}$ of unit rank can be represented as $\boldsymbol{s}\boldsymbol{t}^{T}$, where $\boldsymbol{s} = (s_i)$ and $\boldsymbol{t} = (t_j)$ are positive vectors, we reduce the problem to that of the form

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

Representation of the objective function in terms of max-algebra yields

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

We now formulate the rank-one approximation problem as to find a matrix $\boldsymbol{X} = \boldsymbol{s}\boldsymbol{t}^{T}$, where $\boldsymbol{s}$ and $\boldsymbol{t}$ are positive vectors that solve the problem

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

The last problem has the same form as problem (1) with $\boldsymbol{x} = \boldsymbol{s}$, $\boldsymbol{y} = (\boldsymbol{t}^T)^-$ and thus admits complete solutions given by the results of section 2.

Theorem 3. *Let $\boldsymbol{A}$ be a non-negative matrix, μ be the spectral radius of the matrix $\boldsymbol{A}\boldsymbol{A}^-$. Then, the minimum error of* log-*Chebyshev approximation is equal to* $\log(\mu)/2$, *and all approximate matrices are given by $\boldsymbol{s}\boldsymbol{t}^T$, where*

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

The next result holds if the approximated matrix has no zero columns.

Theorem 4. *Let $\boldsymbol{A}$ be a non-negative matrix without zero columns, μ be the spectral radius of $\boldsymbol{A}\boldsymbol{A}^-$. Then, the minimal error of* log-*Chebyshev approximation is equal to* $\log \mu^{1/2}$, *and all approximate matrices are given by $\boldsymbol{s}\boldsymbol{t}^T$, where*

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

References

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

[2] A. Aissa-El-Bey, K. Seghouane *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, China. IEEE, 2016. P. 4678–4682.

[3] T. Saaty *The Analytic Hierarchy Process: Planning, Priority Setting, Resource Allocation*, New York, McGraw-Hill, 1980. 281 p.

[4] N.K. Krivulin *Methods of idempotent algebra for problems in modeling and analysis of complex systems*, St. Petersburg, St. Petersburg Univ. Publ., 2009. 256 p. (in Russian)

[5] N.K. Krivulin, E.Yu. Romanova *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.

[6] N. Krivulin *Extremal properties of tropical eigenvalues and solutions to tropical optimization problems* // Linear Algebra Appl. 2015. Vol. 468. P. 211–232.

N. Krivulin
Faculty of Mathematics and Mechanics
Saint Petersburg State University
Saint Petersburg, Russia
e-mail: `nkk@math.spbu.ru`

E. Romanova
Faculty of Mathematics and Mechanics
Saint Petersburg State University
Saint Petersburg, Russia
e-mail: `romanova.ej@gmail.com`