Available online at www.sciencedirect.com
SCIENCE DIRECT®
LINEAR ALGEBRA AND ITS APPLICATIONS
ELSEVIER Linear Algebra and its Applications 367 (2003) 313–335
www.elsevier.com/locate/laa

Max-algebra: the linear algebra of combinatorics?

Peter Butkovič

School of Mathematics and Statistics, The University of Birmingham, Edgbaston, Birmingham B15 2TT, UK

Received 29 August 2002; accepted 14 October 2002

Submitted by R.A. Brualdi

Abstract

Let $a \oplus b = \max(a, b)$, $a \otimes b = a + b$ for $a, b \in \overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty\}$. By max-algebra we understand the analogue of linear algebra developed for the pair of operations $(\oplus, \otimes)$ extended to matrices and vectors. Max-algebra, which has been studied for more than 40 years, is an attractive way of describing a class of nonlinear problems appearing for instance in machine-scheduling, information technology and discrete-event dynamic systems. This paper focuses on presenting a number of links between basic max-algebraic problems like systems of linear equations, eigenvalue–eigenvector problem, linear independence, regularity and characteristic polynomial on one hand and combinatorial or combinatorial optimisation problems on the other hand. This indicates that max-algebra may be regarded as a linear-algebraic encoding of a class of combinatorial problems. The paper is intended for wider readership including researchers not familiar with max-algebra.

AMS classification: 15A15; 05C50; 90C27

Keywords: Max-algebra; Digraph; Permanent; Permutation

1. Introduction

Let

$$a \oplus b = \max(a, b)$$

and

$$a \otimes b = a + b$$

E-mail address: p.butkovic@bham.ac.uk (P. Butkovič).

0024-3795/03/$ - see front matter © 2003 Elsevier Science Inc. All rights reserved.
doi:10.1016/S0024-3795(02)00655-9

for $a, b \in \overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty\}$. By max-algebra we understand in this paper the analogue of linear algebra developed for the pair of operations $(\oplus, \otimes)$ extended to matrices and vectors formally in the same way as in linear algebra, that is if $A = (a_{ij}),\ B = (b_{ij})$ and $C = (c_{ij})$ are matrices with elements from $\overline{\mathbb{R}}$ of compatible sizes, we write $C = A \oplus B$ if $c_{ij} = a_{ij} \oplus b_{ij}$ for all i, j, $C = A \otimes B$ if $c_{ij} = \sum_k^{\oplus} a_{ik} \otimes b_{kj} = \max_k(a_{ik} + b_{kj})$ for all i, j and $\alpha \otimes A = (\alpha \otimes a_{ij})$ for $\alpha \in \mathbb{R}$.

So $2 \oplus 3 = 3$, $2 \otimes 3 = 5$, $(4, -1) \otimes (3, 6)^{\mathrm{T}} = 7$ and the system

$$\begin{pmatrix} 1 & -3 \\ 5 & 2 \end{pmatrix} \otimes \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 3 \\ 7 \end{pmatrix}$$

in the conventional notation reads

$$\max(1 + x_1, -3 + x_2) = 3,$$
$$\max(5 + x_1, 2 + x_2) = 7.$$

The attractivity of max-algebra is related to the fact that $(\overline{\mathbb{R}}, \oplus, \otimes)$ is a commutative semiring and $(\overline{\mathbb{R}}^n, \oplus, \otimes)$ is a semimodule. Hence many of the basic tools we are used to from linear algebra are available in max-algebra as well.

Max-algebra enables us to describe and study a class of nonlinear problems appearing for instance in machine-scheduling, information technology and discrete-event dynamic systems, by applying the linear-algebraic approach. The following simple example should give an indication how the need for this special transformation arises in real-life: Suppose two trains leave two different stations but arrive at the same station from which a third train, connecting to the first two, departs. Let us denote the departure times of the trains as x_1, x_2 and x_3, respectively, and the duration of the journeys of the first two trains (including the necessary times for changing the trains) by a_1 and a_2, respectively. Then

$$x_3 = \max(x_1 + a_1, x_2 + a_2)$$

which in the max-algebraic notation reads

$$x_3 = x_1 \otimes a_1 \oplus x_2 \otimes a_2$$

thus a max-algebraic scalar product of the vectors (x_1, x_2) and (a_1, a_2). So if the departure time of the third train is given say, as a constant b, and we need to find the necessary departure times of the first two trains, we have to solve a simple max-algebraic linear equation

$$x_1 \otimes a_1 \oplus x_2 \otimes a_2 = b.$$

This paper aims at presenting an overview of results which demonstrate strong links between basic max-algebraic problems and combinatorial or combinatorial optimization problems. These links are so significant that max-algebra may be regarded as a linear-algebraic encoding of a class of combinatorial problems. Instances of such combinatorial problems are: the set covering problem (which in max-algebra is the solvability problem of a linear system), the minimal set covering problem (unique solvability of a linear system), existence of a directed cycle (strong regularity of a matrix), existence of an even directed cycle (regularity of a matrix), maximal cycle mean (eigenvalue), longest-distances (eigenvectors), best principal submatrices (coefficients of a characteristic polynomial). Some results are related to matrix scaling and enable us to formulate a link between combinatorial problems so different as the assignment problem and the longest-distances problem.

All results mentioned in this paper remain valid if $\otimes$ is defined as conventional multiplication and a, b are restricted to non-negative reals. This is especially important for the application to matrix scaling. Almost all (but not all) results straightforwardly generalize to the case when the underlying algebraic structure is a linearly ordered commutative group in which $\otimes$ is the group operation. So our previous two definitions of max-algebra are derived from the additive group of reals (denoted by $\mathscr{G}_0$) and multiplicative group of positive reals ($\mathscr{G}_1$), respectively. Unless stated otherwise everywhere in this paper we deal with max-algebra based on $\mathscr{G}_0$ which is called the *principal interpretation* [13].

It should also be noted that if $\oplus$ is set to the minimum rather than to the maximum then we speak about the min-algebra in which all the results remain valid after appropriately amending the formulation, for instance after replacing $\leqslant$ with $\geqslant$, "positive" with "negative", "longest" with "shortest", etc.

It is beyond the scope of this paper to provide a comprehensive survey of all major existing results and works. But the following may help a reader who would like to know more about max-algebra: Although first papers related to max-algebra appeared as soon as in 1950s or even before, the first comprehensive monograph seems to be [13] which was later extended and updated [15]. Algebraic aspects are comprehensively studied in [27]. Other recommended sources are: monograph [1] with many applications and various different approaches, papers [19,21] and thesis [17]. A summer school's proceedings [18] is a collection of a number of the state-of-the-art papers including a thorough historical overview of the field. A chapter on max-algebra can be found in [2].

Throughout the paper we shall use the following standard notation:

$$\operatorname{diag}(a, b, c, \ldots) = \begin{pmatrix} a & & & & \\ & b & & -\infty & \\ & & c & & \\ & -\infty & & \ddots & \\ & & & & \ddots \end{pmatrix},$$

$$I = \operatorname{diag}(0, 0, \ldots).$$

Obviously

$$A \otimes I = I \otimes A = A$$

for any matrices A and I of compatible sizes. The matrix $\mathrm{diag}(a, b, c, \ldots)$ will be called *diagonal* if all of $a, b, c, \ldots$ are real numbers.

Any matrix which can be obtained from I by permuting the rows and/or columns will be called a *permutation matrix*. A matrix arising as a product of a diagonal matrix and a permutation matrix will be called a *generalized permutation matrix*.

If A is a square matrix then the iterated product $A \otimes A \otimes \cdots \otimes A$ in which the letter A stands k-times will be denoted as $A^{(k)}$. The symbol $a^{(k)}$ applies similarly to scalars, thus $a^{(k)}$ is simply ka. This definition immediately extends to $a^{(x)} = xa$ for any x real and in particular $a^{(-1)}$ denotes what in the conventional notation is $-a$. As an exception and for simplicity of notation the last symbol will be written simply a^{-1} as the conventional inverse will not be used in this paper. The symbol 0 will denote both the number and the vector whose every component is zero. Throughout the paper the letters M and N will stand for the sets $\{1, \ldots, m\}$ and $\{1, \ldots, n\}$, respectively. Most of the results will be shown for matrices with real (finite) entries as these demonstrate the aim of the paper most clearly.

If $A = (a_{ij}) \in \bar{\mathbb{R}}^{n\times n}$ then the symbols $F_A[Z_A]$ will stand for the digraphs with the node set N and arc set

$$E = \{(i, j); a_{ij} > -\infty\} \quad [E = \{(i, j); a_{ij} = 0\}].$$

Z_A will be called the *zero digraph* of the matrix A. D_A will denote the arc-weighted digraph arising from F_A by assigning the weight a_{ij} to arc (i, j) for all $i, j \in N$.

Conversely, if $D = (N, E)$ is an arc-weighted digraph with weight function $w : E \to \mathbb{R}$ then A_D will denote the matrix $A = (a_{ij}) \in \bar{\mathbb{R}}^{n\times n}$ defined by

$$a_{ij} = \begin{cases} w(i, j) & \text{if } (i, j) \in E \\ -\infty & \text{else} \end{cases} \quad \text{for all } i, j \in N.$$

A_D will be called the *direct-distances matrix* of the digraph D.

2. Linear system is set covering

If $A = (a_{ij}) \in \mathbb{R}^{m\times n}$ and $b = (b_1, \ldots, b_m)^{\mathrm{T}} \in \mathbb{R}^m$ then the max-algebraic linear system ("max-algebraic" will usually be omitted in the rest of the paper)

$$A \otimes x = b$$

written in the conventional notation is the nonlinear system

$$\max_{j=1,\ldots,n} (a_{ij} + x_j) = b_i \quad (i = 1, \ldots, m).$$

By subtracting the right-hand side values we get

$$\max_{j=1,\ldots,n} (a_{ij} - b_i + x_j) = 0 \quad (i = 1, \ldots, m).$$

A linear system whose all right hand-side constants are zero will be called *normalized* and the above process will be called *normalization*. So for instance the system

$$\begin{pmatrix} -2 & 2 & 2 \\ -5 & -3 & -2 \\ -5 & -3 & 3 \\ -3 & -3 & 2 \\ 1 & 4 & 6 \end{pmatrix} \otimes \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 3 \\ -2 \\ 1 \\ 0 \\ 5 \end{pmatrix}$$

after normalization becomes

$$\begin{pmatrix} -5 & -1 & -1 \\ -3 & -1 & 0 \\ -6 & -4 & 2 \\ -3 & -3 & 2 \\ -4 & -1 & 1 \end{pmatrix} \otimes \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Normalization is of course nothing else than multiplying the system by the matrix

$$B = \operatorname{diag}\left(b_1^{-1}, b_2^{-1}, \ldots, b_m^{-1}\right)$$

from the left, that is

$$B \otimes A \otimes x = B \otimes b = 0.$$

Notice that the first equation of the normalized system above reads

$$\max(x_1 - 5, x_2 - 1, x_3 - 1) = 0.$$

So, if $x = (x_1, x_2, x_3)^{\mathrm{T}}$ is a solution to this system then $x_1 \leqslant 5, x_2 \leqslant 1, x_3 \leqslant 1$ and at least one of these inequalities will be satisfied with equality. For x_1 we then get from the other equations: $x_1 \leqslant 3, x_1 \leqslant 6, x_1 \leqslant 3, x_1 \leqslant 4$, thus $x_1 \leqslant \min(5, 3, 6, 3, 4) = -\max(-5, -3, -6, -3, -4) = \bar{x}_1$ where $-\bar{x}_1$ is the column 1 maximum. Clearly for all j then $x_j \leqslant \bar{x}_j$ where $-\bar{x}_j$ is the column j maximum. On the other hand equality must be attained in some of these inequalities so that in every row there is at least one column maximum which is attained by x_j. This was observed in 1970s [26] and is precisely formulated in the theorem below in which it is assumed that we study a system

$$A \otimes x = 0,$$

where $A = (a_{ij}) \in \mathbb{R}^{m \times n}$, and we denote

$$S = \left\{x \in \mathbb{R}^n;\ A \otimes x = 0\right\},$$

$$\bar{x}_j = -\max_i a_{ij} \quad \text{for all } j \in N,$$

$$\bar{x} = (\bar{x}_1, \ldots, \bar{x}_n)^{\mathrm{T}},$$

$$M_j = \left\{k \in M;\ a_{kj} = \max_i a_{ij}\right\} \quad \text{for all } j \in N.$$

Theorem 2.1 [26]. *$x \in S$ if and only if*

(a) $x \leqslant \bar{x}$ *and*
(b) $\bigcup_{j \in N_x} M_j = M$, *where* $N_x = \{j \in N; x_j = \bar{x}_j\}$.

It follows immediately that $A \otimes x = 0$ has no solution if $\bar{x}$ is not a solution. Therefore $\bar{x}$ is called the *principal solution* [13]. More precisely we have

Corollary 2.1. *The following three statements are equivalent*:

(a) $S \neq \emptyset$,
(b) $\bar{x} \in S$,
(c) $\bigcup_{j \in N} M_j = M$.

Corollary 2.2. $S = \{\bar{x}\}$ *if and only if*

(i) $\bigcup_{j \in N} M_j = M$ *and*
(ii) $\bigcup_{j \in N'} M_j \neq M$ *for any* $N' \subseteq N$, $N' \neq N$.

Let us consider the following problems:

[UNIQUE] SOLVABILITY. Given $A = (a_{ij}) \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$ does the system $A \otimes x = b$ have a [unique] solution?

[MINIMAL] SET COVERING. Given a finite set M and subsets $M_1, \ldots, M_n$ of M, is

$$\bigcup_{j=1}^{n} M_j = M$$

[is

$$\bigcup_{j=1}^{n} M_j = M \quad \text{but} \quad \bigcup_{\substack{j=1 \\ j \neq k}}^{n} M_j \neq M$$

for any $k \in \{1, \ldots, n\}$]?

Corollaries 2.1 and 2.2 show that for every linear system it is possible to straightforwardly find a finite set and a collection of its subsets so that SOLVABILITY is equivalent to SET COVERING and UNIQUE SOLVABILITY is equivalent to MINIMAL SET COVERING.

However, this correspondence is two-way, as the statement below suggests. Let us assume without loss of generality that the given finite set is $M = \{1, \ldots, m\}$ and its subsets are $M_1, \ldots, M_n$. Define $A = (a_{ij}) \in \mathbb{R}^{m \times n}$ as follows:

$$a_{ij} = \begin{cases} 1 & \text{if } i \in M_j \\ 0 & \text{else} \end{cases} \quad \text{for all } i \in M,\ j \in N,$$

$$b = 0.$$

The following are easily proved.

Theorem 2.2. $\bigcup_{j\in N} M_j = M$ *if and only if* $(\exists x) A \otimes x = b$.

Theorem 2.3. $\bigcup_{j\in N} M_j = M$ *and* $\bigcup_{j\in N'} M_j \neq M$ *for any* $N' \subseteq N, N' \neq N$ *if and only if* $(\exists ! x) A \otimes x = b$.

This section has demonstrated that every linear system is an algebraic representation of a set covering problem (as formulated above), and conversely.

3. Permanent is the optimal assignment problem value

Let $A = (a_{ij}) \in \mathbb{R}^{n\times n}$ and denote by P_n the set of all permutations of the set N. Then the max-algebraic permanent of A can straightforwardly be defined as

$$\mathrm{maper}(A) = \bigoplus_{\pi\in P_n} \bigotimes_{i\in N} a_{i,\pi(i)}$$

which in the conventional notation reads

$$\mathrm{maper}(A) = \max_{\pi\in P_n} \sum_{i\in N} a_{i,\pi(i)}.$$

For $\pi \in P_n$ the quantity

$$w(A, \pi) = \bigotimes_{i\in N} a_{i,\pi(i)} = \sum_{i\in N} a_{i,\pi(i)}$$

is called the *weight* of the permutation π. The problem of finding a permutation $\pi \in P_n$ of maximum weight (called *optimal permutation* or *optimal solution*) is the well known *assignment problem* for the matrix A (see for instance [22] or other textbooks on combinatorial optimization). There are a number of efficient solution methods [5] for finding an optimal solution, one of the best known is the Hungarian method of computational complexity $O(n^3)$. The set of all optimal permutations will be denoted by $\mathrm{ap}(A)$, that is

$$\mathrm{ap}(A) = \left\{ \pi \in P_n;\ \mathrm{maper}(A) = \sum_{i\in N} a_{i,\pi(i)} \right\}.$$

If $A \in \mathbb{R}^{n\times n}$ is a zero–one matrix then the maximum number of zeros in this matrix such that no two are from the same row or column is called the *term rank* of A [3]. It is easily seen that this value is equal to $\mathrm{maper}(\bar{A})$ where $\bar{A}$ arises from A by replacing zeros by ones and vice versa.

A matrix $A \in \mathbb{R}^{n\times n}$ will be called *diagonally dominant* if $\mathrm{id} \in \mathrm{ap}(A)$. A diagonally dominant matrix is called *definite* if all diagonal elements are zero. Thus there are no positive cycles in D_A if A is definite (since any positive cycle could be extended using the complementary diagonal zeros to a permutation of N of positive weight).

A non-positive matrix with zero diagonal is called *normal* (thus every normal matrix is definite). A normal matrix whose all off-diagonal elements are negative is called *strictly normal*. If $|\text{ap}(A)| = 1$ then A is said to have *strong permanent.* Obviously, a strictly normal matrix has strong permanent.

The permanent plays a key role in many max-algebraic problems because of the absence of the determinant due to the lack of subtraction. It turns out that the structure of the optimal solution set $\text{ap}(A)$ is related to some max-algebraic properties and especially to questions such as regularity of matrices.

If a constant c is added to a row or column of A then for any permutation $\pi \in P_n$ exactly one term in $w(A, \pi)$ will change by c and therefore $w(A, \pi)$ will change by this constant and thus $\text{ap}(A)$ remains unchanged. This very simple but crucial idea is a basic principle of the Hungarian method which transforms any matrix A to a non-positive matrix B in which at least one permutation has zero weight and therefore $\text{ap}(A) = \{\pi \in P_n; w(B, \pi) = 0\}$. Since the term rank of B is n, by a suitable permutation of the rows (or columns) of B we obtain a normal matrix. Although this final step is usually not part of the Hungarian method we will assume here for simplicity that this is the case. This enables us to state the following theorem which plays a fundamental role in combinatorial matrix theory and which is a direct consequence of the Hungarian method. Here and in the rest of the paper we shall say that two matrices are *similar* if one can be obtained from the other by adding constants to the rows and columns and by permuting the rows and columns. Similarity is clearly an equivalence relation.

Theorem 3.1. *Every real square matrix is similar to a normal matrix.*

One can easily see that adding constants $c_1, \ldots, c_n$ to the rows of a matrix A has the same effect as the product $C \otimes A$ where $C = \text{diag}(c_1, \ldots, c_n)$. Similarly adding $d_1, \ldots, d_n$ to the columns produces the matrix $A \otimes D$ where $D = \text{diag}(d_1, \ldots, d_n)$. Permutation of the rows [columns] of A can be described by the max-algebraic multiplication of A from the left [from the right] by a permutation matrix. Hence we have:

Proposition 3.1. *Two matrices A, B are similar if and only if $B = P \otimes A \otimes Q$ where P and Q are generalized permutation matrices.*

The above mentioned basic principle of the Hungarian method can be expressed as follows:

Proposition 3.2. *Let A, B be two $n \times n$ matrices.*

(a) *If $B = C \otimes A \otimes D$ where C and D are diagonal matrices then* $\text{ap}(A) = \text{ap}(B)$.
(b) *If $B = P \otimes A \otimes Q$ where P and Q are generalized permutation matrices then* $|\text{ap}(A)| = |\text{ap}(B)|$.

An alternative version of Theorem 3.1 can now be formulated as follows:

Theorem 3.2. *For every real square matrix A there exist generalized permutation matrices P and Q such that $P \otimes A \otimes Q$ is normal.*

Notes

- A normal matrix similar to a matrix A may not be unique. Any such matrix will be called a *normal form* of A.
- Not every square matrix is similar to a strictly normal one (for instance constant matrices). This question is related to so called strong regularity of matrices in max-algebra which is discussed in the following section.
- The present section is closely related to matrix scaling as all discussion of this section could straightforwardly be repeated in max-algebra based on $\mathscr{G}_1$, that is on the multiplicative group of positive reals. Similarity of two positive matrices here would mean that one of the matrices can be obtained from the other by multiplying the rows and columns by positive constants and by permuting the rows and columns. A normal matrix is a positive matrix whose all entries are 1 or less and all diagonal entries are 1.

4. Strong regularity and cycles; regularity and even cycles

In this section we denote

$$S(A, b) = \{x \in \mathbb{R}^n;\ A \otimes x = b\} \quad \text{for } A \in \mathbb{R}^{m\times n} \text{ and } b \in \mathbb{R}^m,$$
$$T(A) = \{|S(A, b)|\,;\ b \in \mathbb{R}^m\} \quad \text{for } A \in \mathbb{R}^{m\times n}.$$

Regularity and linear independence are closely related to the number of solutions of linear systems. One good thing is that like in the conventional case the number of solutions to a linear system can only be 0, 1 or ∞:

Theorem 4.1 [6]. $|S(A, b)| \in \{0, 1, \infty\}$ *for any* $A \in \mathbb{R}^{m\times n}$ *and* $b \in \mathbb{R}^m$.

On the other hand even if a system $A \otimes x = b$ is uniquely solvable it is always possible to find vectors b so that the system has no solution and so that it has an infinite number of solutions (without changing A). However, the structure of $T(A)$ is simpler than in the conventional case and every matrix belongs to one of two classes:

Theorem 4.2 [6]. $T(A)$ *is either* $\{0, \infty\}$ *or* $\{0, 1, \infty\}$ *for any* $A \in \mathbb{R}^{m\times n}$.

We say that the columns of A are *strongly linearly independent* if $1 \in T(A)$. A square matrix with strongly linearly independent columns is called *strongly regular*.

It follows from Corollary 2.2 that if a normalized $m \times n$ system has a unique solution then for every $j \in N$ there is at least one $i \in M_j$ such that $i \notin M_k$ for all $k \neq j$. Let us denote this index i by i_j (take any in the case of a tie). Consider the subsystem with row indices $i_1, i_2, \ldots, i_n$ (and with all columns). This is an $n \times n$ system with unique column maximum in every column and in every row. Hence again by Corollary 2.2 this system has a unique solution and so we have:

Proposition 4.1. *Every normalized $m \times n$ system which has a unique solution contains an $n \times n$ subsystem which has a unique solution.*

This statement has already been known for some time in a stronger form [10]:

Theorem 4.3. *A matrix $A \in \mathbb{R}^{m\times n}$ has strongly linearly independent columns if and only if it contains a strongly regular $n \times n$ submatrix.*

Corollary 4.1. *If a matrix $A \in \mathbb{R}^{m\times n}$ has strongly linearly independent columns then $m \geqslant n$.*

Unique column maximum in every column and in every row is a property which characterizes every normalized uniquely solvable square system. To see this just realize that we require the sets $M_1, \ldots, M_n$ to form a minimal covering of the set $N = \{1, \ldots, n\}$. It is easily verified that this is only possible if all the sets are one-element and pairwise-disjoint.

If a square matrix has a unique column maximum in every column and in every row then the column maxima determine a permutation of the set N whose weight is strictly greater than the weight of any other permutation and thus this matrix has strong permanent. In other words, if A is a square matrix and $A \otimes x = 0$ has a unique solution then A has strong permanent. However, we know from Section 2 that normalization of a system $A \otimes x = b$ means to multiply A by a diagonal matrix from the left. By Proposition 3.2(a) this does not effect $\mathrm{ap}(A)$ and hence we have proved:

Proposition 4.2. *If A is strongly regular then A has strong permanent.*

The converse is also true. This statement has been known for some time [10] (note that the original long proof has now been substantially shortened [8]):

Theorem 4.4. *A square matrix is strongly regular if and only if it has strong permanent.*

This result leads naturally to the idea of checking the strong regularity of a matrix by checking whether this matrix has strong permanent. We now briefly discuss this question.

Let $A = (a_{ij}) \in \mathbb{R}^{n\times n}$. Due to the Hungarian method we can easily find a normal matrix B similar to A. By Corollary 3.2(b) B has strong permanent if and only if A has this property. Every permutation is a product of cycles, therefore if $w(B, \pi) = 0$ for some $\pi \neq \mathrm{id}$ then at least one of the constituent cycles of π is of length two or more or, equivalently there is a cycle of the length two or more in the digraph Z_B. Conversely, every such cycle can be extended using the complementary diagonal zeros in B to a permutation of zero weight with respect to B, different from id. Thus we have:

Theorem 4.5 [6]. *A square matrix is strongly regular if and only if the zero digraph of any* (*and thus of all*) *of its normal form is acyclic.*

We observe here that the transformation in the opposite direction is also possible: Given a digraph $D = (N, E)$ we may assume without loss of generality that D has no loops. Set $A = (a_{ij}) \in \mathbb{R}^{n\times n}$ where

$$a_{ij} = \begin{cases} 0 & \text{if } (i, j) \in E \text{ or } i = j \\ -1 & \text{else} \end{cases} \quad \text{for all } i, j \in N.$$

Then $\mathrm{maper}(A) = 0$, $\mathrm{id} \in \mathrm{ap}(A)$ and $|\mathrm{ap}(A)| = 1$ if and only if D is acyclic. Hence D is acyclic if and only if A is strongly regular.

It should be noted that an early paper [16] on matrix scaling contains results which are closely related to Theorem 4.5.

We conclude that strong regularity is an algebraic way of describing acyclic digraphs.

There are several ways of defining linear independence and regularity which are non-equivalent in max-algebra although they would be equivalent in linear algebra. In contrast to the concepts defined earlier in this section we now define (max-algebraic) linear independence and regularity. Let $A_1, \ldots, A_n$ be the columns of the matrix $A = (a_{ij}) \in \mathbb{R}^{m\times n}$. We say that $A_1, \ldots, A_n$ are *linearly dependent* if

$$\sum_{j\in S}^{\oplus} \lambda_j \otimes A_j = \sum_{j\in T}^{\oplus} \lambda_j \otimes A_j$$

holds for some real numbers λ_j and two non-empty, disjoint subsets S and T of the set N. If such a representation is impossible, then the columns of A are called *linearly independent*. A square matrix with linearly independent columns is called *regular*.

To formulate a regularity criterion we need to introduce some more symbols: P_n^+ [P_n^-] is the set of even [odd] permutations of the set N.

Theorem 4.6 (Gondran–Minoux [20]). *A square matrix A of order n is regular if and only if* $\mathrm{ap}(A) \subseteq P_n^+$ *or* $\mathrm{ap}(A) \subseteq P_n^-$.

Corollary 4.2. *Every strongly regular matrix is regular.*

It follows straight from the definition of regularity that this property is not affected by the similarity transformations. Hence a matrix A is regular if and only if its normal form B is regular. Since id $\in$ ap(B) and id is even, the regularity of the normal matrix $B = (b_{ij})$ is equivalent to the non-existence of an odd permutation $\sigma \in P_n$ such that $b_{i,\sigma(i)} = 0$ for all $i \in N$. If the parity of a permutation σ is odd then the parity of at least one of its constituent cycles is odd. Since the parity of a cyclic permutation is odd if and only if this cycle is of an even length, we have that a necessary condition for a permutation σ to be odd is that at least one of its constituent cycles be even or, equivalently there is an even cycle in Z_B. On the other hand, obviously, if Z_B contains an even cycle then this cycle can be extended by the complementary diagonal zeros to an odd optimal permutation. We have proved:

Corollary 4.3 [7]. *Let $A \in \mathbb{R}^{n\times n}$ and let B be any normal form of A. Then A is regular if and only if Z_B does not contain an even cycle.*

Note that the question of checking the existence of an even cycle in a digraph is a hard combinatorial question and its polynomial solvability was open for many years until 1999 when it was affirmatively answered [23].

It is not clear whether an analogue of Theorem 4.3 is true for linear independence but we can prove an analogue of its corollary, Corollary 4.1:

Theorem 4.7. *If a matrix $A \in \mathbb{R}^{m\times n}$ has linearly independent columns then $m \geqslant n$.*

Proof. Let $A = (a_{ij}) \in \mathbb{R}^{m\times n}$ and $m < n$. We shall show that A has linearly dependent columns.

Since the linear independence of columns is not affected by $\otimes$ multiplying the columns by constants, we may assume without loss of generality that the last row of A is zero. Let B be an $m \times m$ submatrix of A with maximum maper(B). We may also assume that B consists of the first m columns of A and that id $\in$ ap(B) (if necessary, we appropriately permute the columns of A). Let C be the $n \times n$ matrix arising by adding $n - m$ zero rows to A. Then clearly maper(C) = maper(B) and ap(C) contains any permutation that is an extension of id from ap(B) to a permutation of N. As A already had one zero row and we have added at least another one, C has at least two zero rows, thus ap(C) contains at least one pair of permutations of different parities (see Fig. 1).

Hence, by Theorem 4.6 C is not regular and if we denote the columns of C by $C_1, \ldots, C_n$ then

$$\sum_{j\in S}^{\oplus} \lambda_j \otimes C_j = \sum_{j\in T}^{\oplus} \lambda_j \otimes C_j$$

holds for some real numbers λ_j and two non-empty, disjoint subsets S and T of the set N. This vector equality restricted to the first m components then yields the linear dependence of the columns of A. $\square$

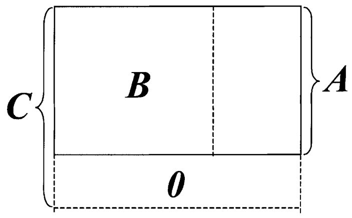

Fig. 1.

Again the conversion in the other direction is possible too: Given a digraph $D = (N, E)$ we may assume without loss of generality that D has no loops (as they cannot be part of any even cycle). Set $A = (a_{ij}) \in \mathbb{R}^{n\times n}$ where

$$a_{ij} = \begin{cases} 0 & \text{if } (i, j) \in E \text{ or } i = j \\ -1 & \text{else} \end{cases} \quad \text{for all } i, j \in N.$$

Then $\text{maper}(A) = 0$, $\text{id} \in \text{ap}(A)$ and $\text{ap}(A) \subseteq P_n^+$ if and only if there is no even cycle in D. Hence D contains no even cycles if and only if A is regular.

We conclude that regularity is an algebraic representation of digraphs without even cycles.

5. Regularity and sign-nonsingularity

A square $(0, 1, -1)$ matrix is called *sign-nonsingular* (abbr. SNS) [3,25] if at least one term in the standard determinant expansion of A is non-zero and all such terms have the same sign. For instance

$$\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$$

is SNS but

$$\begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

are not. Obviously SNS is not affected by permutations of the rows and columns.

There is a direct connection between (max-algebraic) regularity and SNS. Before we can see this we need to mention a remarkable combinatorial property of square matrices which readily follows from the Hungarian method (and which may not be immediately evident without it):

Theorem 5.1. *Let* $A = (a_{ij}) \in \mathbb{R}^{n\times n}$ *and* $\pi \in P_n$. *If for every* $i \in N$ *there is a* $\sigma \in \text{ap}(A)$ *such that* $\pi(i) = \sigma(i)$ *then* $\pi \in \text{ap}(A)$.

Proof. Without loss of generality $\text{id} \in \text{ap}(A)$. Every matrix is similar to a matrix in normal form, thus A is similar to a normal matrix $B = P \otimes A \otimes Q$ where P and Q

are diagonal. Then by Corollary 3.2(a) $\mathrm{ap}(A) = \mathrm{ap}(B)$. Since for every $i \in N$ there is a $\sigma \in \mathrm{ap}(B)$ such that $b_{i,\pi(i)} = b_{i,\sigma(i)} = 0$ we have that $\pi \in \mathrm{ap}(B)$. $\square$

Theorem 5.2. *Let $A = (a_{ij}) \in \mathbb{R}^{n\times n}$ and $\bar{A} = (\bar{a}_{ij})$ be the $n \times n$ zero–one matrix defined as follows*:

$$\bar{a}_{ij} = \begin{cases} 1 & \text{if } j = \pi(i) \text{ for some } \pi \in \mathrm{ap}(A) \\ 0 & \text{else} \end{cases} \quad \text{for all } i, j \in N.$$

Then A is regular if and only if $\bar{A}$ is SNS.

Proof. First note that due to Theorem 5.1 non-zero terms in the standard determinant expansion of $\bar{A}$ exactly correspond to permutations from $\mathrm{ap}(A)$ and at least one such always exists since $\mathrm{ap}(A) \neq \emptyset$ for every matrix A. The sign of every non-zero term is fully determined by the parity of the permutation as the non-zero entries of $\bar{A}$ are all one and so the theorem statement follows. $\square$

Conversely, the question whether a square zero–one matrix $B = (b_{ij})$ is SNS can be converted to the regularity question for the matrix $C = (c_{ij})$ defined as

$$c_{ij} = \begin{cases} 0 & \text{if } b_{ij} = 1 \\ -1 & \text{if } b_{ij} = 0 \end{cases} \quad \text{for all } i, j \in N.$$

Namely, B can be assumed to have at least one non-zero term in its determinant expansion (otherwise there is nothing to solve), thus $\mathrm{maper}(C) = 0$ and there is a one-to-one correspondence between permutations of zero weight w.r.t. C (that is permutations in $\mathrm{ap}(C)$) and non-zero terms in $\det(B)$. Since B is zero–one, the sign of every non-zero term is fully determined by the parity of the permutation and hence regularity of C means the same as SNS of B.

We conclude that regularity is an alternative algebraic description of zero–one sign-nonsingularity.

6. The eigenproblem and graph balancing

The eigenvalue–eigenvector problem (shortly: the *eigenproblem*)

$$A \otimes x = \lambda \otimes x \tag{6.1}$$

for $A = (a_{ij}) \in \mathbb{R}^{n\times n}$ was one of the first problems studied in max-algebra. Here we only discuss the case when A does not contain $-\infty$ as this case most visibly exposes combinatorial features of the eigenproblem. In such a case the situation is transparent—every matrix has exactly one eigenvalue:

Theorem 6.1 (Cuninghame-Green [13]). *For every $A = (a_{ij}) \in \mathbb{R}^{n\times n}$ there is a unique value of $\lambda = \lambda(A)$ (called the eigenvalue of A) to which there is an $x \in \mathbb{R}^n$ satisfying* (6.1). *The unique eigenvalue is the maximum cycle mean in D_A that is*

$$\lambda(A) = \max_{\sigma} \frac{\omega(A, \sigma)}{l(\sigma)} \tag{6.2}$$

where $\sigma = (i_1, \ldots, i_k)$ *denotes an elementary cycle* (*that is a cycle with no repeated node except the first and last one*) *in* D_A, $\omega(A, \sigma) = a_{i_1 i_2} + \cdots + a_{i_k i_1}$ *is the weight of* σ *and* $l(\sigma) = k$ *is the length of* σ. *The maximization is taken over elementary cycles of all lengths in* D_A, *including the loops.*

Note that (6.2) in the max-algebraic notation reads

$$\lambda(A) = \bigoplus_{\sigma} \sqrt[l(\sigma)]{\omega(A, \sigma)}. \tag{6.3}$$

It is beyond the aim and scope of the present paper to discuss numerous generalizations of Theorem 6.1 not only to the cases when $-\infty$ are among the entries of A but also to various algebraic structures from which the entries of A are taken, and infinite-dimensional generalizations. We also do not deal here with algorithms for finding the eigenvalue (let us mention at least Karp's $O(n^3)$ algorithm [15] and Howard's algorithm [12] of unproved computational complexity showing excellent algorithmic performance). It follows from (6.3) that the existence of roots with respect to $\otimes$ is essential (cf. "radicability" [13]) and therefore the statement of Theorem 6.1 does not hold for instance in the max-algebra based on the additive group of integers or multiplicative group of positive rationals.

If $A = (a_{ij}) \in \mathbb{R}^{n \times n}$ then the matrix $\Gamma(A) = A \oplus A^{(2)} \oplus \cdots \oplus A^{(n)}$ is called the *metric matrix* associated with A. Note that $I \oplus \Gamma(A)$ describes the transitive closure of a digraph $D = (N, E)$ if A is defined as follows:

$$a_{ij} = \begin{cases} 0 & \text{if } (i, j) \in E \\ -1 & \text{else} \end{cases} \quad \text{for all } i, j \in N.$$

Any vector x satisfying (6.1) is called an *eigenvector*, the set of all eigenvectors will be denoted $\mathrm{sp}(A)$ and called the *eigenspace* of A.

Theorem 6.2 ([13]). *Let* $A = (a_{ij}) \in \mathbb{R}^{n \times n}$ *and suppose that* $\Gamma(\lambda^{-1} \otimes A) = (g_{ij})$ *has columns* $g_1, g_2, \ldots, g_n$. *Then*

$$\mathrm{sp}(A) = \left\{ \bigoplus_{j \in N_0} \alpha_j \otimes g_j; \alpha_j \in \mathbb{R} \right\}, \tag{6.4}$$

where $N_0 = \{j \in N; g_{jj} = 0\}$.

Since the vectors g_j, $j \in N_0$, are the generators of the eigenspace they are called the *fundamental eigenvectors* [13]. Note that it is not necessary to include all fundamental eigenvectors in (6.4) and it would be sufficient to keep one representative for each of the critical cycles from the computation of the eigenvalue [13].

There are several combinatorial interpretations of the eigenproblem. One has already appeared in Theorem 6.1: the eigenvalue has a very clear combinatorial meaning—it is the biggest of all cycle means. Another one is associated with graph balancing [24]. One of the questions which are studied in graph balancing is: Given an arc-weighted digraph D, amend the arc weights so that the cycle weights remain unchanged but the maximum arc weight of arcs leaving a node is the same for every node. Since the transformation of arc-weights $a_{ij} \to a_{ij} + x_j - x_i$ does not change cycle weights, we may be looking for values $x_1, \ldots, x_n$ such that $\max_{j=1,\ldots,n}(a_{ij} + x_j - x_i)$ is a constant independent of i. If we denote this constant by λ then we have $\max_{j=1,\ldots,n}(a_{ij} + x_j - x_i) = \lambda$ for all $i \in N$ or, equivalently

$$\max_{j=1,\ldots,n} (a_{ij} + x_j) = \lambda + x_i \quad \text{for all } i \in N$$

which in the max-algebraic notation is (6.1). Hence one combinatorial interpretation of the eigenproblem is graph balancing.

7. The eigenproblem and the longest distances problem

One of the most classical combinatorial optimization problems is: Given an $n \times n$ matrix A of direct distances between n places, find the matrix A^* of shortest distances (that is the matrix of the lengths of shortest paths between any pair of places). We may assume without loss of generality that all diagonal elements of A are 0. It is known that the shortest-distances matrix exists if and only if there are no negative cycles in D_A in which case $A^* = A^{(n-1)}$. This is perhaps a min-algebraic property best known outside min-algebra.

To remain consistent with the rest of the paper we shall formulate this very basic but important result in max-algebraic terms. It follows from the definition of definite matrices (see Section 3) that all diagonal entries of a definite matrix A are 0 and all cycles in D_A have non-positive weights. Hence, $\lambda(A) = 0$ and thus eigenvectors are exactly the fixed points of the mapping $x \mapsto A \otimes x$. Since $I \leqslant A$, we also have

$$I \leqslant A \leqslant A^{(2)} \leqslant \cdots$$

and due to the absence of positive cycles

$$A^{(n-1)} = A^{(n)} = A^{(n+1)} = \cdots \tag{7.1}$$

which altogether yield $\Gamma(A) = A^{(n-1)}$. We have:

Theorem 7.1. *If $A \in \mathbb{R}^{n\times n}$ is a definite direct-distances matrix then the fundamental eigenvectors of A are all columns of the longest-distances matrix, that is vectors of longest-distances to all nodes of D_A.*

8. The eigenproblem: a link between the longest distances problem and the assignment problem

When solving the assignment problem for a matrix A it may be useful not only to know an optimal solution (permutation) but to get a more comprehensive information, for instance to find alternative optimal solutions, or a best solution satisfying certain conditions, etc. A very suitable workplace for this purpose is a normal form of A in which optimal solutions are easily identified by zeros. It is an efficient encoding of a set of a size up to $n!$ using n^2 entries. If the Hungarian method is used then the normal form is a by-product of the solution method but when a different method is used then an optimal solution may only be known without actually having a normal form. If an optimal solution is known then it may be assumed to be id (otherwise we appropriately permute the rows or columns). Due to Proposition 3.2(a) we may also assume that all diagonal elements of A are zero. Then to find a normal form by a similarity transformation (matrix scaling) means to find $x_1, \ldots, x_n$ and $y_1, \ldots, y_n$ so that

$$\begin{aligned} &y_i + a_{ij} + x_j \leqslant 0 \quad \text{for all } i, j \in N, \\ &y_i + a_{ii} + x_i = 0 \quad \text{for all } i \in N. \end{aligned} \tag{8.1}$$

Hence $y_i = -x_i$ for all $i \in N$ and the inequalities (8.1) are equivalent to

$$\max_{j \in N}(-x_i + a_{ij} + x_j) = 0 \quad \text{for all } i \in N$$

or,

$$\max_{j \in N}(a_{ij} + x_j) = x_i \quad \text{for all } i \in N$$

or, in the max-algebraic notation

$$A \otimes x = x. \tag{8.2}$$

Note that maper$(A) = 0$, yielding that A is definite since any positive cycle could be extended by complementary diagonal zeros to a permutation of positive weight. Hence $\lambda(A) = 0$ and therefore (8.2) has a solution, namely every eigenvector of A is a solution.

We have arrived at the following:

Theorem 8.1. *The set of vectors which can be used for scaling a definite matrix A to a normal form is* sp(A). *That is if A is definite then*

$$\operatorname{sp}(A) = \left\{x = (x_1, \ldots, x_n)^{\mathrm{T}};\ (x_i + a_{ij} - x_j) \text{ is normal}\right\}.$$

In particular, every longest-distances vector corresponding to the definite direct-distances matrix A can be used to scale A to a normal form.

Note that to the author's knowledge the relationship between the structure of the set of normal forms and the structure of the eigenspace has not been studied yet.

9. The eigenproblem and matrix scaling

It follows from the definitions that every strictly normal matrix is normal and every normal matrix is definite. We have also previously seen that every square matrix is similar to a normal one (which can be found by the Hungarian method). Not every matrix is similar to a strictly normal matrix and it follows from Theorem 4.4 that strong regularity is a necessary condition for a matrix to be similar to a strictly normal one. We will now show that this condition is also sufficient. This will enable us to describe the set of vectors that can be used for scaling to strictly normal form.

Assume that $A = (a_{ij}) \in \mathbb{R}^{n\times n}$ is definite and that $b \in \mathbb{R}^n$ is a vector for which the system $A \otimes x = b$ has a unique solution, thus $\mathrm{ap}(A) = \{\mathrm{id}\}$. Let

$$B = \mathrm{diag}\left(b_1^{-1}, b_2^{-1}, \ldots, b_n^{-1}\right) \otimes A \otimes \mathrm{diag}(b_1, b_2, \ldots, b_n).$$

Then (Proposition 3.2(a)) $\mathrm{ap}(B) = \mathrm{ap}(A) = \{\mathrm{id}\}$ and B has a unique column maximum in every row and column and zero diagonal. Hence B is strictly normal (and thus strong regularity is a sufficient condition). Conversely, if A is definite and

$$\mathrm{diag}(c_1, \ldots, c_n) \otimes A \otimes \mathrm{diag}(b_1, b_2, \ldots, b_n)$$

is strictly normal then $c_i \otimes b_i = 0$ for all $i \in N$, yielding $c_i = b_i^{-1}$ for all $i \in N$. Therefore in

$$\mathrm{diag}\left(b_1^{-1}, b_2^{-1}, \ldots, b_n^{-1}\right) \otimes A$$

all columna maxima are on the diagonal only and thus $A \otimes x = b$ has a unique solution. We have proved:

Theorem 9.1. *Let $A \in \mathbb{R}^{n\times n}$ be definite. Then $\mathrm{diag}(c_1, \ldots, c_n) \otimes A \otimes \mathrm{diag}(b_1, b_2, \ldots, b_n)$ is normal if and only if $c_i = b_i^{-1}$ for all $i \in N$ and the system $A \otimes x = b$ has a unique solution.*

This result prompts us to investigate how to find a vector b for which the system $A \otimes x = b$ has a unique solution. Recall that we have defined matrices for which such a vector exists as strongly regular and note that Theorem 4.5 gives a practical criterion for checking that a matrix is strongly regular. Nevertheless it is still not quite clear how to find such a vector b. First of all notice that if A is a strongly regular definite matrix and $A \otimes x = b$ has a unique solution $\bar{x}$ then $\bar{x} = b$. Hence $b \in \mathrm{sp}(A)$. However, not every eigenvector of A is satisfying the requirements for b. To give a full answer to this question we introduce the following notation (here as before $A = (a_{ij}) \in \mathbb{R}^{n\times n}$):

$$\mathrm{Im}(A) = \left\{A \otimes x; x \in \mathbb{R}^n\right\},$$

$$S_A = \left\{b \in \mathbb{R}^n; A \otimes x = b \text{ has a unique solution}\right\}.$$

$\mathrm{Im}(A)$ and S_A are called the *image set* and the *simple image set* of A, respectively. Since $A^{(k+1)} \otimes x = A^{(k)} \otimes (A \otimes x)$ we have $\mathrm{Im}(A^{(k+1)}) \subseteq \mathrm{Im}(A^{(k)})$ for every k natural. It follows then from (7.1) that for a definite matrix A,

$$\mathrm{Im}(A) \supseteq \mathrm{Im}\left(A^{(2)}\right) \supseteq \mathrm{Im}\left(A^{(3)}\right) \supseteq \cdots$$
$$\cdots \supseteq \mathrm{Im}\left(A^{(n-1)}\right) = \mathrm{Im}\left(A^{(n)}\right) = \mathrm{Im}\left(A^{(n+1)}\right) = \cdots = \mathrm{sp}(A),$$

see Fig. 2.

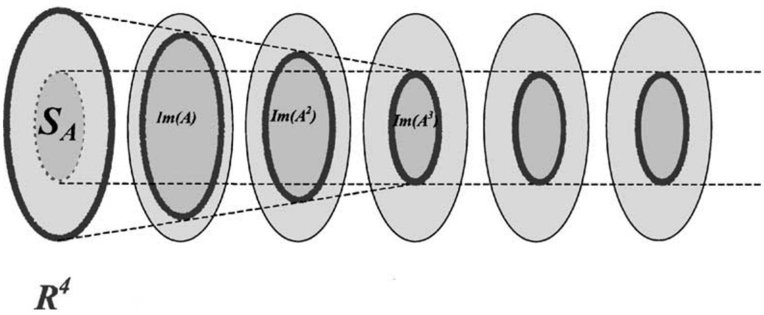

Fig. 2.

Theorem 9.2 ([8]). *If A is definite and strongly regular then* $S_A = \text{int}(\text{sp}(A))$.

So, our task of describing the scaling vectors of a definite, strongly regular matrix A to a strictly normal form has now been reduced to that of finding internal points of the eigenspace of A.

We will need the following notation: $\tilde{A}$ is the square matrix of the same order as A whose all diagonal elements are $-\infty$ and off-diagonal ones are same as in A. The symbol $V_A(g)$ for $g \in \mathbb{R}$ will stand for the set

$$\left\{v \in \mathbb{R}^n;\ A \otimes v \leqslant g \otimes v\right\}.$$

Theorem 9.3 ([8]). *If A is definite then*

$$S_A = \bigcup_{\lambda(\tilde{A}) \leqslant g < 0} V_{\tilde{A}}(g).$$

Corollary 9.1. *If A is definite and strongly regular then* $\text{sp}(\tilde{A}) \subseteq S_A$.

Proof. If A is strongly regular then $\lambda(\tilde{A}) < 0$ and $\tilde{A} \otimes v = \lambda(\tilde{A}) \otimes v \leqslant g \otimes v$ for any $g \geqslant \lambda(\tilde{A})$. □

We summarise that to scale a definite [definite, strongly regular] matrix A to a normal [strictly normal] form we may use any eigenvector of A [$\tilde{A}$].

10. Characteristic polynomial, the best principal submatrices and the exact cycle cover

The max-algebraic characteristic polynomial called *characteristic maxpolynomial* of a square matrix $A = (a_{ij}) \in \mathbb{R}^{n \times n}$ is defined [14] by

$$\chi_A(x) := \text{maper}(A \oplus x \otimes I).$$

Equivalently, it is the max-algebraic permanent of the matrix

$$\begin{pmatrix} a_{11} \oplus x & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} \oplus x & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \oplus x \end{pmatrix}.$$

Hence

$$\chi_A(x) = \delta_0 \oplus \delta_1 \otimes x \oplus \cdots \oplus \delta_{n-1} \otimes x^{(n-1)} \oplus x^{(n)}$$

for some $\delta_0, \ldots, \delta_{n-1} \in \mathbb{R}$ or, written using conventional notation

$$\chi_A(x) = \max\big(\delta_0, \delta_1 + x, \ldots, \delta_{n-1} + (n-1)x, nx\big).$$

Thus, viewed as a function in x, the characteristic maxpolynomial of a matrix A is a piecewise linear convex function whose slopes are from the set $\{0, 1, \ldots, n\}$.

Theorem 10.1 ([4]). *$\chi_A(x)$ as a function can be found in $O(n^4)$ steps.*

The method in [4] is based on ideas from computational geometry combined with solving assignment problems.

If for some $k \in \{0, \ldots, n-1\}$ the inequality

$$\delta_k \otimes x^{(k)} \leqslant \sum_{i \neq k}^{\oplus} \delta_i \otimes x^{(i)}$$

holds for every real x then the term $\delta_k \otimes x^{(k)}$ is called *inessential*, otherwise it is called *essential*. Obviously, if $\delta_k \otimes x^{(k)}$ is inessential then

$$\chi_A(x) = \sum_{i \neq k}^{\oplus} \delta_i \otimes x^{(i)}$$

holds for every real x. Thus inessential terms are not needed for the description of $\chi_A(x)$ as a function of x. Note also that they will not be found by the method described in [4].

It is known [15] that every max-algebraic polynomial can be expressed as a max-algebraic product of a constant and of expressions of the form $(x \oplus \beta)^{(p)}$ where β and p are real constants. The constants β are called the corners of the max-polynomial. The characteristic maxpolynomials have a remarkable property resembling the conventional characteristic polynomials:

Theorem 10.2 ([14]). *The greatest corner of $\chi_A(x)$ is $\lambda(A)$.*

The reader is referred to [1,14] for more information about characteristic max-polynomials including an analogue of the Cayley–Hamilton theorem.

We will now deal with combinatorial aspects of the characteristic maxpolynomial. Let $A = (a_{ij})$ be an $n \times n$ matrix. As usual, any matrix of the form

$$\begin{pmatrix} a_{i_1 i_1} & a_{i_1 i_2} & \cdots & a_{i_1 i_k} \\ a_{i_2 i_1} & a_{i_2 i_2} & \cdots & a_{i_2 i_k} \\ \vdots & \vdots & & \vdots \\ a_{i_k i_1} & a_{i_k i_2} & \cdots & a_{i_k i_k} \end{pmatrix}$$

with $1 \leqslant i_1 < i_2 < \cdots < i_k \leqslant n$ is called a $k \times k$ *principal submatrix of* A. The set of all $k \times k$ *principal submatrices of* A will be denoted as $A(k)$.

The *best principal submatrix problem* (BPSM) is defined as follows: *Given a matrix* $A = (a_{ij}) \in \mathbb{R}^{n\times n}$ and an integer k, $1 \leqslant k \leqslant n$, *find*

$$\max_{B\in A(k)} \operatorname{maper}(B).$$

Theorem 10.3 ([14]). *If* $A \in \mathbb{R}^{n\times n}$ *then* $\delta_k = \max_{B\in A(n-k)} \operatorname{maper}(B)$ *for all* $k = 0, 1, \ldots, n-1$.

So, in particular, $\delta_0 = \operatorname{maper}(A)$ and $\delta_{n-1} = \max(a_{11}, a_{22}, \ldots, a_{nn})$. It follows from Theorem 10.1 that all coefficients δ_k corresponding to essential terms of the characteristic maxpolynomial can be found in $O(n^4)$ time. Therefore BPSM for A can be solved in polynomial time if all terms of $\chi_A(x)$ are essential.

However, no polynomial method is known to the author for solving the best principal submatrix problem in general. We present now a few problems in the conventional terminology which are equivalent to BPSM. All of these are related to the special binary case when the matrices are over the set $T = \{0, -\infty\}$. The set of $n \times n$ matrices over T will be denoted by $T^{n\times n}$. Clearly, $\delta_k = 0$ or $-\infty$ for every k if $A \in T^{n\times n}$. The question whether $\delta_k = 0$ for a given k is very closely related to the exact cycle cover problem in digraphs and can also easily be formulated in terms of classical permanents:

Theorem 10.4 ([9]). *Let* $A = (a_{ij}) \in T^{n\times n}$ *and* $k \in \{0, \ldots, n-1\}$. *Then the following are equivalent*:

- $\delta_{n-k} = 0$.
- *In the digraph* Z_A *there is a set of* k *nodes which is the union of the sets of nodes of node-disjoint cycles in* Z_A.
- *There exists a* $k \times k$ *submatrix of the zero–one matrix* $e^A = (e^{a_{ij}})$ *with positive permanent.*

Note that for some special types of matrices (such as diagonally dominant or permutation matrices) BPSM can be solved in polynomial time [9]. Yet, no efficient solution method seems to be known for general matrices over T. It may therefore be of interest that the biggest value of k for which the answer to the three equivalent problems mentioned in Theorem 10.4 is affirmative, can always be found relatively easily:

Theorem 10.5 ([9]). *If* $A = (a_{ij}) \in T^{n\times n}$ *then the biggest value of k for which* $\delta_{n-k} = 0$ *is* $n + \mathrm{maper}(A \oplus (-1) \otimes I)$ *and therefore it can be determined in* $\mathrm{O}(n^3)$ *time*.

Note that in [11] a generalization of this result has been proved for matrices over $\bar{\mathbb{R}}$.

11. Overview of links

Max-algebra	Combinatorics	Combinatorial Optimization
maper(A)	Term rank	Linear assignment problem
$A \otimes x = b$		
$\exists x$	Set covering	
$\exists! x$	Minimal set covering	
$\exists b\ \exists! x$	Digraph acyclic	Unique opt permutation
$\Gamma(A)$ if A definite	Transitive closure	Longest-distances matrix
$A \otimes x = \lambda \otimes x$		
λ	$\exists$ positive cycle	Maximum cycle mean
x		Vector of balancing coefficients
x if A definite	Connectivity to a node	Longest-distances to a node
x if A definite		Scaling to normal form
$\tilde{A} \otimes x = \lambda \otimes x$		
x if A definite, SR		Scaling to strictly normal form
Regularity	$\nexists$ even directed cycle	All optimal permutations of the same parity
	0–1 sign-nonsingularity	
Char. polynomial	$\exists$ exact cycle cover	Best principal submatrix
	$\exists$ principal submatrix with positive permanent	

References

[1] F.L. Baccelli, G. Cohen, G.-J. Olsder, J.-P. Quadrat, Synchronization and Linearity, John Wiley, Chichester, New York, 1992.
[2] R.B. Bapat, T.E.S. Raghavan, Nonnegative Matrices and Applications, Cambridge University Press, London, 1997.
[3] R.A. Brualdi, H. Ryser, Combinatorial Matrix Theory, Cambridge University Press, London, 1981.

[4] R.E. Burkard, P. Butkovič, Finding all essential terms of a characteristic maxpolynomial, SFB Report No. 249, Institute of Mathematics, Graz University of Technology, 2002.
[5] R.E. Burkard, E. Çela, Linear assignment problems and extensions, in: P.M. Pardalos, D.-Z. Du (Eds.), Handbook of Combinatorial Optimization, Supplement vol. A, Kluwer Academic Publishers, 1999, pp. 75–149.
[6] P. Butkovič, Strong regularity of matrices—a survey of results, Discrete Appl. Math. 48 (1994) 45–68.
[7] P. Butkovič, Regularity of matrices in min-algebra and its time-complexity, Discrete Appl. Math. 57 (1995) 121–132.
[8] P. Butkovič, Simple image set of (max, +) linear mappings, Discrete Appl. Math. 105 (2000) 73–86.
[9] P. Butkovič, On the coefficients of the max-algebraic characteristic polynomial and equation, in: Proceedings of the Workshop on Max-algebra, Symposium of the International Federation of Automatic Control, Prague, 2001.
[10] P. Butkovič, F. Hevery, A condition for the strong regularity of matrices in the minimax algebra, Discrete Appl. Math. 11 (1985) 209–222.
[11] P. Butkovič, L. Murfitt, Calculating essential terms of a characteristic maxpolynomial, CEJOR 8 (2000) 237–246.
[12] J. Cochet-Terrasson, G. Cohen, S. Gaubert, M. McGettrick, J.-P. Quadrat, Numerical computation of spectral elements in max-plus algebra, in: IFAC Conference on System Structure and Control, 1998.
[13] R.A. Cuninghame-Green, Minimax Algebra, Lecture Notes in Economics and Math. Systems, vol. 166, Springer, Berlin, 1979.
[14] R.A. Cuninghame-Green, The characteristic maxpolynomial of a matrix, J. Math. Anal. Appl. 95 (1983) 110–116.
[15] R.A. Cuninghame-Green, Minimax algebra and applications, in: Advances in Imaging and Electron Physics, vol. 90, pp. 1–121, Academic Press, New York, 1995.
[16] M. Fiedler, V. Pták, Diagonally dominant matrices, Czechoslovak Math. J. 92 (1967) 420–433.
[17] S. Gaubert, Théorie des systèmes linéaires dans les dioïdes, Thèse, Ecole des Mines de Paris, 1992.
[18] S. Gaubert et al., Algèbres Max-Plus et applications en informatique et automatique, 26ème école de printemps d'informatique théorique, Noirmoutier, 1998.
[19] M. Gondran, Path algebra and algorithms, in: B. Roy (Ed.), Combinatorial programming: methods and applications (Proc. NATO Advanced Study Inst., Versailles, 1974) NATO Advanced Study Inst. Ser., Ser. C: Math. and Phys. Sci., 19, Reidel, Dordrecht, pp. 137–148, 1975.
[20] M. Gondran, M. Minoux, L'indépendance linéaire dans les dioïdes, Bulletin de la Direction Etudes et Recherches, EDF, Série C 1 67-90, 1978.
[21] M. Gondran, M. Minoux, Linear algebra of dioïds: a survey of recent results, Ann. Discrete Math. 19 (1984) 147–164.
[22] C.H. Papadimitriou, K. Steiglitz, Combinatorial Optimization-Algorithms and Complexity, Dover, 1998.
[23] N. Robertson, P.D. Seymour, R. Thomas, Permanents, Pfaffian orientations and even directed circuits, Ann. Math. 150 (2) (1999) 929–975.
[24] H. Schneider, M.H. Schneider, Max-balancing weighted directed graphs and matrix scaling, Math. Oper. Res. 16 (1) (1991) 208–222.
[25] C. Thomassen, Sign-nonsingular matrices and even cycles in directed graphs, Linear Algebra Appl. 75 (1986) 27–41.
[26] K. Zimmermann, Extremální algebra, Výzkumná publikace Ekonomicko—matematické laboratoře při Ekonomickém ústavě ČSAV, 46, Praha, 1976 (in Czech).
[27] U. Zimmermann, Linear and Combinatorial Optimization in Ordered Algebraic Structures, in: Ann. Discrete Math., vol. 10, North-Holland, Amsterdam, 1981.