A Solution of a Tropical Linear Vector Equation

NIKOLAI KRIVULIN
Faculty of Mathematics and Mechanics
St. Petersburg State University
28 Universitetsky Ave., St. Petersburg, 198504
RUSSIA
nkk@math.spbu.ru

Abstract: A linear vector equation is considered defined in terms of idempotent mathematics. To solve the equation, we apply an approach that is based on the analysis of distances between vectors in idempotent vector spaces and reduces the solution of the equation to that of a tropical optimization problem. Based on the approach, existence and uniqueness conditions are established for the solution, and a general solution to the equation is given.

Key-Words: Idempotent vector space, Linear equation, Tropical optimization problem, Existence condition

1 Introduction

Many applications of tropical mathematics [1–7] involve the solution of linear equations defined on finite-dimensional semimodules over idempotent semifields (idempotent vector spaces). One of the equation that often arise in the applications has 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, efficient solution methods for the equation are of both practical and theoretical interest.

There are several existing solution methods, including those in [1–3]. In this paper another solution approach is described which uses the analysis of distances between vectors in idempotent vector spaces. The results presented are based on implementation and further refinement of solutions first published in the papers [8–10] and not 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.

2 Preliminaries

In this section, we present algebraic definitions, notations, and results based on [8, 10] to provide a background for subsequent analysis and solutions. Additional details and further results are found in [1–7].

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

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 and involves that $x \leq y$ if and only if $x \oplus y = y$. Below, the relation symbols and the operator $\min$ are thought in terms of this linear order.

Note that we have $x \geq \mathbb{0}$ for all $x \in \mathbb{X}$. We also assume that the set $\mathbb{X}$ can always be extended by an element ∞ such that $x \leq \infty$ for all $x \in \mathbb{X}$.

As an example of the semifields under consideration, one can consider a semifield of real numbers $\mathbb{R}_{\max,+} = \langle \mathbb{R} \cup \{-\infty\}, -\infty, 0, \max, + \rangle$.

2.1 Idempotent Vector Space

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

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

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

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

Fig. 1 illustrates the operations in the space $\mathbb{R}^2_{\max,+}$ with the Cartesian coordinates on the plane.
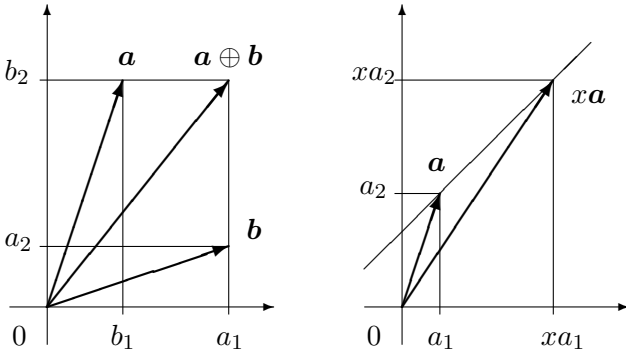

Figure 1: Vector addition (left) and multiplication by scalars (right) in $\mathbb{R}^2_{\max,+}$.

Multiplication of a matrix $A = (a_{ij}) \in \mathbb{X}^{m \times n}$ by a vector $\boldsymbol{x} = (x_i) \in \mathbb{X}^n$ is routinely defined to result in a vector with components

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

All above operations are component-wise isotone in each argument.

A matrix is regular if it has no zero rows.

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 $x_i \neq \mathbb{0}$, and $x_i^- = \mathbb{0}$ otherwise.

For any regular vectors $\boldsymbol{x}$ and $\boldsymbol{y}$, the component-wise inequality $\boldsymbol{x} \leq \boldsymbol{y}$ implies $\boldsymbol{x}^- \geq \boldsymbol{y}^-$.

If $\boldsymbol{x}$ is a regular vector, then $\boldsymbol{x}^-\boldsymbol{x} = \mathbb{1}$ and $\boldsymbol{x}\boldsymbol{x}^- \geq I$, where I is an identity matrix with the elements equal to $\mathbb{1}$ on the diagonal, and $\mathbb{0}$ elsewhere.

2.2 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 $\boldsymbol{b} = x_1\boldsymbol{a}_1 \oplus \cdots \oplus x_n\boldsymbol{a}_n$, where $x_1, \ldots x_n \in \mathbb{X}$.

The set of all linear combinations of $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ forms a linear span denoted by $\mathrm{span}\{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n\}$. A geometrical example of a linear span in $\mathbb{R}^2_{\max,+}$ is given in Fig. 2.
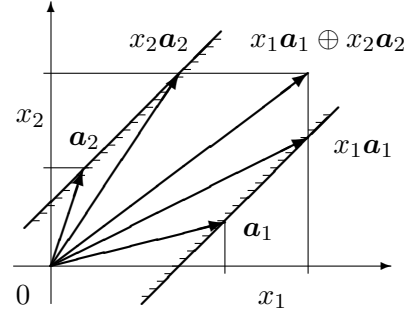
Figure 2: A linear span of vectors $\boldsymbol{a}_1, \boldsymbol{a}_2$ in $\mathbb{R}^2_{\max,+}$.

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 this vector 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 $x_i' \neq x_i$ for some index $i = 1, \ldots, n$.

Assuming for definiteness that $x_i' < x_i$, 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 $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$.

2.3 Distance Function

For any vector $\boldsymbol{a} \in \mathbb{X}^m$, we introduce the support as the index set $\mathrm{supp}(\boldsymbol{a}) = \{i | a_i \neq \mathbb{0}\}$.

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

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

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

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

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

3 Evaluation of Distances

Let $a_1, \ldots, a_n \in \mathbb{X}^m$ be given vectors. Denote by $A = (a_1, \ldots, a_n)$ a matrix composed of the vectors, and by $\mathcal{A} = \mathrm{span}\{a_1, \ldots, a_n\}$ their linear span.

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

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

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

$$\rho(\mathcal{A}, d) = \min_{x \in \mathbb{X}^n} \rho(Ax, d). \tag{2}$$

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

Let some of the vectors $a_1, \ldots, a_n$ 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 $\rho(\mathcal{A}, d) = \mathbb{0}$ if $d = \mathbb{0}$, and $\rho(\mathcal{A}, d) = \infty$ otherwise.

From here on we assume that $d \neq \mathbb{0}$ and $a_i \neq \mathbb{0}$ for all $i = 1, \ldots, n$.

Suppose the vector $d = (d_i)$ 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 d. In the matrix $\widehat{A}$, these columns have all entries that are not located at the intersections set to zero. The matrix $\widehat{A}$ and the vector d are said to be consistent with each other.

Note that when d is regular, we have $\widehat{A} = A$.

Proposition 1. *For all $x = (x_i) \in \mathbb{X}^n$ it holds that*

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

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

Suppose that $\rho(Ax, d) < \infty$, which occurs only under the condition $\mathrm{supp}(Ax) = \mathrm{supp}(d)$. 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 must 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 $\rho(Ax, d) < \infty$ unchanged.

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

With the above result, we may now concentrate only on the problems when A is consistent with d.

In order to describe the solution of problem (2), we need the following notation. For any consistent matrix A and a vector d, we define a residual value

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

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

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

Below we find the solution when the vector d is regular and then extend this result to irregular vectors.

3.1 Regular Vector

Suppose that d is a regular vector. First we verify that the minimum of $\rho(Ax, d)$ over $\mathbb{X}^n$ in (2) can be found by examining only regular vectors $x \in \mathbb{X}_+^n$.

Proposition 2. *If the vector d is regular, then*

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

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

Suppose $y = (y_1, \ldots, y_m)^T$ is regular. Assume a corresponding vector x to have a zero component, say $x_j = \mathbb{0}$. We define the set $I = \{i | a_{ij} > \mathbb{0}\} \neq \emptyset$ and find the 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 y together with the minimum value of $\rho(Ax, d)$ remain unchanged. $\square$

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

Lemma 1. *If the vector d is regular, then it holds that*

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

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

Proof. Let the matrix A be irregular. Then we have $\operatorname{supp}(A\boldsymbol{x}) \neq \operatorname{supp}(\boldsymbol{d})$ and $\rho(\mathcal{A}, \boldsymbol{d}) = \infty$. Since, by definition, $\Delta = \infty$, the statement is true in this case.

Suppose A is regular. Taking into account (1) and (2), 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 $\boldsymbol{x}^-$ gives $r\boldsymbol{x}^- \geq \boldsymbol{d}^- A\boldsymbol{x}\boldsymbol{x}^- \geq \boldsymbol{d}^- A$. Then we obtain $\boldsymbol{x} \leq r(\boldsymbol{d}^- A)^-$ and $(A\boldsymbol{x})^- \geq r^{-1}(A(\boldsymbol{d}^- A)^-)^-$.

Substitution into the second inequality results in $r \geq r^{-1}(A(\boldsymbol{d}^- A)^-)^- \boldsymbol{d} = r^{-1}\Delta^2$, and so in $r \geq \Delta$.

It remains to verify that $r = \Delta$ when we take $\boldsymbol{x} = \Delta(\boldsymbol{d}^- A)^-$. Indeed, substitution of the $\boldsymbol{x}$ 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 $\boldsymbol{y} = \Delta A(\boldsymbol{d}^- A)^- \in \mathcal{A}$. $\square$

Examples of a linear span $\mathcal{A} = \operatorname{span}(\boldsymbol{a}_1, \boldsymbol{a}_2)$ and vectors $\boldsymbol{d}$ in the space $\mathbb{R}^2_{\max,+}$ are given in Fig. 3.
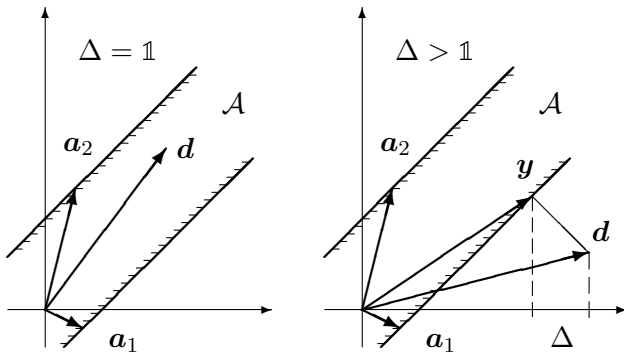
Figure 3: A linear span $\mathcal{A}$ and vectors $\boldsymbol{d}$ in $\mathbb{R}^2_{\max,+}$ when $\Delta = \mathbb{1}$ (left) and $\Delta > \mathbb{1}$ (right).

3.2 Arbitrary Nonzero Vector

Now we examine the distance between the linear span $\mathcal{A}$ and a vector $\boldsymbol{d} \neq \mathbb{0}$. It will suffice to consider only the case when the matrix A is consistent with $\boldsymbol{d}$.

Theorem 1. *Suppose the matrix A is consistent with the vector $\boldsymbol{d} \neq \mathbb{0}$. Then 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. We define sets of indices $I = \{i | d_i = \mathbb{0}\}$ and $J = \{j | a_{ij} > \mathbb{0}, i \in I\}$.

To provide minimum of $\rho(A\boldsymbol{x}, \boldsymbol{d})$, we have to set $x_j = \mathbb{0}$ 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 regular or irregular simultaneously.

Suppose that both matrices are 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 $\boldsymbol{x} = \Delta(\boldsymbol{d}^- A)^-$. $\square$

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

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

It follows from Lemma 1 that $\Delta \geq \mathbb{1}$. The equality $\Delta = \mathbb{1}$ means that the vector $\boldsymbol{d}$ belongs to the linear span $\mathcal{A} = \operatorname{span}\{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n\}$, whereas the inequality $\Delta > \mathbb{1}$ implies that $\boldsymbol{d}$ is outside $\mathcal{A}$. In other words, we have the following statement.

Lemma 2. *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 criterion that a system $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ is linearly independent. We denote by $A_i = (\boldsymbol{a}_1, \ldots, \boldsymbol{a}_{i-1}, \boldsymbol{a}_{i+1}, \ldots, \boldsymbol{a}_n)$ 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 3. *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 Theorem 1 that here none of the vectors $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ is a linear combination of others, and therefore, the system of the vectors is linearly independent. $\square$

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 linearly independent system, we implement a sequential procedure that examines the vectors one by one to decide whether to remove a vector 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 $\widetilde{\boldsymbol{a}}_1, \ldots, \widetilde{\boldsymbol{a}}_k$, 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 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 3, the system $\widetilde{\boldsymbol{a}}_1, \ldots, \widetilde{\boldsymbol{a}}_k$ is linearly independent. $\square$

5 Linear Equations

Given a matrix $A \in \mathbb{X}^{m \times n}$ and a vector $\boldsymbol{d} \in \mathbb{X}^m$, consider the problem of finding an unknown vector $\boldsymbol{x} \in \mathbb{X}^n$ to satisfy the equation

$$A\boldsymbol{x} = \boldsymbol{d}. \tag{3}$$

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

If a matrix $A = (\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n)$ has a zero column, say $\boldsymbol{a}_i$, then the solution of equation (3) reduces to that of an equation that is obtained from (3) 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 (3) to have a set of solutions, where x_i takes all values in $\mathbb{X}$.

Suppose that $A = \mathbb{0}$. In this case, any vector $\boldsymbol{x} \in \mathbb{X}^n$ is a solution provided that $\boldsymbol{d} = \mathbb{0}$, and there is no solution otherwise. If $\boldsymbol{d} = \mathbb{0}$, then equation (3) has 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.

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

5.1 Existence and Uniqueness of Solution

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

Theorem 2. *If a matrix A is consistent with a vector $\boldsymbol{d} \neq \mathbb{0}$, then the following statements are valid:*

(1) *Equation (3) has solutions if and only if $\Delta = \mathbb{1}$.*

(2) *If solvable, the equation has a 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 result from Theorem 1. The uniqueness condition follows from representation of a vector as a unique linear combination of its minimal set of generators. $\square$

We define a pseudo-solution to equation (3) as a vector that satisfies the equation $A\boldsymbol{x} = \Delta A(\boldsymbol{d}^- A)^-$.

Note that the last equation always has a solution that is given by $\boldsymbol{x} = \Delta(\boldsymbol{d}^- A)^-$. Clearly, when $\Delta = \mathbb{1}$ this pseudo-solution becomes an actual solution. Moreover, it follows from Theorem 1 that the pseudo-solution provides the minimum distance to the vector $\boldsymbol{d}$ over all vectors in the linear span of columns of the matrix A in the sense of the metric ρ.

5.2 General Solution

To describe a general solution to equation (3), we first prove an auxiliary result.

Lemma 4. *Suppose that I is a subset of column indices of a matrix A and $d \in \operatorname{span}\{a_i | i \in I\}$.*

Then any vector $x_I = (x_i)$ with components $x_i = (d^- a_i)^-$ if $i \in I$, and $x_i \le (d^- a_i)^-$ otherwise, is a solution to equation (3).

Proof. To verify the statement, we first consider that $d \in \operatorname{span}\{a_i | i \in I\} \subset \operatorname{span}\{a_1, \ldots, a_n\}$ and thus equation (3) has a solution x_I. For the components x_i of the solution, we can write

$$d = \bigoplus_{i=1}^{n} x_i a_i = \bigoplus_{i \in I} x_i a_i \oplus \bigoplus_{i \notin I} x_i a_i.$$

We note that the condition $d \in \operatorname{span}\{a_i | i \in I\}$ is equivalent to an equality

$$d = \bigoplus_{i \in I} x_i a_i,$$

which is valid when $x_i = (d^- a_i)^-$ for all $i \in I$.

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

$$d \ge \bigoplus_{i \notin I} x_i a_i \ge x_i a_i.$$

It remains to solve the inequalities to conclude that for all $i \notin I$, we can take any $x_i \le (d^- a_i)^-$. $\square$

Now suppose that the columns with indices in I form a minimal generating system for d. Denote by $\mathcal{I}$ a set of all such index sets I. Clearly, $\mathcal{I} \ne \emptyset$ only when equation (3) has at least one solution.

By applying Lemma 4, it is not difficult to verify that the following statement holds.

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

$$x_i = (d^- a_i)^-, \quad \textit{if } i \in I,$$
$$x_i \le (d^- a_i)^-, \quad \textit{if } i \notin I.$$

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

Graphical illustration of unique and non-unique solutions to equation (3) are given in Fig. 4.
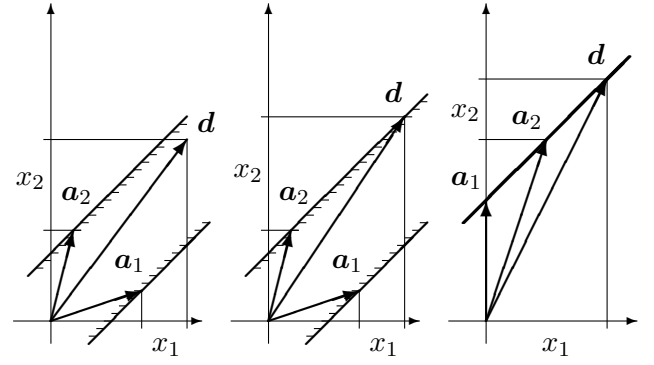

Figure 4: A unique (left) and non-unique (middle and right) solutions to linear equations in $\mathbb{R}^2_{\max,+}$.

References:

[1] N. N. Vorobjev, *Extremal algebra of positive matrices*, Elektronische Informationsverarbeitung und Kybernetik, Vol. 3, No. 1, pp. 39-72. (in Russian)

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

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

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

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

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

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

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

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

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