Solution of linear equations and inequalities in idempotent vector spaces

Nikolai Krivulin

***Abstract*—**Linear vector equations and inequalities are considered defined in terms of idempotent mathematics. To solve the equations, we apply an approach that is based on the analysis of distances between vectors in idempotent vector spaces. The approach reduces the solution of the equation to that of an optimization problem in the idempotent algebra setting. Based on the approach, existence and uniqueness conditions are established for the solution of equations, and a general solution to both linear equations and inequalities are given. Finally, a problem of simultaneous solution of equations and inequalities is also considered.

***Keywords*—**Idempotent vector space, Linear equation, Linear inequality, Tropical optimization problem, Existence condition

I. Introduction

Many applications of idempotent mathematics [1]–[9] involve the solution of linear equations and inequalities defined on finite-dimensional semimodules over idempotent semifields (idempotent vector spaces). One of the problems that often arise is to solve an equation having the form

$$A\boldsymbol{x} = \boldsymbol{d},$$

where A and $\boldsymbol{d}$ are given matrix and vector, $\boldsymbol{x}$ is an unknown vector, and multiplication is thought of in terms of idempotent algebra. Since the equation can be considered as representing linear dependence between vectors, the development of reasonable solution to the problem holds both practical and theoretical interest. Of particular importance are the methods that offer solutions in a compact vector form suitable for development of efficient computational algorithms and related software tools, including those intended for implementation in vector and parallel computers.

Along with the above equation, an inequality

$$A\boldsymbol{x} \leq \boldsymbol{d}$$

that is considered component-wise constitutes another problem of interest in the idempotent algebra setting.

Among the early investigations of the problem of solving the equation and of its interplay with linear dependence of vectors are the works [1]–[3]. Further development of the question is given in many studies, including [4], [9]–[14].

To solve the equation where the matrix has no (idempotent) zero entries, an approach based on the concept of a covering set of rows for the matrix A is proposed in [1], [2]. With the approach, existence conditions are established and a procedure to find all solutions of the equation is described in terms of the covering sets. The maximum solution to the equation is given in the form $\boldsymbol{x} = A^- \otimes \boldsymbol{d}$, where A^- is a pseudoinverse matrix in the initial idempotent semimodule (called there extremal inverse matrix), and $\otimes$ denotes matrix-vector multiplication in a dual semimodule. In [10], [11], the above approach is extended to investigate linear dependence in idempotent semimodules.

The development of the theory and methods in [3] is aimed in particular at the solution of equations when the matrix A may have zero entries. The operation of pseudoinversion is extended to such matrices (the matrix A^- is called conjugate to A). For the solution, existence conditions in the form of an equality $A(A^- \otimes \boldsymbol{d}) = \boldsymbol{d}$, where $\otimes$ is the multiplication in a dual semimodule, are given and uniqueness conditions are established. A procedure is proposed to determine the linear dependence between vectors. The results are further developed in [9] to offer a combinatorial and an algebraic techniques for the solution of equations.

In [4], [12], [13], a notion of a subsolution to the equation is introduced as any vector $\boldsymbol{x}$ that satisfies the condition $A\boldsymbol{x} \leq \boldsymbol{d}$. A residuation operation $\backslash$ is defined so that $A\backslash\boldsymbol{d}$ represents the maximal subsolution of the equation. It is shown that when an ordinary solution exists, it can be written in terms of a dual semimodule and then $A\backslash\boldsymbol{d} = A^- \otimes \boldsymbol{d}$. For an extended equation $A\boldsymbol{x} \oplus \boldsymbol{b} = \boldsymbol{d}$, where $\oplus$ denotes idempotent vector addition, a necessary and sufficient condition for the existence of its subsolutions is given in [4], [13] in the form of an inequality $\boldsymbol{b} \leq \boldsymbol{d}$ that however suggests only necessary conditions for the actual solution.

Another approach that is based on the application of an idempotent analogue for the matrix determinant, known as dominant, is proposed in [14]. A solution technique is developed which uses Cramer's rule with the dominant in place of determinant. The implementation of the approach requires, however, that some sufficient constraints for both the matrix A and the vector $\boldsymbol{d}$ to satisfy.

In this paper another solution approach is described which uses the analysis of distances between vectors in idempotent vector spaces. As a metric, we take a distance function that involves only main binary operations of the semimodule supplemented with the operation of pseudoinversion. This allows to represent subsequent results in a compact vector form only in terms of the initial semimodule and give them clear and natural geometrical interpretation in the plane with the Cartesian coordinates. The results presented are based

Manuscript received October 31, 2012.

N. Krivulin is with the Faculty of Mathematics and Mechanics, St. Petersburg State University, 28 Universitetsky Ave., St. Petersburg, 198504, Russia (e-mail: nkk < at > math.spbu.ru).

on implementation and further refinement of solutions that were first published in the papers [15]–[17] (see also [18]) and were not fully available in English.

We start with a brief overview of preliminary algebraic definitions and results. Furthermore, the problem of solving the equation under study reduces to an optimization problem of finding the minimal distance from a vector to a linear span of vectors. We derive a comprehensive solutions to the optimization problem under quite general conditions. The obtained results are applied to give existence and uniqueness condition as well as to offer a general solution of the equation. Furthermore, a complete solution to the inequality is given. Finally, a problem of simultaneous solution of equations and inequalities is also considered.

II. Preliminaries

In this section, we present algebraic definitions, notations, and results based on [15], [17] to provide a background for subsequent analysis and solutions. Additional details and further results can be found in [2]–[9].

A. Idempotent Semifield

We consider a set $\mathbb{X}$ endowed with addition $\oplus$ and multiplication $\otimes$ and equipped with the zero $\mathbb{0}$ and the identity $\mathbb{1}$. The system $\langle \mathbb{X}, \mathbb{0}, \mathbb{1}, \oplus, \otimes \rangle$ is assumed to be a linearly ordered radicable commutative semiring with idempotent addition and invertible multiplication, and it is commonly called idempotent semifield.

Idempotency of addition implies that $x \oplus x = x$ for all $x \in \mathbb{X}$. For any $x \in \mathbb{X}_+$, where $\mathbb{X}_+ = \mathbb{X} \setminus \{\mathbb{0}\}$, there exists an inverse x^{-1} such that $x^{-1} \otimes x = \mathbb{1}$. Furthermore, the power x^q is defined for any $x \in \mathbb{X}_+$ and a rational q. Specifically, for any integer $p \geq 0$, we have

$$x^0 = \mathbb{1}, \qquad x^p = x^{p-1}x, \qquad x^{-p} = (x^{-1})^p.$$

In what follows, we drop the multiplication sign $\otimes$ and use the power notation only in the above sense.

The linear order defined on $\mathbb{X}$ is assumed to be consistent with a partial order that is induced by idempotent addition to involve that $x \leq y$ if and only if $x \oplus y = y$. From the last definition it follows that addition possesses an extremal property in the form of inequalities

$$x \leq x \oplus y, \qquad y \leq x \oplus y,$$

and that both addition and multiplication are isotonic.

Below, the relation symbols and the operator $\min$ are thought in terms of the above linear order.

Note that we have $x \geq \mathbb{0}$ for all $x \in \mathbb{X}$. We also assume that the set $\mathbb{X}$ includes (or can be extended by) a maximal element ∞ such that $x \leq \infty$ for all $x \in \mathbb{X}$.

Examples of the linearly ordered radicable idempotent semifield under consideration include

$$\begin{aligned}
\mathbb{R}_{\max,+} &= \langle \mathbb{R} \cup \{-\infty\}, -\infty, 0, \max, + \rangle, \\
\mathbb{R}_{\min,+} &= \langle \mathbb{R} \cup \{+\infty\}, +\infty, 0, \min, + \rangle, \\
\mathbb{R}_{\max,\times} &= \langle \mathbb{R}_+ \cup \{0\}, 0, 1, \max, \times \rangle, \\
\mathbb{R}_{\min,\times} &= \langle \mathbb{R}_+ \cup \{+\infty\}, +\infty, 1, \min, \times \rangle,
\end{aligned}$$

where $\mathbb{R}$ is the set of real numbers, $\mathbb{R}_+ = \{x \in \mathbb{R} | x > 0\}$.

Specifically, the semifield $\mathbb{R}_{\max,+}$ has its null and identity defined as $\mathbb{0} = -\infty$ and $\mathbb{1} = 0$. For each $x \in \mathbb{R}$, there exists an inverse x^{-1} equal to $-x$ in conventional arithmetic. For any $x, y \in \mathbb{R}$, the power x^y corresponds to the arithmetic product xy. The order induced by the idempotent addition coincides with the natural linear order on $\mathbb{R}$. The maximal element is given by $+\infty$.

In $\mathbb{R}_{\min,\times}$, we have $\mathbb{0} = +\infty$ and $\mathbb{1} = 1$. The inverse and power notations have the same interpretation as in the conventional algebra. The relation $\leq$ defines a reverse order to the linear order on $\mathbb{R}$. The role of the maximal element is performed by 0.

The semifields $\mathbb{R}_{\max,+}$, $\mathbb{R}_{\min,+}$, $\mathbb{R}_{\max,\times}$, and $\mathbb{R}_{\min,\times}$ are isomorphic to each other. Fig. 1 offers a diagram that represents isomorphism maps for these semifields.

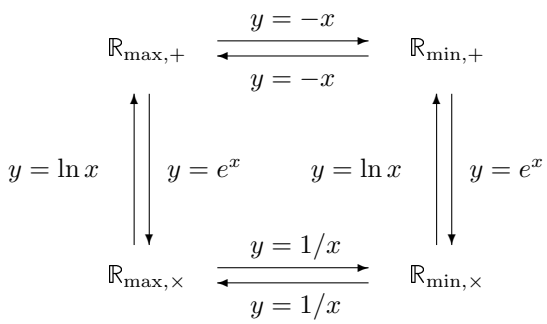

Fig. 1. Isomorphism of $\mathbb{R}_{\max,+}$, $\mathbb{R}_{\min,+}$, $\mathbb{R}_{\max,\times}$, and $\mathbb{R}_{\min,\times}$.

As an example of an idempotent semiring that is not a simifield, one can consider

$$\mathbb{R}_{\max,\min} = \langle \mathbb{R} \cup \{-\infty, +\infty\}, -\infty, +\infty, \max, \min \rangle.$$

In this semiring, it holds that $\mathbb{0} = -\infty$ and $\mathbb{1} = +\infty$. Inverse elements with respect to the multiplication $\otimes$ defined to be $\min$ do not exist, whereas the power notation is undefined. The order induced by the addition $\oplus$ corresponds to the natural linear order. The maximal element is $+\infty$.

Now we introduce a distance function ρ on $\mathbb{X}$ as follows. For any $x, y \in \mathbb{X}_+$, we define

$$\rho(x, y) = y^{-1}x \oplus x^{-1}y.$$

Since the function ρ takes values in the segment $[\mathbb{1}, \infty)$, it is natural to put $\rho(x, y) = \mathbb{1}$ when $x = y = \mathbb{0}$. We also assume that $\rho(x, y) = \infty$ if one of the arguments x and y is zero while the other is not.

In the semifield $\mathbb{R}_{\max,+}$ for all $x, y \in \mathbb{R}$, the function ρ coincides with the ordinary distance $d(x,y) = |x-y|$. Due to the isomorphism between semifields, the function ρ induces a distance function in each semifield $\mathbb{R}_{\max,\times}$, $\mathbb{R}_{\min,+}$, and $\mathbb{R}_{\min,\times}$. Specifically, in $\mathbb{R}_{\max,\times}$, we have

$$\rho'(x,y) = \ln(y^{-1}x \oplus x^{-1}y).$$

The function ρ possesses the symmetry property and satisfies the triangle inequality in all semifields $\mathbb{R}_{\max,+}$, $\mathbb{R}_{\max,\times}$, $\mathbb{R}_{\min,+}$, and $\mathbb{R}_{\min,\times}$. Moreover, for every such semifield, the function can always be converted into an actual metric by scaling its value by an appropriate isomorphism. Below, for simplicity, we leave out the isomorphism maps, and take the function ρ as a metric for all semifields.

B. Idempotent Vector Space

Consider the Cartesian power $\mathbb{X}^m$ with column vectors as its elements. A vector with all elements equal to $\mathbb{0}$ is the zero vector. A vector is regular if it has no zero elements.

For any two vectors $\boldsymbol{a} = (a_i)$ and $\boldsymbol{b} = (b_i)$ in $\mathbb{X}^m$, and a scalar $x \in \mathbb{X}$, addition and scalar multiplication are defined component-wise as follows

$$\{\boldsymbol{a} \oplus \boldsymbol{b}\}_i = a_i \oplus b_i, \qquad \{x\boldsymbol{a}\}_i = xa_i.$$

Endowed with these operations, the set $\mathbb{X}^m$ forms a semimodule over the idempotent semifield $\mathbb{X}$ and it is referred to as the idempotent vector space.

Fig. 2 gives geometrical illustrations of the addition in the space $\mathbb{R}^2_{\max,+}$ with the Cartesian coordinates on the plane. Note that the left example can equally refer to $\mathbb{R}^2_{\max,\times}$.

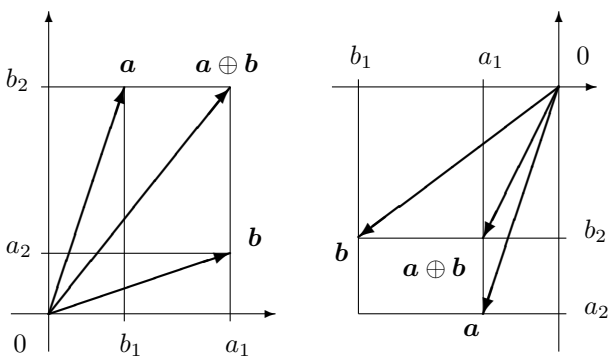

Fig. 2. Vector addition in $\mathbb{R}^2_{\max,+}$.

Idempotent addition of two vectors in $\mathbb{R}^2_{\max,+}$ follows the "rectangle rule" that defines the sum as the upper right vertex of a rectangle formed by lines drawn through the end points of the vectors, parallel to the coordinate axes.

Scalar multiplication of vectors in $\mathbb{R}^2_{\max,+}$ is equivalent to shifting the end point of the vector in the direction at $45°$ to the axes (see Fig. 3, left), whereas in $\mathbb{R}^2_{\max,\times}$ it has conventional geometrical interpretation (Fig. 3, right).

For any vectors $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{X}^m$, the extremal property of addition leads to component-wise vector inequalities

$$\boldsymbol{a} \le \boldsymbol{a} \oplus \boldsymbol{b}, \qquad \boldsymbol{b} \le \boldsymbol{a} \oplus \boldsymbol{b}.$$

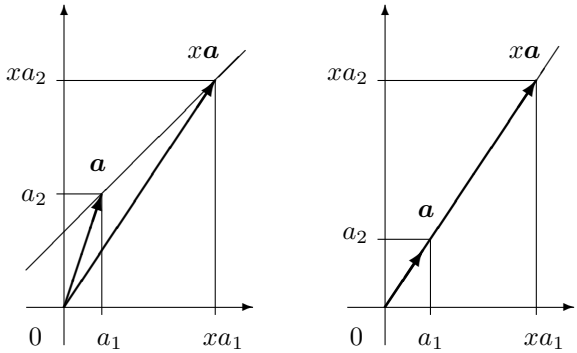

Fig. 3. Scalar multiplication in $\mathbb{R}^2_{\max,+}$ (left) and in $\mathbb{R}^2_{\max,\times}$ (right).

Furthermore, vector and scalar inequalities $\boldsymbol{a} \le \boldsymbol{b}$ and $x \le y$ imply that the inequalities

$$\boldsymbol{a} \oplus \boldsymbol{c} \le \boldsymbol{b} \oplus \boldsymbol{c}, \qquad x\boldsymbol{a} \le x\boldsymbol{b}, \qquad x\boldsymbol{c} \le y\boldsymbol{c}$$

are valid for all $\boldsymbol{c} \in \mathbb{X}^m$, and so both vector addition and scalar multiplication are isotone in each argument.

C. Matrix Algebra

Consider matrices having entries in $\mathbb{X}$. For conforming matrices $A = (a_{ij})$, $B = (b_{ij})$, and $C = (c_{ij})$, matrix addition and multiplication together with multiplication by a scalar $x \in \mathbb{X}$ follow the conventional rules

$$\{A \oplus B\}_{ij} = a_{ij} \oplus b_{ij}, \qquad \{BC\}_{ij} = \bigoplus_k b_{ik}c_{kj},$$

$$\{xA\}_{ij} = xa_{ij}.$$

Specifically, a matrix $A = (a_{ij}) \in \mathbb{X}^{m\times n}$ is multiplied by a vector $\boldsymbol{x} = (x_i) \in \mathbb{X}^n$ to result in a vector with elements

$$\{A\boldsymbol{x}\}_i = a_{i1}x_1 \oplus \cdots \oplus a_{in}x_n.$$

The matrix operations is component-wise isotone in each argument.

A matrix with all entries equal to zero is called the zero matrix and denoted by $\mathbb{0}$.

A matrix is row regular (column regular) if it has no zero rows (columns).

A square matrix is diagonal if its off-diagonal entries are zero. The diagonal matrix $I = \mathrm{diag}(\mathbb{1}, \ldots, \mathbb{1})$ is the identity.

For any nonzero column vector $\boldsymbol{x} = (x_i) \in \mathbb{X}^n$, we define a row vector $\boldsymbol{x}^- = (x_i^-)$, where $x_i^- = x_i^{-1}$ if $x_i \ne \mathbb{0}$, and $x_i^- = \mathbb{0}$ otherwise.

If $\boldsymbol{x}$ is a nonzero vector, then $\boldsymbol{x}^-\boldsymbol{x} = \mathbb{1}$.

Suppose that both $\boldsymbol{x}$ and $\boldsymbol{y}$ are regular vectors. The component-wise inequality $\boldsymbol{x} \le \boldsymbol{y}$ implies $\boldsymbol{x}^- \ge \boldsymbol{y}^-$. Furthermore, it is not difficult to verify the inequality

$$\boldsymbol{x}\boldsymbol{y}^- \ge (\boldsymbol{x}^-\boldsymbol{y})^{-1}I. \tag{1}$$

Indeed, since $\boldsymbol{x}^-\boldsymbol{y} = x_1^{-1}y_1 \oplus \cdots \oplus x_n^{-1}y_n \ge x_i^{-1}y_i$, we have $x_iy_i^{-1} \ge (\boldsymbol{x}^-\boldsymbol{y})^{-1}$ for all $i = 1, \ldots, n$, and so

$$\boldsymbol{x}\boldsymbol{y}^- \ge \mathrm{diag}(x_1y_1^{-1}, \ldots, x_ny_n^{-1}) \ge (\boldsymbol{x}^-\boldsymbol{y})^{-1}I.$$

When $\boldsymbol{y} = \boldsymbol{x}$ the inequality (1) takes the form $\boldsymbol{x}\boldsymbol{x}^- \geq I$. By applying (1) in this form, one can get the inequality

$$(A\boldsymbol{x})^- A \leq \boldsymbol{x}^-, \tag{2}$$

which is valid for any nonzero matrix $A \in \mathbb{X}^{m\times n}$. In fact, if $\boldsymbol{x} \in \mathbb{X}_+^n$, then we have $(A\boldsymbol{x})^- A \leq (A\boldsymbol{x})^- A\boldsymbol{x}\boldsymbol{x}^- = \boldsymbol{x}^-$.

D. Linear Dependence

Consider a system of vectors $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n \in \mathbb{X}^m$. As usual, a vector $\boldsymbol{b} \in \mathbb{X}^m$ is linearly dependent on the system if it admits representation as a linear combination

$$\boldsymbol{b} = x_1\boldsymbol{a}_1 \oplus \cdots \oplus x_n\boldsymbol{a}_n$$

with coefficients $x_1, \ldots, x_n \in \mathbb{X}$.

The linear span of $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ is defined as the set of all linear combinations that form an idempotent subspace

$$\operatorname{span}(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m) = \left\{ \bigoplus_{i=1}^{m} x_i\boldsymbol{a}_i \,\middle|\, x_1, \ldots, x_m \in \mathbb{X} \right\}.$$

Geometrical examples of linear spans are given in Fig. 4.

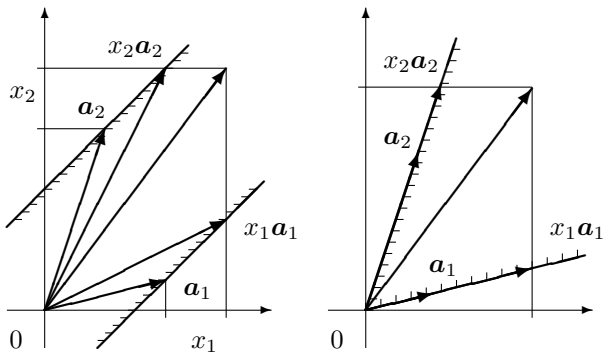

Fig. 4. Linear span of vectors in $\mathbb{R}^2_{\max,+}$ (left) and in $\mathbb{R}^2_{\max,\times}$ (right).

The linear span of vectors $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$ in the idempotent space $\mathbb{R}^2_{\max,+}$ is a strip bounded by the lines drawn through the end points of the vectors (see Fig. 4, left). In $\mathbb{R}^2_{\max,\times}$, the linear span has the form of a cone (Fig. 4, right).

A system of vectors $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ is linearly dependent if at least one its vector is linearly dependent on others, and it is linear independent otherwise.

A system of nonzero vectors $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ is a minimal generating system for a vector $\boldsymbol{b}$, if $\boldsymbol{b}$ is linearly dependent on the system and independent of any of its subsystems.

Let us verify that if vectors $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ are a minimal generating system for a vector $\boldsymbol{b}$, then representation of $\boldsymbol{b}$ as a linear combination of $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ is unique. Suppose there are two linear combinations

$$\boldsymbol{b} = x_1\boldsymbol{a}_1 \oplus \cdots \oplus x_n\boldsymbol{a}_n = x'_1\boldsymbol{a}_1 \oplus \cdots \oplus x'_n\boldsymbol{a}_n,$$

where $x'_i \neq x_i$, say $x'_i < x_i$, for some index $i = 1, \ldots, n$.

Assuming, for the sake of simplicity, that the vector $\boldsymbol{a}_i$ is regular, we have $\boldsymbol{b} \geq x_i\boldsymbol{a}_i > x'_i\boldsymbol{a}_i$. Therefore, the term $x'_i\boldsymbol{a}_i$ does not affect $\boldsymbol{b}$ and so may be omitted, which contradicts with the minimality of the system $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$.

E. Distance Function

For any vector $\boldsymbol{a} \in \mathbb{X}^m$, we introduce its support as the index set

$$\operatorname{supp}(\boldsymbol{a}) = \{i | a_i \neq \mathbb{0}, i = 1, \ldots, m\}.$$

The distance between nonzero vectors $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{X}^m$ with $\operatorname{supp}(\boldsymbol{a}) = \operatorname{supp}(\boldsymbol{b})$ is defined by a function

$$\rho(\boldsymbol{a}, \boldsymbol{b}) = \bigoplus_{i \in \operatorname{supp}(\boldsymbol{a})} \left(b_i^{-1}a_i \oplus a_i^{-1}b_i\right) = \boldsymbol{b}^-\boldsymbol{a} \oplus \boldsymbol{a}^-\boldsymbol{b}. \tag{3}$$

We put $\rho(\boldsymbol{a}, \boldsymbol{b}) = \infty$ when $\operatorname{supp}(\boldsymbol{a}) \neq \operatorname{supp}(\boldsymbol{b})$, and $\rho(\boldsymbol{a}, \boldsymbol{b}) = \mathbb{1}$ if $\boldsymbol{a} = \boldsymbol{b} = \mathbb{0}$.

Note that the function ρ in $\mathbb{R}^m_{\max,+}$ coincides for all vectors $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^m$ with the Chebyshev metric

$$\rho_\infty(\boldsymbol{a}, \boldsymbol{b}) = \max_{1 \leq i \leq m} |b_i - a_i|.$$

III. Evaluation of Distances

Let $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n \in \mathbb{X}^m$ be given vectors. We denote by $A = (\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n)$ a matrix having the vectors as columns, and by $\mathcal{A} = \operatorname{span}\{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n\}$ a linear span of the vectors.

Take a vector $\boldsymbol{d} \in \mathbb{X}^m$ and consider the problem of computing the distance from $\boldsymbol{d}$ to $\mathcal{A}$ defined as

$$\rho(\mathcal{A}, \boldsymbol{d}) = \min_{\boldsymbol{a} \in \mathcal{A}} \rho(\boldsymbol{a}, \boldsymbol{d}).$$

As another problem of interest, we examine the distance from $\boldsymbol{d}$ to the sets (half-spaces)

$$\mathcal{A}_1 = \{\boldsymbol{a} \in \mathcal{A} | \boldsymbol{a} \leq \boldsymbol{d}\}, \qquad \mathcal{A}_2 = \{\boldsymbol{a} \in \mathcal{A} | \boldsymbol{a} \geq \boldsymbol{d}\}.$$

Taking into account that every vector $\boldsymbol{a} \in \mathcal{A}$ can be represented as $\boldsymbol{a} = A\boldsymbol{x}$ for some vector $\boldsymbol{x} \in \mathbb{X}^n$, we arrive at the problem of calculating

$$\rho(\mathcal{A}, \boldsymbol{d}) = \min_{\boldsymbol{x} \in \mathbb{X}^n} \rho(A\boldsymbol{x}, \boldsymbol{d}). \tag{4}$$

Suppose $\boldsymbol{d} = \mathbb{0}$. Considering that $\mathcal{A}$ always contains the zero vector, we obviously get $\rho(\mathcal{A}, \boldsymbol{d}) = \mathbb{1}$.

Let some of the vectors $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ be zero. Since zero vectors do not affect the linear span $\mathcal{A}$, they can be removed with no change of distances. When all vectors are zero and thus $\mathcal{A} = \{\mathbb{0}\}$, we have $\rho(\mathcal{A}, \boldsymbol{d}) = \mathbb{0}$ if $\boldsymbol{d} = \mathbb{0}$, and $\rho(\mathcal{A}, \boldsymbol{d}) = \infty$ otherwise.

From here on we assume that $\boldsymbol{d} \neq \mathbb{0}$ and $\boldsymbol{a}_i \neq \mathbb{0}$ for all $i = 1, \ldots, n$, and so the matrix A is column regular.

Suppose the vector $\boldsymbol{d} = (d_i)$ may have zero components and so be irregular. For the matrix $A = (a_{ij})$, we introduce a matrix $\widehat{A} = (\widehat{a}_{ij})$ as follows. We define two sets of indices $I = \{i | d_i = \mathbb{0}\}$ and $J = \{j | a_{ij} > \mathbb{0}, i \in I\}$, and then determine the entries in $\widehat{A}$ according to the conditions

$$\widehat{a}_{ij} = \begin{cases} \mathbb{0}, & \text{if } i \notin I \text{ and } j \in J; \\ a_{ij}, & \text{otherwise.} \end{cases}$$

The matrix A may differ from $\widehat{A}$ only in those columns that have nonzero intersections with the rows corresponding to zero components in $\boldsymbol{d}$. In the matrix $\widehat{A}$, these columns have all entries that are not located at the intersections set to zero. The matrix $\widehat{A}$ with the above properties and the vector $\boldsymbol{d}$ are said to be consistent with each other.

Note that when $\boldsymbol{d}$ is regular, we have $\widehat{A} = A$; that is, the matrix $\widehat{A}$ obtained from the matrix A to provide consistency with $\boldsymbol{d}$ appears to coincide with A.

Proposition 1: For all $\boldsymbol{x}$ it holds that

$$\rho(A\boldsymbol{x}, \boldsymbol{d}) = \rho(\widehat{A}\boldsymbol{x}, \boldsymbol{d}).$$

Proof: With a regular $\boldsymbol{d}$ the statement becomes trivial and so assume $\boldsymbol{d} \neq \mathbb{0}$ to have zero components.

Suppose that $\rho(A\boldsymbol{x}, \boldsymbol{d}) < \infty$, which occurs only under the condition $\mathrm{supp}(A\boldsymbol{x}) = \mathrm{supp}(\boldsymbol{d})$. The fulfillment of the condition is equivalent to equalities $a_{i1}x_1 \oplus \cdots \oplus a_{in}x_n = \mathbb{0}$ that must be true whenever $d_i = \mathbb{0}$. To provide the equalities, we put $x_j = \mathbb{0}$ for all indices j such that $a_{ij} \neq \mathbb{0}$ for at least one index i with $d_i = \mathbb{0}$. In this case, replacing A with $\widehat{A}$ leaves the value of $\rho(A\boldsymbol{x}, \boldsymbol{d}) < \infty$ unchanged.

Since the condition $\mathrm{supp}(A\boldsymbol{x}) \neq \mathrm{supp}(\boldsymbol{d})$ implies $\mathrm{supp}(\widehat{A}\boldsymbol{x}) \neq \mathrm{supp}(\boldsymbol{d})$ and vice versa, the statement is also true when $\rho(A\boldsymbol{x}, \boldsymbol{d}) = \infty$. ■

With the above result, we may now concentrate only on the problems when A is consistent with $\boldsymbol{d}$.

In order to describe the solution of problem (4), we need the following notation. For any consistent matrix A and vector $\boldsymbol{d}$, we define a residual value

$$\Delta_A(\boldsymbol{d}) = \sqrt{(A(\boldsymbol{d}^- A)^-)^- \boldsymbol{d}}$$

if A is row regular, and $\Delta_A(\boldsymbol{d}) = \infty$ otherwise.

In what follows, we drop subscripts and arguments in $\Delta_A(\boldsymbol{d})$ and write Δ if no confusion arises.

Below we find the solution when the vector $\boldsymbol{d}$ is regular and then extend this result to irregular vectors.

A. Regular Vector

Suppose that $\boldsymbol{d}$ is a regular vector. First we verify that the minimum of $\rho(A\boldsymbol{x}, \boldsymbol{d})$ over $\mathbb{X}^n$ in (4) can be found by examining only regular vectors $\boldsymbol{x} \in \mathbb{X}^n_+$.

Proposition 2: If a vector $\boldsymbol{d}$ is regular, then

$$\rho(\mathcal{A}, \boldsymbol{d}) = \min_{\boldsymbol{x} \in \mathbb{X}^n_+} \rho(A\boldsymbol{x}, \boldsymbol{d}).$$

Proof: Take a vector $\boldsymbol{y} = A\boldsymbol{x}$ such that $\rho(A\boldsymbol{x}, \boldsymbol{d})$ achieves the minimum value. If $\boldsymbol{y}$ is irregular and so has zero components, then $\mathrm{supp}(\boldsymbol{y}) \neq \mathrm{supp}(\boldsymbol{d})$, and thus $\rho(A\boldsymbol{x}, \boldsymbol{d}) = \infty$ for all $\boldsymbol{x}$, including regular vectors $\boldsymbol{x}$.

Suppose $\boldsymbol{y} = (y_1, \ldots, y_m)^T$ is regular. Assume a corresponding vector $\boldsymbol{x}$ to have a zero component, say $x_j = \mathbb{0}$. Now we define a set $I = \{i | a_{ij} > \mathbb{0}\} \neq \emptyset$ and find a number $\varepsilon = \min\{a_{ij}^{-1} y_i | i \in I\} > \mathbb{0}$.

It remains to note that with $x_j = \varepsilon$ in place of $x_j = \mathbb{0}$, all components of $\boldsymbol{y}$ together with the minimum value of $\rho(A\boldsymbol{x}, \boldsymbol{d})$ remain unchanged. Therefore, to get the minimum it is suffice to examine only regular vectors $\boldsymbol{x} \in \mathbb{X}^n_+$. ■

The next statement reveals the meaning of the residual value $\Delta = \Delta_A(\boldsymbol{d})$ in terms of distances.

Lemma 1: If a vector $\boldsymbol{d}$ is regular, then it holds that

$$\rho(\mathcal{A}, \boldsymbol{d}) = \Delta,$$

where the minimum is attained at

$$\boldsymbol{x} = \Delta(\boldsymbol{d}^- A)^-.$$

Proof: Suppose that the matrix A is not row regular. Then we have $\mathrm{supp}(A\boldsymbol{x}) \neq \mathrm{supp}(\boldsymbol{d})$ and $\rho(\mathcal{A}, \boldsymbol{d}) = \infty$. Since, by definition, $\Delta = \infty$, the statement is true in this case.

Let A be row regular. Taking into account (3) and (4), we arrive at an optimization problem to find

$$\min_{\boldsymbol{x} \in \mathbb{X}^n_+} (\boldsymbol{d}^- A\boldsymbol{x} \oplus (A\boldsymbol{x})^- \boldsymbol{d}).$$

Take any vector $\boldsymbol{y} = A\boldsymbol{x}$ such that $\boldsymbol{x} > \mathbb{0}$, and define

$$r = \boldsymbol{d}^- A\boldsymbol{x} \oplus (A\boldsymbol{x})^- \boldsymbol{d} > \mathbb{0}.$$

From the definition of r, we have two inequalities

$$r \geq \boldsymbol{d}^- A\boldsymbol{x}, \qquad r \geq (A\boldsymbol{x})^- \boldsymbol{d}.$$

Right multiplication of the first inequality by $\boldsymbol{x}^-$ together with (1) give $r\boldsymbol{x}^- \geq \boldsymbol{d}^- A\boldsymbol{x}\boldsymbol{x}^- \geq \boldsymbol{d}^- A$. Then we obtain $\boldsymbol{x} \leq r(\boldsymbol{d}^- A)^-$ and $(A\boldsymbol{x})^- \geq r^{-1}(A(\boldsymbol{d}^- A)^-)^-$.

Further substitution into the second inequality results in

$$r \geq r^{-1}(A(\boldsymbol{d}^- A)^-)^- \boldsymbol{d} = r^{-1}\Delta^2,$$

whence it follows directly that $r \geq \Delta$.

It remains to verify the equality $r = \Delta$ when we take $\boldsymbol{x} = \Delta(\boldsymbol{d}^- A)^-$. Indeed, substitution of this vector $\boldsymbol{x}$ gives

$$r = \Delta \boldsymbol{d}^- A(\boldsymbol{d}^- A)^- \oplus \Delta^{-1}(A(\boldsymbol{d}^- A)^-)^- \boldsymbol{d} = \Delta.$$

Finally note that the above vector $\boldsymbol{x}$ corresponds to the vector $\boldsymbol{y} = \Delta A(\boldsymbol{d}^- A)^- \in \mathcal{A}$. ■

Examples of a subspace $\mathcal{A} = \mathrm{span}(\boldsymbol{a}_1, \boldsymbol{a}_2)$ and a vector $\boldsymbol{d}$ in the idempotent space $\mathbb{R}^2_{\max,+}$ are given in Fig. 5.

Now we turn to evaluation of the distance from the vector $\boldsymbol{d}$ to the half-spaces $\mathcal{A}_1$ and $\mathcal{A}_2$.

Lemma 2: If a vector $\boldsymbol{d}$ is regular, then it holds that

$$\rho(\mathcal{A}_1, \boldsymbol{d}) = \min_{A\boldsymbol{x} \leq \boldsymbol{d}} \rho(A\boldsymbol{x}, \boldsymbol{d}) = \Delta^2,$$
$$\rho(\mathcal{A}_2, \boldsymbol{d}) = \min_{A\boldsymbol{x} \geq \boldsymbol{d}} \rho(A\boldsymbol{x}, \boldsymbol{d}) = \Delta^2,$$

where the minimum is respectively attained at

$$\boldsymbol{x}_1 = (\boldsymbol{d}^- A)^-, \qquad \boldsymbol{x}_2 = \Delta^2(\boldsymbol{d}^- A)^-.$$

Proof: Similarly as in Lemma 1 we can verify the equality $\rho(\mathcal{A}_1, \boldsymbol{d}) = \rho(\mathcal{A}_2, \boldsymbol{d}) = \Delta^2$ provided that A is not a row regular matrix. Let us show that the equality remains valid when the matrix A is row regular.

By multiplying the inequality $A\boldsymbol{x} \leq \boldsymbol{d}$ by $\boldsymbol{x}^-$ from the right and applying (1) we get $A \leq A\boldsymbol{x}\boldsymbol{x}^- \leq \boldsymbol{d}\boldsymbol{x}^-$. Further

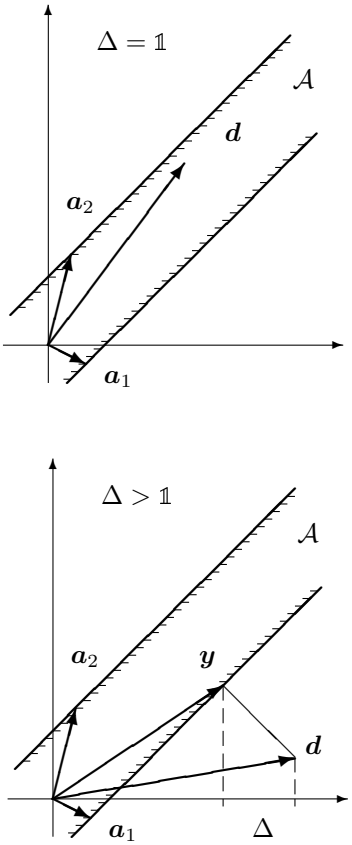

Fig. 5. A linear span $\mathcal{A}$ and a vector $\boldsymbol{d}$ in $\mathbb{R}^2_{\max,+}$ when $\Delta = \mathbb{1}$ (top) and $\Delta > \mathbb{1}$ (bottom).

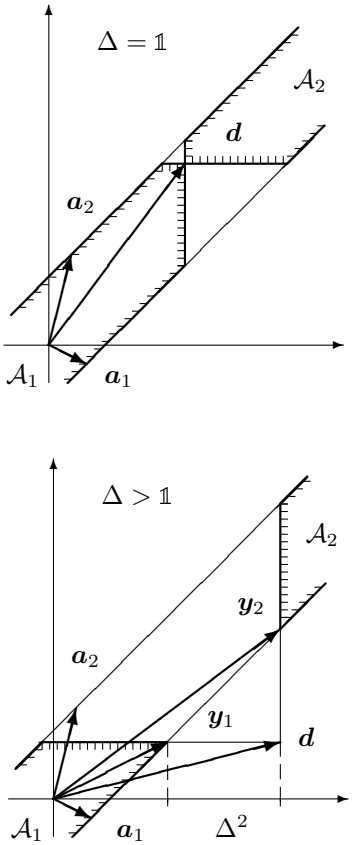

Fig. 6. The sets $\mathcal{A}_1$ and $\mathcal{A}_2$, and the vector $\boldsymbol{d}$ in $\mathbb{R}^2_{\max,+}$ when $\Delta = \mathbb{1}$ (top) and $\Delta > \mathbb{1}$ (bottom).

multiplication by $\boldsymbol{d}^-$ from the left results in the inequality $\boldsymbol{d}^- A \leq \boldsymbol{x}^-$, which then gives $\boldsymbol{x} \leq (\boldsymbol{d}^- A)^-$.

Therefore, for any vector $A\boldsymbol{x} \in \mathcal{A}_1$, we have

$$\rho(A\boldsymbol{x}, \boldsymbol{d}) = (A\boldsymbol{x})^- \boldsymbol{d} \geq (A(\boldsymbol{d}^- A)^-)^- \boldsymbol{d} = \Delta^2.$$

It is clear that $\rho(A\boldsymbol{x}_1, \boldsymbol{d}) = \Delta^2$ if $\boldsymbol{x}_1 = (\boldsymbol{d}^- A)^-$.

Consider an arbitrary vector $A\boldsymbol{x} \in \mathcal{A}_2$. Application of (2) to the vector $(\boldsymbol{d}^- A)^-$ yields $\boldsymbol{d}^- A \geq (A(\boldsymbol{d}^- A)^-)^- A$. Taking into account the condition that $A\boldsymbol{x} \geq \boldsymbol{d}$, we further have $\boldsymbol{d}^- A\boldsymbol{x} \geq (A(\boldsymbol{d}^- A)^-)^- A\boldsymbol{x} \geq (A(\boldsymbol{d}^- A)^-)^- \boldsymbol{d}$.

Now we can conclude that for $A\boldsymbol{x} \in \mathcal{A}_2$, it holds

$$\rho(A\boldsymbol{x}, \boldsymbol{d}) = \boldsymbol{d}^- A\boldsymbol{x} \geq (A(\boldsymbol{d}^- A)^-)^- \boldsymbol{d} = \Delta^2.$$

It remains to see that substitution $\boldsymbol{x}_2 = \Delta^2 (\boldsymbol{d}^- A)^-$ gives $\rho(A\boldsymbol{x}_2, \boldsymbol{d}) = \Delta^2 \boldsymbol{d}^- A(\boldsymbol{d}^- A)^- = \Delta^2$.

Note that when $\Delta < \infty$ the minimum distance from $\boldsymbol{d}$ to the half-spaces $\mathcal{A}_1$ and $\mathcal{A}_2$ is achieved at the respective vectors $\boldsymbol{y}_1 = A(\boldsymbol{d}^- A)^-$ and $\boldsymbol{y}_2 = \Delta^2 A(\boldsymbol{b}^- A)^-$. ■

A geometric illustration of the above result in the idempotent space $\mathbb{R}^2_{\max,+}$ is given in Fig. 6.

B. Arbitrary Nonzero Vector

Now we examine the distance between the linear span $\mathcal{A}$ and an arbitrary vector $\boldsymbol{d} \neq \mathbb{0}$.

Theorem 1: For any vector $\boldsymbol{d} \neq \mathbb{0}$ it holds that

$$\rho(\mathcal{A}, \boldsymbol{d}) = \min_{\boldsymbol{x} \in \mathbb{X}^n_+} \rho(A\boldsymbol{x}, \boldsymbol{d}) = \Delta,$$

where the minimum is attained at $\boldsymbol{x} = \Delta(\boldsymbol{d}^- A)^-$.

Proof: Note that for the case when $\boldsymbol{d}$ is regular, the proof is given in Lemma 1. Now we suppose that the vector $\boldsymbol{d} \neq \mathbb{0}$ has zero components. Due to Proposition 1, it will suffice to examine only the case when A and $\boldsymbol{d}$ are consistent.

Let us define the sets of indices $I = \{i | d_i = \mathbb{0}\}$ and $J = \{j | a_{ij} > \mathbb{0}, i \in I\}$. In order to provide the minimum of $\rho(A\boldsymbol{x}, \boldsymbol{d})$, we must put $x_j = \mathbb{0}$ for all $j \in J$. This makes it possible to exclude from consideration all components of $\boldsymbol{d}$ and the rows of A with indices in I, as well as all columns of A with indices in J. By eliminating these elements, we obtain a new matrix A' and a new vector $\boldsymbol{d}'$.

Denote the linear span of the columns in A' by $\mathcal{A}'$. Considering that the vector $\boldsymbol{d}'$ has no zero components, we apply Lemma 1 to get

$$\rho(\mathcal{A}, \boldsymbol{d}) = \rho(\mathcal{A}', \boldsymbol{d}') = \Delta_{A'}(\boldsymbol{d}') = \Delta'.$$

Furthermore, we note that the minimum $\rho(A'\boldsymbol{x}', \boldsymbol{d}')$ is attained if $\boldsymbol{x}' = \Delta'(\boldsymbol{d}'^- A')^-$, where $\boldsymbol{x}'$ is a vector of order less than n.

The matrix A differs from A' only in that it has extra zero rows and columns. Clearly, both matrices appear to be row regular or irregular simultaneously.

Suppose that both matrices are row regular. Taking into account that the vector $\boldsymbol{d}'$ is obtained from $\boldsymbol{d}$ by removing zero components, we have

$$\Delta' = \sqrt{(A'(\boldsymbol{d}'^{-}A')^{-})^{-}\boldsymbol{d}'} = \sqrt{(A(\boldsymbol{d}^{-}A)^{-})^{-}\boldsymbol{d}} = \Delta.$$

Since the optimal vector $\boldsymbol{x}$ differs from $\boldsymbol{x}'$ only in extra zero components, we conclude that $\rho(A\boldsymbol{x}, \boldsymbol{d})$ achieves minimum at $\boldsymbol{x} = \Delta(\boldsymbol{d}^{-}A)^{-}$. ■

Using the same proof scheme as above, it is not difficult to extend the result of Lemma 2 as follows.

Lemma 3: For any vector $\boldsymbol{d} \neq \mathbb{0}$, it holds that

$$\rho(\mathcal{A}_1, \boldsymbol{d}) = \min_{A\boldsymbol{x} \leq \boldsymbol{d}} \rho(A\boldsymbol{x}, \boldsymbol{d}) = \Delta^2,$$
$$\rho(\mathcal{A}_2, \boldsymbol{d}) = \min_{A\boldsymbol{x} \geq \boldsymbol{d}} \rho(A\boldsymbol{x}, \boldsymbol{d}) = \Delta^2,$$

where the minimums are respectively attained at

$$\boldsymbol{x}_1 = (\boldsymbol{d}^{-}A)^{-}, \qquad \boldsymbol{x}_2 = \Delta^2(\boldsymbol{d}^{-}A)^{-}.$$

To conclude this section, let us formulate a direct consequence of Theorem 1. First note that the residual Δ satisfies the condition $\Delta \geq \mathbb{1}$. The equality $\Delta = \mathbb{1}$ means that the vector $\boldsymbol{d}$ belongs to the linear span $\mathcal{A} = \mathrm{span}\{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n\}$, whereas the inequality $\Delta > \mathbb{1}$ implies that $\boldsymbol{d}$ is outside $\mathcal{A}$. In other words, the following assertion is valid.

Lemma 4: A vector $\boldsymbol{d}$ belongs to a linear span of columns of a matrix A if and only if $\Delta = \mathbb{1}$, and if so it holds that $\boldsymbol{d} = A\boldsymbol{x}$, where $\boldsymbol{x} = (\boldsymbol{d}^{-}A)^{-}$.

In the next sections, we consider applications of the result to analysis of linear dependence and to solution of linear equations and inequalities.

IV. Linear Dependence

First we give conditions for a vector $\boldsymbol{d} \in \mathbb{X}^m$ to be linearly dependent on vectors $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n \in \mathbb{X}^m$, or equivalently, to admit a representation in the form of a linear combination $\boldsymbol{d} = x_1\boldsymbol{a}_1 \oplus \cdots \oplus x_n\boldsymbol{a}_n$.

We define the matrix $A = (\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n)$ and then calculate the residual $\Delta = \Delta_A(\boldsymbol{d}) = \sqrt{(A(\boldsymbol{d}^{-}A)^{-})^{-}\boldsymbol{d}}$.

As a slight modification of Lemma 4, we arrive at the following statement.

Lemma 5: A vector $\boldsymbol{d}$ is linearly dependent on vectors $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ if and only if $\Delta = \mathbb{1}$.

Now we formulate a formal criterion that a system $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ is linearly independent. We use the notation $A_i = (\boldsymbol{a}_1, \ldots, \boldsymbol{a}_{i-1}, \boldsymbol{a}_{i+1}, \ldots, \boldsymbol{a}_n)$ to represent a matrix obtained from A by removing column i, and introduce

$$\delta(A) = \min_{1 \leq i \leq n} \Delta_{A_i}(\boldsymbol{a}_i).$$

Lemma 6: The system of vectors $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ is linearly independent if and only if $\delta(A) > \mathbb{1}$.

Proof: Clearly, the condition $\delta(A) > \mathbb{1}$ involves that $\Delta_{A_i}(\boldsymbol{a}_i) > \mathbb{1}$ for all $i = 1, \ldots, n$. It follows from Lemma 5 that in this case none of the vectors $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ is a linear combination of others, and so the system of vectors is linearly independent. ■

Let $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ and $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_k$ be two systems of nonzero vectors. These systems are considered to be equivalent if each vector of one system is a linear combination of vectors of the other system.

Consider a system $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ that can include linearly dependent vectors. To construct an equivalent independent system (a basis of the system), we implement a sequential procedure that examines the vectors one by one to decide whether to remove a vector from the system or not.

At each step $i = 1, \ldots, n$, the vector $\boldsymbol{a}_i$ is removed if $\Delta_{\widetilde{A}_i}(\boldsymbol{a}_i) = \mathbb{1}$, where the matrix $\widetilde{A}_i$ is composed of those columns in A_i, that are retained after the previous steps. Upon completion of the procedure, we get a new system of vectors $\widetilde{\boldsymbol{a}}_1, \ldots, \widetilde{\boldsymbol{a}}_k$, where $k \leq n$.

Proposition 3: The system $\widetilde{\boldsymbol{a}}_1, \ldots, \widetilde{\boldsymbol{a}}_k$ is a linearly independent system that is equivalent to $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$.

Proof: According to the way of constructing the system $\widetilde{\boldsymbol{a}}_1, \ldots, \widetilde{\boldsymbol{a}}_k$, each vector $\widetilde{\boldsymbol{a}}_i$ coincides with a vector of the initial system $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$. Since at the same time, for each $\boldsymbol{a}_j$, it holds that $\boldsymbol{a}_j \in \mathrm{span}\{\widetilde{\boldsymbol{a}}_1, \ldots, \widetilde{\boldsymbol{a}}_k\}$, both systems are equivalent. Finally, due to Lemma 6, the system $\widetilde{\boldsymbol{a}}_1, \ldots, \widetilde{\boldsymbol{a}}_k$ is linearly independent. ■

V. Linear Equations and Inequalities

Suppose there are given a matrix $A \in \mathbb{X}^{m \times n}$ and a vector $\boldsymbol{d} \in \mathbb{X}^m$. Consider problems of finding an unknown vector $\boldsymbol{x} \in \mathbb{X}^n$ to satisfy the equation

$$A\boldsymbol{x} = \boldsymbol{d}, \tag{5}$$

and the inequality

$$A\boldsymbol{x} \leq \boldsymbol{d}. \tag{6}$$

In what follows, we assume that the matrix A is already put into a form that is consistent with the vector $\boldsymbol{d}$, and use the notation $\Delta = \Delta_A(\boldsymbol{d}) = \sqrt{(A(\boldsymbol{d}^{-}A)^{-})^{-}\boldsymbol{d}}$.

If a matrix $A = (\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n)$ has a zero column, say $\boldsymbol{a}_i$, then the solution of equation (5) reduces to that of an equation that is obtained from (5) by removing the component x_i in the vector $\boldsymbol{x}$ together with eliminating the column $\boldsymbol{a}_i$ in A. Each solution of the reduced equation causes equation (5) to have a set of solutions, where x_i takes all values in $\mathbb{X}$. The solution of inequality (6) with a matrix A having a zero column reduces in the same way.

Suppose that $A = \mathbb{0}$. In this case, any vector $\boldsymbol{x} \in \mathbb{X}^n$ is a solution of (5) provided that $\boldsymbol{d} = \mathbb{0}$, and there is no solution otherwise. The solution of (6) is any vector $\boldsymbol{x} \in \mathbb{X}^n$.

If $\boldsymbol{d} = \mathbb{0}$, then both equation (5) and inequality (6) have a trivial solution $\boldsymbol{x} = \mathbb{0}$, which is unique when the matrix A has no zero columns.

From here on we assume that the vector $\boldsymbol{d}$ and all columns in the matrix A are nonzero, and so A is column regular.

A solution $\boldsymbol{x}_0$ of equation (5) is called maximal if it holds that $\boldsymbol{x} \leq \boldsymbol{x}_0$ for any solution $\boldsymbol{x}$.

The next result gives a complete solution of inequality (6).

Lemma 7: For any column regular matrix A and vector $\boldsymbol{b} \neq \mathbb{0}$, the solution of inequality (6) exists and is given by

$$\boldsymbol{x} \leq (\boldsymbol{d}^- A)^-. \tag{7}$$

Proof: Let us ensure that inequalities (7) and (6) are equivalent to each other. First assume the vector $\boldsymbol{d}$ to be regular. In much the same way as in Lemma 2, we verify that inequality (7) follows from (6).

Suppose that inequality (7) holds. Then we have

$$A\boldsymbol{x} \leq A(\boldsymbol{b}^- A)^- \leq \boldsymbol{b}\boldsymbol{b}^- A(\boldsymbol{b}^- A)^- = \boldsymbol{b},$$

and thus inequality (6) holds as well.

Now assume that the vector $\boldsymbol{d} \neq \mathbb{0}$ is not regular. In this case, we use the same proof scheme as in Theorem 1 to reduce the problem to that with a regular vector and then apply the above result. ■

In the following, we examine conditions for the solution of equation (5) to exist and to be unique, and then describe the general solution to the equation.

A. Existence and Uniqueness of Solution

Application of previous results brings us to a position to arrive at the next assertion.

Theorem 2: For any column regular matrix A and nonzero vector $\boldsymbol{b}$, the following statements are true:

1) Equation (5) has solutions if and only if $\Delta = \mathbb{1}$.
2) If solvable, the equation has a maximal solution
$$\boldsymbol{x} = (\boldsymbol{d}^- A)^-.$$
3) If all columns in A form a minimal system that generates $\boldsymbol{d}$, then the above solution is unique.

Proof: The existence condition and the form of a solution follows from Lemma 4. The result of Lemma 7 says that this solution is maximal. The uniqueness condition follows from representation of the vector as a unique linear combination of its minimal set of generators. ■

Suppose that $\Delta > \mathbb{1}$. In this case equation (5) has no solution. However, we can define a pseudo-solution to (5) as a solution of the equation

$$A\boldsymbol{x} = \Delta A(\boldsymbol{d}^- A)^-,$$

which always exists and takes the form

$$\boldsymbol{x}_0 = \Delta(\boldsymbol{d}^- A)^-.$$

By Theorem 1, the pseudo-solution yields the minimum deviation between the vectors $\boldsymbol{y} = A\boldsymbol{x}$ and the vector $\boldsymbol{d}$ in the sense of the metric ρ. When $\Delta = \mathbb{1}$, the pseudo-solution obviously coincides with the maximum solution.

Consider a problem of finding two vectors $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ that provide the minimum deviation between both sides of (5), while satisfying the respective inequalities

$$A\boldsymbol{x} \leq \boldsymbol{d}, \qquad A\boldsymbol{x} \geq \boldsymbol{d}.$$

It follows from Lemma 3 that the problem has a solution that is given by

$$\boldsymbol{x}_1 = (\boldsymbol{d}^- A)^-, \qquad \boldsymbol{x}_2 = \Delta^2(\boldsymbol{d}^- A)^-.$$

B. General Solution

To describe a general solution to equation (5), we first give an auxiliary result that solves (5) when the vector $\boldsymbol{d}$ is linearly dependent on a subset of columns in the matrix A.

Lemma 8: Let $A = (\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n)$ be a matrix, I be a subset of column indices of A, and $\boldsymbol{d} \in \operatorname{span}\{\boldsymbol{a}_i | i \in I\}$.

Then any vector $\boldsymbol{x}_I = (x_i)$, where $x_i = (\boldsymbol{d}^- \boldsymbol{a}_i)^-$ if $i \in I$, and $x_i \leq (\boldsymbol{d}^- \boldsymbol{a}_i)^-$ otherwise, is a solution to (5).

Proof: Since $\boldsymbol{d} \in \operatorname{span}\{\boldsymbol{a}_i | i \in I\} \subset \operatorname{span}\{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n\}$, there is a solution $\boldsymbol{x}_I$ of equation (5), and therefore,

$$\boldsymbol{d} = A\boldsymbol{x}_I = \bigoplus_{i=1}^{n} x_i \boldsymbol{a}_i = \bigoplus_{i \in I} x_i \boldsymbol{a}_i \oplus \bigoplus_{i \notin I} x_i \boldsymbol{a}_i.$$

Furthermore, the condition $\boldsymbol{d} \in \operatorname{span}\{\boldsymbol{a}_i | i \in I\}$ yields an equality

$$\boldsymbol{d} = \bigoplus_{i \in I} x_i \boldsymbol{a}_i,$$

which is valid when $x_i = (\boldsymbol{d}^- \boldsymbol{a}_i)^-$ for all $i \in I$.

The remaining components with indices $i \notin I$ must be set so as to satisfy inequalities

$$\boldsymbol{d} \geq \bigoplus_{i \notin I} x_i \boldsymbol{a}_i \geq x_i \boldsymbol{a}_i.$$

It remains to solve the inequalities to conclude that for each $i \notin I$, we can take any $x_i \leq (\boldsymbol{d}^- \boldsymbol{a}_i)^-$. ■

Let I be a set of indices of those columns in A that form a minimal generating system for the vector $\boldsymbol{d}$. We denote the set of all such index sets I by $\mathcal{I}$. It is clear that $\mathcal{I} \neq \emptyset$ only when equation (5) has at least one solution.

By applying Lemma 8, we arrive at the following result.

Theorem 3: The general solution to equation (5) is a (possible empty) family of solutions $\{\boldsymbol{x}_I | I \in \mathcal{I}\}$, where each solution $\boldsymbol{x}_I = (x_i)$ is given by

$$\begin{aligned} x_i &= (\boldsymbol{d}^- \boldsymbol{a}_i)^-, \qquad \text{if } i \in I, \\ x_i &\leq (\boldsymbol{d}^- \boldsymbol{a}_i)^-, \qquad \text{if } i \notin I. \end{aligned} \tag{8}$$

Let us examine a case when the family reduces to one solution set. Suppose that the columns in A are linearly independent. Then there may exist only one subset of columns that form a minimal generating system for $\boldsymbol{d}$. If the subset coincides with the set of all columns, then the solution reduces to a unique vector $\boldsymbol{x} = (\boldsymbol{d}^- A)^-$.

As an illustration, consider equation (5) in the idempotent vector space $\mathbb{R}^2_{\max,+}$ under the conditions that

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \qquad \boldsymbol{d} = \begin{pmatrix} d_1 \\ d_2 \end{pmatrix},$$

where $a_{11}, a_{12}, a_{21}, a_{22} > \mathbb{0}$ and $d_1, d_2 > \mathbb{0}$.

Suppose that $\Delta = \Delta_A(\boldsymbol{d}) = (A(\boldsymbol{d}^- A)^-)^- \boldsymbol{d} = \mathbb{1}$. The maximal solution of the equation takes the form

$$\boldsymbol{x} = \begin{pmatrix} (\boldsymbol{d}^- \boldsymbol{a}_1)^{-1} \\ (\boldsymbol{d}^- \boldsymbol{a}_2)^{-1} \end{pmatrix} = \begin{pmatrix} (d_1^{-1} a_{11} \oplus d_2^{-1} a_{21})^{-1} \\ (d_1^{-1} a_{12} \oplus d_2^{-1} a_{22})^{-1} \end{pmatrix}.$$

If the vector $\boldsymbol{d}$ is not collinear with any of vectors $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$, then the solution is unique (see Fig. 7, upper left).

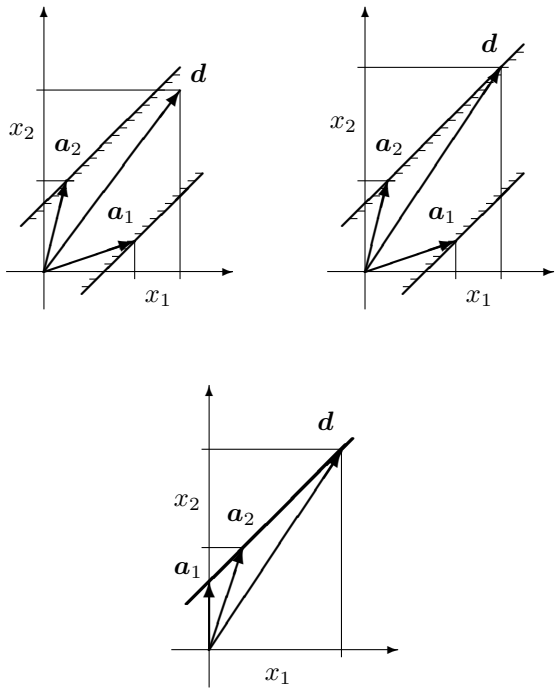

Fig. 7. A unique (upper left) and non-unique (upper right and bottom) solutions to the linear equation $A\boldsymbol{x} = \boldsymbol{d}$ in $\mathbb{R}^2_{\max,+}$.

Two cases when equation (5) has more than one solution are shown on the Fig. 7 (upper left and bottom).

In the upper left, the vector $\boldsymbol{d}$ is collinear with $\boldsymbol{a}_1$ and is not collinear with $\boldsymbol{a}_2$. The solution of the equation is any vector $\boldsymbol{x}$ with components

$$x_1 = (d_1^{-1}a_{11} \oplus d_2^{-1}a_{21})^{-1},$$
$$x_2 \leq (d_1^{-1}a_{12} \oplus d_2^{-1}a_{22})^{-1}.$$

In the case depicted in Fig. 7 in the bottom, both vectors $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$ are collinear with each other and with the vector $\boldsymbol{d}$. Under these conditions, there are two families of solution vectors $\boldsymbol{x}' = (x_1', x_2')^T$ and $\boldsymbol{x}'' = (x_1'', x_2'')$.

The vectors in the first family are given by

$$x_1' = (d_1^{-1}a_{11} \oplus d_2^{-1}a_{21})^{-1},$$
$$x_2' \leq (d_1^{-1}a_{12} \oplus d_2^{-1}a_{22})^{-1},$$

whereas those in the second family are given by

$$x_1'' \leq (d_1^{-1}a_{11} \oplus d_2^{-1}a_{21})^{-1},$$
$$x_2'' = (d_1^{-1}a_{12} \oplus d_2^{-1}a_{22})^{-1}.$$

C. Systems of Equations and Inequalities

Given matrices A and C, and vectors $\boldsymbol{d}$ and $\boldsymbol{b}$, consider a problem to find vectors $\boldsymbol{x}$ that simultaneously solves an equation and an inequality combined into one system

$$\begin{aligned} A\boldsymbol{x} &= \boldsymbol{d}, \\ C\boldsymbol{x} &\leq \boldsymbol{b}. \end{aligned} \tag{9}$$

To solve the problem, we denote by I a subset of indices for those columns in the matrix A that form a minimal generating system for the vector $\boldsymbol{b}$, and by $\mathcal{I}$ a set of all such subsets. Furthermore, we define

$$\widetilde{\mathcal{I}} = \{I \in \mathcal{I} | \boldsymbol{b}^- \boldsymbol{c}_i \leq \boldsymbol{d}^- \boldsymbol{a}_i, i \in I\} \subset \mathcal{I},$$

where $\boldsymbol{a}_i$ and $\boldsymbol{c}_i$ are columns i in the matrices A and C, and recall the notation $\Delta = \Delta_A(\boldsymbol{d}) = \sqrt{(A(\boldsymbol{d}^- A)^-)^- \boldsymbol{d}}$.

Lemma 9: For any column regular matrices A and C, and nonzero vectors $\boldsymbol{d}$ and $\boldsymbol{b}$, system (9) has solutions if and only if $\Delta = \mathbb{1}$ and $\widetilde{\mathcal{I}} \neq \emptyset$. The general solution to the system is a (possible empty) family of solutions $\{\boldsymbol{x}_I | I \in \widetilde{\mathcal{I}}\}$, where each solution $\boldsymbol{x}_I = (x_i)$ is given by

$$\begin{aligned} x_i &= (\boldsymbol{d}^- \boldsymbol{a}_i)^-, && \text{if } i \in I, \\ x_i &\leq (\boldsymbol{d}^- \boldsymbol{a}_i \oplus \boldsymbol{b}^- \boldsymbol{c}_i)^-, && \text{if } i \notin I. \end{aligned}$$

Proof: System (9) is solvable if and only if there exists a solution to the equation in the system to satisfy the condition $\boldsymbol{x} \leq (\boldsymbol{b}^- C)^-$, which is equivalent to the inequality at (9).

According to Theorem 3, the solution to the equation is a family $\{\boldsymbol{x}_I | I \in \mathcal{I}\}$, where each member is given by (8). Consider a solution that corresponds to a subset $I \in \mathcal{I}$. The solution is the vectors $\boldsymbol{x}_I = (x_i)$ with components $x_i = (\boldsymbol{d}^- \boldsymbol{a}_i)^-$, if $i \in I$, and $x_i \leq (\boldsymbol{d}^- \boldsymbol{a}_i)^-$, otherwise.

These vectors include solutions to the inequality at (9) if and only if $(\boldsymbol{d}^- \boldsymbol{a}_i)^- \leq (\boldsymbol{b}^- \boldsymbol{c}_i)^-$ for all $i \in I$. By collecting all sets I that provides this condition, we get the set $\widetilde{\mathcal{I}}$.

It remains to see that each set $I \in \widetilde{\mathcal{I}}$ gives a solution $\boldsymbol{x}_I = (x_i)$, where $x_i = (\boldsymbol{d}^- \boldsymbol{a}_i \oplus \boldsymbol{b}^- \boldsymbol{c}_i) = (\boldsymbol{d}^- \boldsymbol{a}_i)^-$, if $i \in I$, and $x_i \leq (\boldsymbol{d}^- \boldsymbol{a}_i \oplus \boldsymbol{b}^- \boldsymbol{c}_i)^-$, otherwise. ■

D. An Extended Equation

Consider a problem to solve with respect to the unknown vector $\boldsymbol{x}$ an extended equation in the form

$$A\boldsymbol{x} \oplus \boldsymbol{b} = \boldsymbol{d}, \tag{10}$$

where A is given matrix, $\boldsymbol{d}$ and $\boldsymbol{b}$ are given vectors.

In what follows, we assume that $\boldsymbol{b} \leq \boldsymbol{d}$ since if it is not the case, then equation 10 obviously has no solutions.

We introduce two sets of indices $I_1 = \{i | b_i < d_i\}$ and $I_2 = \{i | b_i = d_i\}$. Let A_1 and A_2 be submatrices composed of the rows in A with indices from I_1 and I_2, respectively. Similarly, we define subvectors $\boldsymbol{d}_1$ and $\boldsymbol{d}_2$ for the vector $\boldsymbol{d}$, and subvectors $\boldsymbol{b}_1$ and $\boldsymbol{b}_2$ for the vector $\boldsymbol{b}$.

Equation (10) is then equivalent to a system

$$\begin{aligned} A_1\boldsymbol{x} &= \boldsymbol{d}_1, \\ A_2\boldsymbol{x} &\leq \boldsymbol{b}_2. \end{aligned}$$

In the same way as above, we construct a set $\mathcal{I}_1$ to include all sets of indices of minimal systems of columns in A_1 that generate $\boldsymbol{d}_1$. Furthermore, we reduce the set $\mathcal{I}_1$ to a set $\widetilde{\mathcal{I}}_1$ consisting of those $I \in \mathcal{I}_1$ that provide common solutions to both equation and inequality. Finally, we introduce the notation $\Delta_1 = \Delta_{A_1}(\boldsymbol{d}_1) = (A(\boldsymbol{d}_1^- A)^-)^- \boldsymbol{d}_1$.

Lemma 10: For any column regular matrix A, and nonzero vectors $\boldsymbol{d}$ and $\boldsymbol{b}$, equation (10) has solutions if and only if $\Delta_1 = \mathbb{1}$ and $\widetilde{\mathcal{I}}_1 \neq \emptyset$. The general solution to the system is a (possible empty) family of solutions $\{\boldsymbol{x}_I | I \in \widetilde{\mathcal{I}}_1\}$, where each solution $\boldsymbol{x}_I = (x_i)$ is given by

$$x_i = (\boldsymbol{d}^- \boldsymbol{a}_i)^-, \qquad \text{if } i \in I,$$
$$x_i \leq (\boldsymbol{d}^- \boldsymbol{a}_i)^-, \qquad \text{if } i \notin I.$$

Proof: By applying Lemma 9, we immediately get the existence conditions as well as a general solution in the form of family $\{\boldsymbol{x}_I | I \in \widetilde{\mathcal{I}}_1\}$ with $\boldsymbol{x}_I = (x_i)$ given by

$$x_i = (\boldsymbol{d}_1^- \boldsymbol{a}_i^1 \oplus \boldsymbol{b}_2^- \boldsymbol{a}_i^2)^-, \qquad \text{if } i \in I,$$
$$x_i \leq (\boldsymbol{d}_1^- \boldsymbol{a}_i^1 \oplus \boldsymbol{b}_2^- \boldsymbol{a}_i^2)^-, \qquad \text{if } i \notin I,$$

where $\boldsymbol{a}_i^1$ and $\boldsymbol{a}_i^2$ denote columns i in A_1 and A_2.

Since $\boldsymbol{b}_2 = \boldsymbol{d}_2$, we have $\boldsymbol{d}_1^- \boldsymbol{a}_i^1 \oplus \boldsymbol{b}_2^- \boldsymbol{a}_i^2 = \boldsymbol{d}^- \boldsymbol{a}_i$ which leads to the solution in the desired form. ■

An illustration of solutions to equation (10) in $\mathbb{R}^2_{\max,+}$ are given in Fig. 8, where the set of vectors $\boldsymbol{y} = A\boldsymbol{x} \oplus \boldsymbol{b}$ for all $\boldsymbol{x} \in \mathbb{R}^2_{\max,+}$ is represented with a shaded part of the strip corresponding to the linear span of columns in A. In the case when the vector $\boldsymbol{b}$ is outside the strip (upper right and bottom), this set is extended by adding vertical segments drawn from the end point of $\boldsymbol{b}$ to the strip.

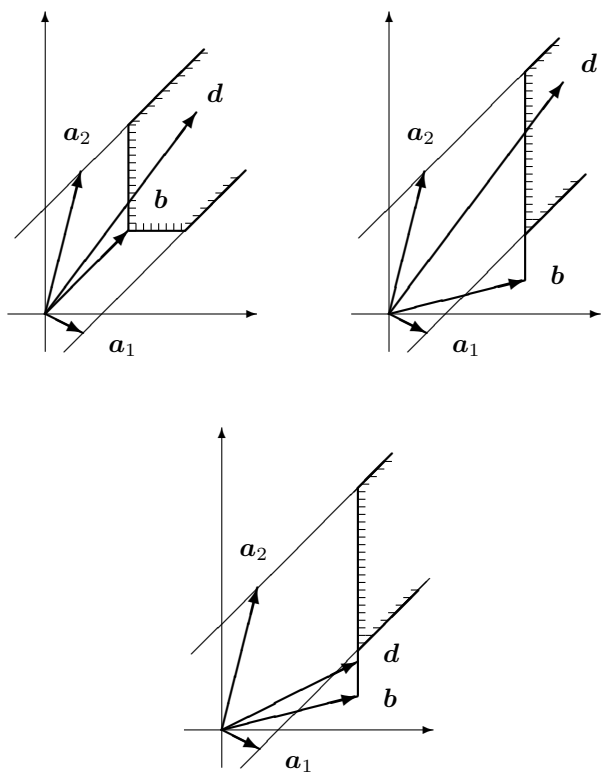

Fig. 8. Solutions to the equation $A\boldsymbol{x} \oplus \boldsymbol{b} = \boldsymbol{d}$ in $\mathbb{R}^2_{\max,+}$

Note that a solution of equation (10) may exist when equation (5) has no solutions as shown in Fig. 8 (bottom).

References

[1] N. N. Vorob'ev, "The extremal matrix algebra," *Soviet Math. Dokl.*, vol. 4, pp. 1220–1223, 1964.

[2] N. N. Vorobjev, "Extremal algebra of positive matrices," *Elektronische Informationsverarbeitung und Kybernetik*, vol. 3, no. 1, pp. 39–72, January–February 1967. (in Russian)

[3] R. Cuninghame-Green, *Minimax Algebra*. Berlin: Springer, 1979. (Lecture Notes in Economics and Mathematical Systems, vol. 166)

[4] F. L. Baccelli, G. Cohen, G. J. Olsder, J.-P. Quadrat, *Synchronization and Linearity*. Chichester: Wiley, 1993.

[5] V. N. Kolokoltsov, V. P. Maslov, *Idempotent Analysis and its Applications*. Dordrecht: Kluwer, 1997.

[6] G. L. Litvinov, V. P. Maslov, "Correspondence principle for idempotent calculus and some computer applications," in *Idempotency*. Cambridge: Cambridge University Press, 1998, pp. 420–443. E-print arXiv:math.GM/0101021

[7] J. S. Golan, *Semirings and Affine Equations Over Them*. New York: Springer, 2003.

[8] B. Heidergott, G. J. Olsder, J. van der Woude, *Max-Plus at Work*. Princeton: Princeton University Press, 2005.

[9] P. Butkovič, *Max-Linear Systems: Theory and Algorithms*. London: Springer, 2010.

[10] A. A. Korbut, "Extremal spaces," *Soviet Math.*, vol. 6, part 2, pp. 1358–1361, 1965.

[11] A. A. Korbut, "Extremal vector spaces and their properties," *Elektronische Informationsverarbeitung und Kybernetik*, vol. 8, no. 8/9, pp. 525–536, 1972. (in Russian)

[12] U. Zimmermann, *Linear and Combinatorial Optimization in Ordered Algebraic Structures*. Amsterdam: North-Holland, 1981. 390 p. (Annals of Discrete Mathematics, vol. 10)

[13] G. Cohen, P. Moller, J.-P. Quadrat, M. Viot, "Algebraic tools for the performance evaluation of discrete event systems," *Proc. IEEE*, vol. 77, no. 1, pp. 39–85, January 1989.

[14] G. J. Olsder, C. Roos, "Cramer and Cayley-Hamilton in the max algebra," *Linear Algebra Appl.*, vol. 101, pp. 87–108, April 1988.

[15] N. K. Krivulin, *Methods of Idempotent Algebra in Problems of Complex Systems Modeling and Analysis*. St. Petersburg: St. Petersburg University Press, 2009. (in Russian)

[16] N. K. Krivulin, "On solution of linear vector equations in idempotent algebra," in *Mathematical Models: Theory and Applications. Issue 5*. St. Petersburg, 2005, pp. 105–113. (in Russian)

[17] N. K. Krivulin, "On solution of a class of linear vector equations in idempotent algebra," *Vestnik St. Petersburg University. Series 10. Applied Mathematics, Informatics, Control Processes*, no. 3, pp. 64-77, September 2009. (in Russian)

[18] N. Krivulin, "A solution of a tropical linear vector equation," in *Advances in Computer Science: Proc. 6th WSEAS European Computing Conf.*, WSEAS Press, 2012, pp. 244–249.

Nikolai Krivulin received a university degree in applied mathematics and operations research in 1983 from St. Petersburg State University (SPbSU), St. Petersburg, Russia. He got his Ph.D. degree in 1990 and D.Sc. degree in 2010 both in applied mathematics from SPbSU. In 1983 he joined the Computer Center at SPbSU as a System Software Engineer, and in 1985 started his Ph.D. study. In 1987 he joined the Faculty of Mathematics and Mechanics at SPbSU as an Assistant Professor, and became there an Associate Professor in 1991 and then a Professor in 2012. From 1999 to 2002 he was the head of the Department of Information Management at the Graduate School of Management of SPbSU.

He is currently a Professor of the Department of Statistical Modelling at St. Petersburg State University. His research interests include theory and applications of idempotent algebra, modelling and performance evaluation of queueing systems, methods of optimization, computational statistics and computer simulation. He is an author and coauthor of more than 70 publications including papers in reviewed journals and conference proceedings, books chapters, textbooks, and a monograph.

Prof. Krivulin was a grantee of national and international foundations, including the Russian Foundation for Basic Research, the Russian Foundation for Humanities Research, the NATO Science Foundation, the USIA and Eurasia Foundation (USA), and the Royal Society (UK). He served as a member of program and organizing committees of international conferences on mathematics, computer sciences, and information technology. He is a member of the St. Petersburg Mathematical Society, AMS, and SIAM.