Constrained rank-one matrix factorization using tropical mathematics

N. Krivulin and E. Romanova

Introduction

Constrained matrix factorization finds wide use in various areas of data analysis [1], such as recommendation systems and image processing. In many applications it is essential to use the factorization by matrices of unit rank (rank-one factorization) [2, 3], which involves the approximation of matrices by products of column and row vectors.

We consider a problem of constrained factorization that can be formulated as the minimization problem

$$\begin{aligned} \min_{\boldsymbol{x},\boldsymbol{y}} \quad & \mathrm{d}(\boldsymbol{A}, \boldsymbol{x}\boldsymbol{y}^{-}), \\ \text{s. t.} \quad & \boldsymbol{a} \leq \boldsymbol{x} \leq \boldsymbol{b},\ \boldsymbol{c} \leq \boldsymbol{y} \leq \boldsymbol{d}, \end{aligned} \tag{1}$$

where d is an approximation error, $\boldsymbol{A}$ is a given matrix under factorization, $\boldsymbol{x}$ is a column vector, $\boldsymbol{y}^{-}$ is the row vector obtained from a column vector $\boldsymbol{y}$ by transposing and replacing each nonzero element by its inverse, and $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$, $\boldsymbol{d}$ are given column vectors.

In this paper we assume the matrix $\boldsymbol{A}$ to be positive and take the Chebyshev distance in the logarithmic scale as an error function d. To find all solutions of the minimization problem we apply methods of tropical (idempotent) mathematics, which deals with the theory and applications of idempotent semifields. We extend the solutions obtained under different assumptions in [4, 5] for an unconstrained rank-one factorization problem to constrained problem (1), where the matrix $\boldsymbol{A}$ can contain missing elements. We start with necessary definitions and notations of tropical mathematics. Next, we formulate and solve a constrained tropical optimization problem under different assumptions. Finally, the obtained solution is applied to the constrained factorization problem in question.

This work was supported in part by the Russian Foundation for Basic Research, Grant No. 20-010-00145.

1. Tropical algebra

We begin with a short overview of basic definitions, notations and preliminary results of tropical mathematics from [6, 7], which are used for the formulation and solution of a tropical optimization problem in the next section. Further details on tropical mathematics can be found, for example, in [8, 9, 10].

Consider a nonempty set $\mathbb{X}$ equipped with addition $\oplus$ and multiplication $\otimes$, which are both associative and commutative, and have respective neutral elements, zero $\mathbb{0}$ and identity $\mathbb{1}$. The addition $\oplus$ is idempotent, which means that $x \oplus x = x$ for all $x \in \mathbb{X}$. The multiplication is distributive over the addition and is invertible, implying that each nonzero $x \in \mathbb{X}$ has its inverse 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 referred to as the idempotent semifield. In what follows we omit the multiplication sign for the sake of brevity.

The addition induces on $\mathbb{X}$ a partial order such that the relation $x \leq y$ holds if and only if $x \oplus y = y$. The partial order is considered as extendable to a total order, and so we assume the semifield to be linearly ordered. Further the relation symbols and optimization problems are considered in the sense of this order.

Below, we use a real idempotent semifield, which is commonly called max-algebra. This semifield is defined on the set of non-negative real numbers and has maximum in the role of addition, and usual arithmetic multiplication in the role of multiplication. Neutral elements coincide with the usual arithmetic 0 and 1. The relation $\leq$ agrees with the natural linear order on the set of non-negative real numbers. The concepts of the inverse and the power are conventional.

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 $\mathbf{0}$. A square matrix with $\mathbb{1}$ on the diagonal and $\mathbb{0}$ elsewhere is identity matrix $\boldsymbol{I}$. In the case of max-algebra, zero and identity matrices have the usual form. Any matrix without zero columns (rows) is called column (row)-regular.

Matrix addition and multiplication and multiplication by scalars follow the standard entry-wise formulas with the arithmetic operations replaced by $\oplus$ and $\otimes$.

The multiplicative conjugate transposition of a nonzero matrix $\boldsymbol{A} = (a_{ij}) \in \mathbb{X}^{m\times n}$ yields 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 a matrix $\boldsymbol{A}$ is the scalar $\lambda = \operatorname{tr} \boldsymbol{A} \oplus \cdots \oplus \operatorname{tr}^{1/n}(\boldsymbol{A}^{n})$.

For a square matrix $\boldsymbol{A}$ we define the matrix $\boldsymbol{A}^{*} = \boldsymbol{I} \oplus \cdots \oplus \boldsymbol{A}^{n-1}$.

The set of column vectors of order n is denoted by $\mathbb{X}^{n}$. Any vector without zero elements is called regular. In max-algebra, the regularity of a vector means that the vector is positive.

The multiplicative conjugate transposition of a nonzero column vector $\boldsymbol{x} = (x_i)$ yields the 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

We consider a problem of tropical optimization with constraints, which is used below in factorization of matrices. Given a matrix $\boldsymbol{A} \in \mathbb{X}^{m\times n}$ and vectors $\boldsymbol{a}$, $\boldsymbol{b} \in \mathbb{X}^{m}$, $\boldsymbol{c}$, $\boldsymbol{d} \in \mathbb{X}^{n}$, the problem is to find all regular vectors $\boldsymbol{x} \in \mathbb{X}^{m}$ and $\boldsymbol{y} \in \mathbb{X}^{n}$ that achieve the minimum

$$\begin{aligned} \min_{\boldsymbol{x},\boldsymbol{y}} \quad & \boldsymbol{x}^{-}\boldsymbol{A}\boldsymbol{y} \oplus \boldsymbol{y}^{-}\boldsymbol{A}^{-}\boldsymbol{x}, \\ \text{s. t.} \quad & \boldsymbol{a} \leq \boldsymbol{x} \leq \boldsymbol{b},\ \boldsymbol{c} \leq \boldsymbol{y} \leq \boldsymbol{d}. \end{aligned} \tag{2}$$

The following result gives a complete solution in an explicit vector form to problem (2) for any nonzero matrix.

Theorem 1. *Suppose that $\boldsymbol{A}$ is a nonzero matrix and μ is a spectral radius of the matrix $\boldsymbol{A}\boldsymbol{A}^{-}$. Let $\boldsymbol{a}$ and $\boldsymbol{c}$ be vectors, $\boldsymbol{b}$ and $\boldsymbol{d}$ be regular vectors such that $\boldsymbol{b}^{-}\boldsymbol{a} \leq \mathbb{1}$ and $\boldsymbol{d}^{-}\boldsymbol{c} \leq \mathbb{1}$. Denote $r = (m+n)/2$ and define a scalar*

$$\theta = \mu^{1/2} \oplus \bigoplus_{k=1}^{\lceil r\rceil} \left(\boldsymbol{b}^{-}\boldsymbol{A}(\boldsymbol{A}^{-}\boldsymbol{A})^{k-1}\boldsymbol{c} \oplus \boldsymbol{d}^{-}\boldsymbol{A}^{-}(\boldsymbol{A}\boldsymbol{A}^{-})^{k-1}\boldsymbol{a}\right)^{1/(2k-1)} \oplus \\ \oplus \bigoplus_{k=1}^{\lfloor r\rfloor} \left(\boldsymbol{b}^{-}(\boldsymbol{A}\boldsymbol{A}^{-})^{k}\boldsymbol{a} \oplus \boldsymbol{d}^{-}(\boldsymbol{A}^{-}\boldsymbol{A})^{k}\boldsymbol{c}\right)^{1/(2k)}. \tag{3}$$

Then, the minimum in problem (2) *is equal to θ and all regular solutions are given by*

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

where $\boldsymbol{v}$ and $\boldsymbol{w}$ are vectors that satisfy the conditions

$$\begin{aligned} \boldsymbol{a} \leq{}& \boldsymbol{v} \leq ((\boldsymbol{b}^{-} \oplus \theta^{-1}\boldsymbol{d}^{-}\boldsymbol{A}^{-})(\theta^{-2}\boldsymbol{A}\boldsymbol{A}^{-})^{*})^{-}, \\ \boldsymbol{c} \leq{}& \boldsymbol{w} \leq ((\theta^{-1}\boldsymbol{b}^{-}\boldsymbol{A} \oplus \boldsymbol{d}^{-})(\theta^{-2}\boldsymbol{A}^{-}\boldsymbol{A})^{*})^{-}. \end{aligned}$$

In the case of a column-regular matrix, a complete solution of problem (2) can be obtained in a different form as follows.

Theorem 2. *Suppose that $\boldsymbol{A}$ is a column-regular matrix and μ is a spectral radius of the matrix $\boldsymbol{A}\boldsymbol{A}^{-}$. Let $\boldsymbol{a}$ and $\boldsymbol{c}$ be vectors, $\boldsymbol{b}$ and $\boldsymbol{d}$ be regular vectors such that $\boldsymbol{b}^{-}\boldsymbol{a} \leq \mathbb{1}$ and $\boldsymbol{d}^{-}\boldsymbol{c} \leq \mathbb{1}$. Then, the minimum in problem* (2) *is equal to* (3) *and all regular solutions are given by*

$$\begin{aligned} \boldsymbol{x} = (\theta^{-2}\boldsymbol{A}\boldsymbol{A}^{-})^{*}\boldsymbol{u}, \quad & \boldsymbol{a} \oplus \theta^{-1}\boldsymbol{A}\boldsymbol{c} \leq \boldsymbol{u} \leq ((\boldsymbol{b}^{-} \oplus \theta^{-1}\boldsymbol{d}^{-}\boldsymbol{A}^{-})(\theta^{-2}\boldsymbol{A}\boldsymbol{A}^{-})^{*})^{-}, \\ & \theta^{-1}\boldsymbol{A}^{-}\boldsymbol{x} \oplus \boldsymbol{c} \leq \boldsymbol{y} \leq (\theta^{-1}\boldsymbol{x}^{-}\boldsymbol{A} \oplus \boldsymbol{d}^{-})^{-}. \end{aligned}$$

If the matrix $\boldsymbol{A}$ does not contain zero rows, a complete solution can be written in the following form.

Theorem 3. *Suppose that $\boldsymbol{A}$ is a row-regular matrix and μ is a spectral radius of the matrix $\boldsymbol{A}\boldsymbol{A}^{-}$. Let $\boldsymbol{a}$ and $\boldsymbol{c}$ be vectors, $\boldsymbol{b}$ and $\boldsymbol{d}$ be regular vectors such that $\boldsymbol{b}^{-}\boldsymbol{a} \leq \mathbb{1}$ and $\boldsymbol{d}^{-}\boldsymbol{c} \leq \mathbb{1}$. Then, the minimum in problem* (2) *is equal to* (3) *and all regular solutions are given by*

$$\boldsymbol{y} = (\theta^{-2}\boldsymbol{A}^{-}\boldsymbol{A})^{*}\boldsymbol{u}, \quad \boldsymbol{c} \oplus \theta^{-1}\boldsymbol{A}^{-}\boldsymbol{a} \leq \boldsymbol{u} \leq ((\boldsymbol{d}^{-} \oplus \theta^{-1}\boldsymbol{b}^{-}\boldsymbol{A})(\theta^{-2}\boldsymbol{A}^{-}\boldsymbol{A})^{*})^{-},$$
$$\theta^{-1}\boldsymbol{A}\boldsymbol{y} \oplus \boldsymbol{a} \leq \boldsymbol{x} \leq (\theta^{-1}\boldsymbol{y}^{-}\boldsymbol{A}^{-} \oplus \boldsymbol{b}^{-})^{-}.$$

3. Application to matrix factorization

Let $\boldsymbol{A} = (a_{ij})$ be a positive matrix with missing elements. First, we fill the missing elements of $\boldsymbol{A}$ with zeroes and consider problem (1), where $\boldsymbol{x} = (x_i)$ is a column vector and $\boldsymbol{y}^{-} = (y_i^{-1})$ is a row vector. We define the function d as the log-Chebyshev distance between the matrix $\boldsymbol{A}$ and the rank-one matrix $\boldsymbol{x}\boldsymbol{y}^{-}$. With the logarithm of a base greater than 1, which is monotone increasing, we have

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

Since logarithm is monotonic, the minimization of the logarithmic function is equivalent to minimizing the argument of this function. After eliminating the logarithm, we rewrite the objective function in terms of max-algebra to obtain

$$\bigoplus_{i,j:a_{ij}\neq 0} (x_i^{-1}a_{ij}y_j \oplus x_i a_{ij}^{-1}y_j^{-1}) = \boldsymbol{x}^{-}\boldsymbol{A}\boldsymbol{y} \oplus \boldsymbol{y}^{-}\boldsymbol{A}^{-}\boldsymbol{x}.$$

Thus, we can reduce the constrained problem of factorization of the positive matrix $\boldsymbol{A}$ to that of the form

$$\begin{aligned} \min_{\boldsymbol{x},\boldsymbol{y}} \quad & \boldsymbol{x}^{-}\boldsymbol{A}\boldsymbol{y} \oplus \boldsymbol{y}^{-}\boldsymbol{A}^{-}\boldsymbol{x}, \\ \text{s. t.} \quad & \boldsymbol{a} \leq \boldsymbol{x} \leq \boldsymbol{b},\ \boldsymbol{c} \leq \boldsymbol{y} \leq \boldsymbol{d}. \end{aligned}$$

As a result, we obtain the problem in the form of (2), which has complete solutions given by theorems 1, 2, and 3. In the application of the solutions to the factorization problem, the minimal error is calculated as the logarithm of (3).

References

[1] N. Gillis. *Introduction to nonnegative matrix factorization* // SIAG/OPT Views and News. 2017. Vol. 25, N 1. P. 7–16.

[2] X. Xiu, L. Kong. *Rank-one and sparse matrix decomposition for dynamic MRI* // Numerical Algebra, Control and Optimization. 2015. Vol. 5, N 2. P. 127–134.

[3] 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 (ICASSP), Shanghai, China. IEEE, 2016. P. 4678–4682.

[4] 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.

[5] N.K. Krivulin, E.Yu. Romanova. *On the rank-one approximation of positive matrices using tropical optimization methods* // Vestnik St. Petersburg Univ. Math. 2019. Vol. 52, N 2. P. 145–153.

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

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

[8] V.P. Maslov, V.N. Kolokoltsov. *Idempotent analysis and its applications to optimal control theory*, Nauka Publ., Moscow, 1994. (In Russian)

[9] J.S. Golan. *Semirings and affine equations over them* / Mathematics and Its Applications. 2003. Vol. 556. Kluwer Acad. Publ., Dordrecht.

[10] B. Heidergott, G.-J. Olsder., J. van der Woude. *Max plus at work* / Princeton Series in Applied Mathematics, Princeton, NJ, Princeton Univ. Press, 2006.

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