SPECIAL AND STRUCTURED MATRICES IN MAX-PLUS ALGEBRA

by

Daniel Jones

A thesis submitted to
The University of Birmingham
for the degree of
Doctor of Mathematics

School of Mathematics
The University of Birmingham
March 2017

Abstract

The aim of this thesis is to present efficient (strongly polynomial) methods and algorithms for problems in max-algebra when certain matrices have special entries or are structured.

First, we describe all solutions to a one-sided parametrised system. Next, we consider special cases of two-sided systems of equations/inequalities. Usually, we describe a set of generators of all solutions but sometimes we are satisfied with finding a non-trivial solution or being able to say something meaningful about a non-trivial solution should it exist. We look at special cases of the generalised eigenproblem, describing the full spectrum usually. Finally, we prove some results on 2×2 matrix roots and generalise these results to a class of $n \times n$ matrices.

Main results include: a description of all solutions to the two-dimensional generalised eigenproblem; observations about a non-trivial solution (should it exist) to essential/minimally active two-sided systems of equations; the full spectrum of the generalised eigenproblem when one of the matrices is an outer-product; the unique candidate for the generalised eigenproblem when the difference of two matrices is symmetric and has a saddle point and finally we explicitly say when a 2×2 matrix has a kth root for a fixed positive integer k.

Acknowledgments

I would like to thank my PhD supervisor Professor Peter Butkovič for introducing me to max-algebra and for his invaluable help in creating this thesis. He has consistently gone above and beyond what is expected of him and his passion for the subject continues to inspire and motivate me - his time and efforts made this work possible. I would also like to thank my close friends and family - they have constantly been there to support me.

Contents

1. Introduction

When one replaces the operation of addition by maximum and the operation of multiplication by addition, one can provide mathematical theory and techniques for solving classical, non-linear problems which take on the form of linear problems in this setting of max-algebra.

1.1 Max-algebra

We assume everywhere that $m, n \geq 1$ are natural numbers and denote $M = \{1, ..., m\}$ and $N = \{1, ..., n\}$; the symbol $\overline{\mathbb{R}}$ stands for $\mathbb{R} \cup \{-\infty\}$ and the symbol $\overline{\overline{\mathbb{R}}}$ stands for $\overline{\mathbb{R}} \cup \{+\infty\}$. We use the convention $\max \emptyset = -\infty$ and $\min \emptyset = +\infty$.

If $a, b \in \overline{\overline{\mathbb{R}}}$ then we set

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

and

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

Note that by definition

$$(-\infty) + (+\infty) = -\infty = (+\infty) + (-\infty).$$

It will also be useful to define the *dual* operation. For $a, b \in \overline{\overline{\mathbb{R}}}$, we set

$$a \oplus' b = \min(a, b)$$

and

$$a \otimes' b = a + b.$$

Throughout this thesis we denote $-\infty$ by ε (the neutral element with respect to $\oplus$) and for convenience we also denote by the same symbol any vector, whose all components are $-\infty$, or a matrix whose all entries are $-\infty$. A similar convention is used for 0 vectors or matrices. If $a \in \mathbb{R}$ then the symbol a^{-1} stands for $-a$. The symbol a^k ($k \geq 1$ integer) stands for the iterated product $a \otimes a \otimes \ldots$ in which the symbol a stands k times (that is ka in conventional notation). By *max-algebra* (also called "tropical linear algebra") we understand the analogue of linear algebra developed for the pair of operations $(\oplus, \otimes)$, extended to matrices and vectors as in conventional linear algebra. That is if $A = (a_{ij}), B = (b_{ij})$ and $C = (c_{ij})$ are matrices of compatible sizes with entries from $\overline{\mathbb{R}}$, we write $C = A \oplus B$ if $c_{ij} = a_{ij} \oplus b_{ij}$ for all i, j and $C = A \otimes B$ if

$$c_{ij} = \bigoplus_k a_{ik} \otimes b_{kj} = \max_k (a_{ik} + b_{kj})$$

for all i and j. If $\alpha \in \overline{\mathbb{R}}$ then $\alpha \otimes A = (\alpha \otimes a_{ij})$.

1.2 Literature overview

Max-algebra has been appearing in books and research papers since the 1960s. The first paper, perhaps, was that of R. A. Cuninghame-Green [33] in 1960, where his work in the Sheffield steelworks made it clear that max-algebra could be a powerful tool in modelling industrial processes or, more generally, interactive processes. Cuninghame-Green went on to produce other articles, including [34], [35], [36] and [37]. A number of other, independent articles were also produced. For example: B. A. Carré [26], L. Elsner [38], G. M. Engel and H. Schneider [39],[40],[60], M. Gondran and M. Minoux [52],[49],[50], [51], B. Giffler [47], [48] and N. N. Vorobyov [64],[65].

The field of max-algebra has been developed intensively since. In 1979, Cuninghame-Green's lecture notes [35] helped to popularise max-algebra and bring it to the attention of the mathematical community. Another milestone was reached in 1981 when U. Zimmermann discussed combinatorial optimisation in ordered algebraic structures [68] - emphasising the importance of this area of mathematics. In 1984, there was a breakthrough in the theory of two-sided systems when a first algorithm for finding a solution was made by Butkovič and Hegedüs [16]. In 1992 Baccelli et al explored the uses of max-algebra in synchronising dynamical systems [8] - emphasising the applications of max-algebra to scheduling problems. Binding and Volkmer (2007) explored the generalised eigenproblem in max-algebra [12] in parallel with Butkovič and Cuninghame-Green [32]. This theoretical problem has inspired much of the work in this thesis. Butkovič's book [25] (2010) was important in unifying well-known classical results and modern results.

Interest in max-algebra is partly due to its ability to take non-linear problems in classical linear algebra and present them in a linear way (discrete event systems [28], [46] for example). The problems arising in max-algebra are often of a managerial nature, arising in areas such as: manufacturing [33], [34], transportation [26], traffic light control [29], allocation of resources [1] and the natural sciences [37]. More recently, there is a description of how to model the entire Dutch railway system using max-algebra [53]. There are also applications in banking using tropical geometry [54]. See [43] for more applications.

For formal definitions, refer to chapter 2. We introduce some of the most important problems related to this thesis with motivating examples and a brief discussion of what is currently known and what is believed to be unknown.

Example 1.1. *Consider m partial products P_i, prepared using n machines. Let a_{ij} be the duration of the work of the jth machine needed to complete the partial product for P_i.*

Denote by x_j the starting time of machine j. Then P_i will be ready at time

$$\max\left(x_1 + a_{i1}, \ldots, x_n + a_{in}\right).$$

If $b_1, \ldots, b_m$ are target completion times, then we have the system of equations

$$(\forall i) \max\left(x_1 + a_{i1}, \ldots, x_n + a_{in}\right) = b_i.$$

The compact form (using max-algebraic notation) is

$$A \otimes x = b$$

and so the problem is to find a vector x such that $A \otimes x = b$.

The matrix A is called the production matrix. A related problem (when we require to not exceed given target times) is

$$A \otimes x \leq b.$$

The problems

$$A \otimes x = b \tag{1.1}$$

and

$$A \otimes x \leq b, x \neq \epsilon \tag{1.2}$$

are examples of scheduling problems and are called *one-sided max-linear systems of equations (inequalities)* respectively.

This model is called the *multi-machine interactive production process* (MMIPP) and is the basis for subsequent models.

Systems (1.1) and (1.2) were studied in the first papers on max-algebra in [33], [64] and the theory has further evolved in the 1960s and 1970s [67], [68], and later [23], [24].

It should be noted that these one-sided max-linear systems can be solved more easily than their linear-algebraic counterparts. Also, unlike in conventional linear algebra, systems of inequalities (1.2) always have a solution and the task of finding a solution to (1.1) is strongly related to the same task for the system of inequalities.

We can describe, in strongly polynomial time, the full set of solutions to systems (1.1) and (1.2), using algebraic and combinatorial methods.

We explore a generalisation of system (1.1) in chapter 3.

Example 1.2. *As part of a wider MMIPP, suppose k other machines produce partial products for products $Q_1, \ldots, Q_m$ and the duration and starting times times are b_{ij} and y_j respectively. The synchronisation problem is to find starting times of all $n+k$ machines (vectors x and y) so that each pair (P_i, Q_i) is completed at the same time, yielding the two-sided max-linear systems of equations*

$$A \otimes x = B \otimes y,$$

or

$$A \otimes x = B \otimes x$$

in the case where starting times x_j and y_j must be the same.

Another variant is when the starting times are linked (a fixed interval between the starting times x_j and y_j). This yields a generalised eigenproblem (GEP):

$$A \otimes x = \lambda \otimes B \otimes x,$$

where we are required to find a pair λ (real number) and starting time vector x.

The systems

$$A \otimes x = B \otimes x, x \neq \epsilon, \quad (1.3)$$

$$A \otimes x = B \otimes y, x \neq \epsilon, y \neq \epsilon \quad (1.4)$$

and

$$A \otimes x = \lambda \otimes B \otimes x, \lambda \neq \epsilon, x \neq \epsilon \quad (1.5)$$

are *synchronisation problems.*

Unlike in conventional linear algebra, moving from the task of finding a solution to a one-sided system (1.1) to finding a solution to a two-sided system (1.3) means a significant change in difficulty of the problem. Systems (1.4) are studied in [31] and can be easily transformed to systems (1.3). Two-sided linear systems (1.3) were first studied in [19] - [22]. We know the solution set of (1.3) is finitely generated [16] and we are reasonably confident in solving such systems (see the pseudopolynomial Alternating Method in [25] and [7]).

We do not yet know whether two-sided systems (1.3) are polynomially solvable. It follows from the results in [11] that two-sided systems (1.3) are polynomially equivalent to mean payoff games, a well known problem in NP $\cap$ co-NP. Informally, NP is the set of all decision problems for which "yes instances" have efficiently verifiable proofs and co-NP is the class of problems for which "no instances" have efficiently verifiable counterexamples given the appropriate certificate. It is known that P $\subseteq$ NP $\cap$ co-NP and it is not known whether equality holds. Thus, there is good reason to hope that two-sided systems are indeed polynomially solvable. There exist algorithms to solve mean payoff games in polynomial time in special cases [5] and the tropical shadow-vertex algorithm [6] solves mean payoff games in polynomial time on average subject to some constraints.

The theory of symmetrised semirings yields necessary conditions for the solvability of (1.3). Chapters 4, 5, 7 and 8 are concerned with solving special cases of such two-sided systems in polynomial time, in the hope of shedding light on the general case. Chapter 8 may prove to shed some light on the general case. The contents of chapter 8 have been submitted as a paper to Discrete Applied Mathematics.

It is likely that GEP is much more difficult than the eigenproblem. This is indicated by the fact that the GEP for a pair of real matrices may have no generalised eigenvalue, a finite number or a continuum of generalised eigenvalues [32]. It is known [63] that the union of any system of closed (possibly one-element) intervals is the set of generalised eigenvalues for suitably taken A and B.

GEP has been studied for the first time in [12] and [32]. The first of these papers solves the problem completely when the matrices have exactly two rows and special cases for general sized matrices; the second solves some other special cases. No solution method seems to be known either for finding a λ or an (non-trivial) x satisfying (1.5) for general real matrices. Obviously, once λ is fixed, the GEP reduces to a system of the form (1.3). We therefore usually concentrate on the task of finding the set of generalised eigenvalues. The level set method [44] is a pseudopolynomial algorithm for finding the generalised eigenvalues.

In chapters 6, 9, 10 and 11 we examine some special cases of GEP. Essential parts of chapters 6, 10 and 11 are the contents of a published paper [17] in the SIAM Journal on Matrix Analysis and Applications.

Other problems are obtained when the MMIPP is considered as a multi-stage process.

Example 1.3. *Suppose the machines work in stages, in which all machines simultaneously produce components necessary for the next stage of some or all other machines. If we let $x_i(r)$ denote the starting time of the rth stage on machine i and let a_{ij} denote the duration of the operation at which the jth machine prepares a component necessary for* *the ith machine at the* $(r+1)$*st stage, then*

$$x(r+1) = A \otimes x(r)$$

in the max-algebraic notation. We say the system reaches steady regime if it moves forward in regular steps after a certain point. That is, for some λ *and* r_0 *we have*

$$x(r+1) = \lambda \otimes x(r) \text{ for all } r \geq r_0.$$

This happens if and only if for some λ *and* $r, x(r)$ *is a solution to*

$$A \otimes x = \lambda \otimes x.$$

System (1.6) below is a *stability problem* and called the *eigenproblem*, where we are required to find a pair λ (real number) and a starting time vector x. Considering also the related *subeigenvector problem*, we have the *stability problems*

$$A \otimes x = \lambda \otimes x, x \neq \epsilon \tag{1.6}$$

and

$$A \otimes x \leq \lambda \otimes x, x \neq \epsilon. \tag{1.7}$$

The eigenproblem (1.6) is of key importance in max-algebra. It has been studied since the 1960s [34] in connection with the analysis of the steady-state behaviour of production systems. All solutions of the eigenproblem, in the case of irreducible matrices, are described in [35] and [49], see also [9] and [64]. A general spectral theorem for reducible matrices has appeared in [45], [10] and partly in [27].

A real number λ is an *eigenvalue* if there exists $x \in \overline{\mathbb{R}}^n, x \neq \epsilon$ such that $A \otimes x = \lambda \otimes x$ and there are at most n such eigenvalues [25] (where A is an $n \times n$ matrix). The full spectrum of eigenvalues and generators of all eigenvectors can be found in $O(n^3)$ time and so system (1.6) is efficiently solved. The spectrum of eigenvalues can be used to bound the size of the eigenvalues of an associated non-negative matrix in classical linear algebra [4]. Also, [58] contains proofs that the max-algebraic roots of a max-algebraic polynomial can provide a good approximation to the classical eigenvalues of an associated matrix polynomial. The advantage of using max-algebra here is that the max-algebraic roots can be calculated in low-order polynomial time, and can then be used as starting points for algorithms which search for classical roots/eigenvalues.

All finite solutions to (1.7) are described in [62].

1.3 Thesis overview

The aim of this thesis is to present efficient (polynomial) methods and algorithms for problems in max-algebra when certain matrices have special entries or are structured, focusing on two-sided systems of equations/inequalities and the generalised eigenproblem.

The work herein can be viewed as a natural progression from our MSci project, which concentrated on the eigenproblem $A \otimes x = \lambda \otimes x$ (refer to chapter 2 for notation), where the matrix A has special and/or structured forms.

The eigenproblem is polynomially solvable in the most general case, whereas no such polynomial algorithms seem to be known for the most general cases of two-sided systems of equations/inequalities and the generalised eigenproblem.

We summarise here the contents of the chapters of this thesis. We also briefly mention what is already known about the relevant problem (including known complexity results of any existing algorithms) and what is new in that chapter.

The first type of max-linear systems we consider are one-sided parametrised systems

$A \otimes x = b(\alpha)$ subject to upper and lower bounds on α. These differ from the known one-sided systems $A \otimes x = b$ with the appearance of a parameter α in the right hand side vector. The regular one-sided systems are well known and can be solved in low order, strongly polynomial time. In chapter 3, we explicitly describe the full set of α for which the one-sided parametrised system has a solution. It then follows from known results that we can describe the full set of solutions for any fixed α. Thus, we can describe all solutions to the one-sided parametrised system in strongly polynomial time.

The next systems we explore are the two-sided systems $A \otimes x \leq B \otimes x, A \otimes x = B \otimes x$ and $A \otimes x = B \otimes y$. We are not aware of any polynomial method for finding a solution in the most general case for these problems, though the alternating method [25] is a pseudopolynomial method for systems of two-sided equations for integer matrices. In chapter 4, we describe a strongly polynomial method for solving the system $A \otimes x \leq B \otimes x$ when B has exactly one finite entry per row. The aggregation method reduces the problem to the well known subeigenvector problem (so allowing us to describe all finite solutions in polynomial time). Continuing with the system $A \otimes x = B \otimes x$, in chapter 5 we consider the case when A and B each have exactly two finite entries per row, appearing in the same position. Again, by reducing to the subeigenvector problem, we are able to describe all finite solutions (after some variables have been trivially set to ϵ).

In chapter 6 we describe all solutions to $A \otimes x = B \otimes x$ when matrices A and B each have exactly two columnns. In the same chapter, we consider another problem, namely the generalised eigenproblem $A \otimes x = \lambda \otimes B \otimes x$. Since for fixed λ, this is a two-sided system, the focus is usually on finding the set of λ for which there exists a corresponding eigenvector. To our knowledge, there is no known method for finding even a single λ in polynomial time in the most general case. In chapter 6 we explicitly describe the full set of λ in the two-dimensional case.

In chapter 7 we explore the two-sided system $A \otimes x = B \otimes y$ when B (say) has exactly two columns. By relating this to the one-sided parametrised systems of chapter 3, we can describe all solutions in polynomial time.

In chapter 8 we demonstrate the pivotal role of the matrix $C = A \oplus B$ and its max-algebraic permanent for solving two-sided linear systems $A \otimes x = B \otimes x$ of *minimally active* or *essential* type where A and B are finite square matrices.

In chapter 9 we show the generalised eigenproblem has a unique solution (which we give explicitly) when the matrices A and B are circulant. Further, we show that the constant vector is a corresponding eigenvector. In chapter 10, we give the full set of λ in the case when B (say) is an outer-product.

In chapter 11 we show that if the matrix $C = A - B$ (in the classical notation) is symmetric and has a saddle point, then the value of the saddle point is the unique candidate for the eigenvalue λ satisfying $A \otimes x = \lambda \otimes B \otimes x$. In the 3×3 case we give a necessary and sufficient condition for the saddle point to be an eigenvalue, we also give a necessary condition in the $n \times n$ case.

Finally, in chapter 12, we consider a separate topic - the problem of finding matrix roots in max-algebra. In general, this is known to be a hard problem. In the 2×2 case we characterise all matrices for which there exists a matrix root and define the kth root of a matrix for a positive integer k. We generalise this to identify a class of $n \times n$ matrices for which there exists a kth root (which we give explicitly).

Among the main results of this thesis are:

- A description of all solutions to the two-dimensional generalised eigenproblem in Section 6.4;

- Lemma 8.3 in Chapter 8 which is crucial for the main results - Theorems 8.6 and 8.29 (which identify active elements of a solution to the minimally active/essential two-sided system) ;

- Theorem 10.7 gives the full spectrum for GEP when B is an outer product;
- Theorem 11.6 identifies the unique candidate for the eigenvalue of GEP in the special case where the difference of matrices A and B is symmetric and has a saddle point and finally
- In Chapter 12 we explicitly say when a 2×2 real matrix has a kth root (for any fixed positive integer k) and define the kth root in such cases. We also generalise to special types of $n \times n$ matrices.

2. Preliminaries

In this section we give the definitions and some basic results which will be used in the formulations and proofs of the results of this thesis. For the proofs and more information about max-algebra the reader is referred to [2], [8], [25] and [53].

Let $A = (a_{ij}) \in \mathbb{R}^{m \times n}$. We will use the following notation:

$$M_j(A) = \left\{ r \in M; a_{rj} = \max_{i \in M} a_{ij} \right\}, \quad j \in N;$$

$$N_i(A) = \{ j \in N; i \in M_j(A) \}, \quad i \in M.$$

We will also write M_j, N_i instead of $M_j(A), N_i(A)$ if no confusion can arise. The following will be useful.

Proposition 2.1. $\bigcup_{j \in N} M_j = M$ *if and only if* $N_i \neq \emptyset$ *for every* $i \in M$.

Proof. Straightforward from definitions. □

Although the use of the symbols $\otimes$ and $\oplus$ is common in max-algebra we will apply the usual convention of not writing the symbol $\otimes$. Thus in what follows the symbol $\otimes$ will not be used and unless explicitly stated otherwise, all multiplications indicated are in max-algebra.

When giving examples of matrices which contain many ϵ entries, we may replace these by a blank space for convenience.

A vector or matrix is called *finite* if all its entries are real numbers. A square matrix is called *diagonal* if all its diagonal entries are real numbers and off-diagonal entries are ε. More precisely, if $x = (x_1, ..., x_n)^T \in \mathbb{R}^n$ then $diag\,(x_1, ..., x_n)$ or just $diag\,(x)$ is the $n \times n$ diagonal matrix

$$\begin{pmatrix} x_1 & \varepsilon & ... & \varepsilon \\ \varepsilon & x_2 & ... & \varepsilon \\ \vdots & \vdots & \ddots & \vdots \\ \varepsilon & \varepsilon & ... & x_n \end{pmatrix}.$$

The matrix $diag\,(0)$ is called the *unit matrix* and denoted I. Obviously, $AI = IA = A$ whenever A and I are of compatible sizes. A matrix obtained from a diagonal matrix [unit matrix] by permuting the rows and/or columns is called a *generalized permutation matrix* [*permutation matrix*]. It is known that in max-algebra, generalized permutation matrices are the only type of invertible matrices [35], [25]. Clearly,

$$(diag\,(x_1, ..., x_n))^{-1} = diag\left(x_1^{-1}, ..., x_n^{-1}\right).$$

Remark 2.2. *Note that* $(.)^{-1}$ *has a different meaning on the left hand side (inverse of a matrix) and right hand side (inverse of a real number) in the above.*

The matrix $A \in \overline{\mathbb{R}}^{m\times n}$ is called *column* (*row*) $\mathbb{R}-$*astic* [35] if $\bigoplus_{i\in M} a_{ij} \in \mathbb{R}$ for every $j \in N$ (if $\bigoplus_{j\in N} a_{ij} \in \mathbb{R}$ for every $i \in M$), that is, when A has no ϵ column (no ϵ row). The matrix A is called *doubly* $\mathbb{R}-$*astic* if it is both column and row $\mathbb{R}$-astic.

Given a matrix $A \in \overline{\mathbb{R}}^{m\times n}$ and a subset $K \subseteq M$, we denote by $A\,[K]$ the $|K| \times n$ sub-matrix of A restricted to the rows of the set K. If, in addition, we have the subset $T \subseteq N$, then denote by $A\,[K:T] \in \overline{\mathbb{R}}^{|K|\times|T|}$ the matrix A restricted to the rows of K and columns of T.

Given a 2×2 matrix $A = (a_{ij})$, we define

$$d(A) := a_{11} a_{22} a_{12}^{-1} a_{21}^{-1}. \tag{2.1}$$

The following statement is probably the historically first result in max-algebra [33] (though the original notation was different); here we denote for $A \in \overline{\mathbb{R}}^{m \times n}$ and $b \in \overline{\mathbb{R}}^{n}$:

$$S(A, b) = \left\{ x \in \overline{\mathbb{R}}^{n}; Ax = b \right\}.$$

Theorem 2.3. *[33], [25] If $A \in \overline{\mathbb{R}}^{m \times n}$ is a matrix with no ε columns, $b \in \mathbb{R}^m$ and $B = (diag\,(b))^{-1} A$ then $S(A, b) \neq \emptyset$ if and only if $\bigcup_{j \in N} M_j(B) = M$.*

Corollary 2.4. *If $A \in \overline{\mathbb{R}}^{m \times n}$ is a matrix with no ε columns, $b \in \mathbb{R}^m$ and $B = (diag\,(b))^{-1} A$ then $S(A, b) \neq \emptyset$ if and only if $N_i(B) \neq \emptyset$ for every $i \in M$.*

Proposition 2.5. *If $A \in \overline{\mathbb{R}}^{m \times n}$ is a matrix with no ε columns, $b \in \mathbb{R}^m$ is a constant vector and $B = (diag\,(b))^{-1} A$, then $M_j(A) = M_j(B)$ for all $j \in N$ and consequently also $N_i(A) = N_i(B)$ for all $i \in M$.*

Proof. Straightforward from definitions. □

Given a matrix $A = (a_{ij}) \in \overline{\mathbb{R}}^{n \times n}$, the symbol D_A will denote the weighted digraph (N, E, w) where $E = \{(i, j) : a_{ij} > \epsilon\}$ and $w(i, j) = a_{ij}$ (or briefly $w(ij)$). If $\pi = (i_1, \ldots, i_p)$ is a path in D_A, then we denote the weight of π by $w(\pi, A) = a_{i_1 i_2} + a_{i_2 i_3} + \cdots + a_{i_{p-1} i_p}$ if $p > 1$ and ϵ if $p = 1$. A path (cycle) is *positive* if it has positive weight.

Given a matrix $A \in \overline{\mathbb{R}}^{m \times n}$, the symbol $\lambda(A)$ will stand for the *maximum cycle mean of A*, that is:

$$\lambda(A) = \max_{\sigma} \mu(\sigma, A)$$

where the maximisation is taken over all *elementary cycles* in D_A (cycles in which there are no repeated vertices) and

$$\mu(\sigma, A) = \frac{w(\sigma, A)}{l(\sigma)}$$

denotes the *mean* of a cycle σ, where $l(\sigma)$ is the length (number of edges) of the cycle σ. Clearly, $\lambda(A)$ always exists since the number of elementary cycles is finite. It follows that D_A is acyclic if and only if $\lambda(A) = \epsilon$.

Given a matrix $A \in \overline{\mathbb{R}}^{m \times n}$ with no ϵ columns and a vector $b \in \overline{\mathbb{R}}$, solving the *one-sided system* (A, b) (see Example 1.1) is the task of finding $x \in \overline{\mathbb{R}}^n$ such that

$$Ax = b. \tag{2.2}$$

Given a square matrix $A \in \overline{\mathbb{R}}^{n \times n}$, finding a solution to the eigenproblem for the matrix A (see Example 1.3) is the task of finding a $\lambda \in \mathbb{R}$ (*eigenvalue*) and $x \in \overline{\mathbb{R}}^n, x \neq \epsilon$ (*eigenvector*) such that

$$Ax = \lambda x. \tag{2.3}$$

Finding a solution to the subeigenproblem for the matrix A (see Example 1.3) is the task of finding a $\lambda \in \mathbb{R}$ (*subeigenvalue*) and $x \in \overline{\mathbb{R}}^n$ (*subeigenvector*) such that

$$Ax \leq \lambda x. \tag{2.4}$$

Given matrices $A, B \in \overline{\mathbb{R}}^{m \times n}$, the problem of finding a non-trivial solution to the two-sided max-linear system of equations (A, B) (see Example 1.2) is the task of finding $x \in \overline{\mathbb{R}}^n, x \neq \epsilon$ such that

$$Ax = Bx. \tag{2.5}$$

Finding a solution to the two-sided max-linear system of inequalities (A, B) (see Example

1.2) is the task of finding $x \in \overline{\mathbb{R}}^n, x \neq \epsilon$ such that

$$Ax \leq Bx. \tag{2.6}$$

If the matrices A and B are finite, then it is easy to see that a non-trivial solution exists for the two-sided system of equations (A, B) if and only if a finite solution exists. As such, we restrict our attention to finding finite solutions to (2.5) in this case. We denote

$$V(A, B) = \{x \in \mathbb{R}^n; Ax = Bx\}. \tag{2.7}$$

Note that some statements remain valid if the finiteness requirement is removed or replaced by the condition that there are no ε columns. We will remind the reader in each chapter if the matrices being considered are finite or not.

Suppose that $A = (a_{ij}) \in \overline{\mathbb{R}}^{m \times n}$ and $B = (b_{ij}) \in \overline{\mathbb{R}}^{m \times n}$ are given. Finding a solution of the *generalized eigenproblem* for (A, B) (see Example 1.2) is the task of finding $x \in \overline{\mathbb{R}}^n, x \neq \varepsilon$ (*generalized eigenvector* or just *eigenvector*) and $\lambda \in \mathbb{R}$ (*generalized eigenvalue* or just *eigenvalue*) such that

$$Ax = \lambda Bx. \tag{2.8}$$

Note that the case $\lambda = \varepsilon$ is trivial and is not discussed here. We denote

$$V(A, B, \lambda) = \left\{x \in \overline{\mathbb{R}}^n; Ax = \lambda Bx, x \neq \varepsilon\right\},$$
$$\Lambda(A, B) = \{\lambda \in \mathbb{R}; V(A, B, \lambda) \neq \emptyset\}.$$

The set $\Lambda(A, B)$ will be called the *spectrum* of the pair (A, B). It is easy to see that if the matrices A and B are finite, then a generalized eigenvector exists if and only if a finite generalized eigenvector exists.

The next two statements [25] provide useful information about the spectrum. Here and in the rest of the thesis (unless said otherwise) we denote

$$C = A - B = (c_{ij}),$$

$$L(C) = \max_{i \in M} \min_{j \in N} c_{ij}$$

and

$$U(C) = \min_{i \in M} \max_{j \in N} c_{ij}.$$

We will write shortly L or U if no confusion can arise.

Proposition 2.6. *[25] $\Lambda(A, B) \subseteq [L, U]$ holds for any $A, B \in \mathbb{R}^{m \times n}$.*

The interval $[L, U]$ will be called the *feasibility interval* for the generalized eigenproblem.

In [12] it is proved that if A and B are symmetric matrices, then $|\Lambda(A, B)| \leq 1$.

For clarity, we use the notation $\bigoplus$ when taking the maximum over a set (max-sum) and the notation $\sum$ when taking the conventional linear sum.

We have the following Lemma ([25], Lemma 7.4.1) which will be used in examples throughout this thesis.

Lemma 2.7 (Cancellation Rule)**.** *Let $v, w, a, b \in \mathbb{R}, a < b$. Then for any real x, we have*

$$v \oplus ax = w \oplus bx$$

if and only if

$$v = w \oplus bx.$$

Let $S \subseteq \overline{\mathbb{R}}^n$. The set S is called a $max - algebraic\ subspace$ if

$$\alpha u \oplus \beta v \in S$$

for every $u, v \in S$ and $\alpha, \beta \in \overline{\mathbb{R}}$. The adjective "max-algebraic" will usually be omitted.

Let $D \in \mathbb{R}^{n\times n}$ and denote by P_n the set of permutations on N. The $max - algebraic\ permanent$ of D is

$$maper\,(D) = \bigoplus_{\sigma\in P_n} \bigotimes_{i\in N} d_{i,\sigma(i)} = \max_{\sigma\in P_n} \sum_{i\in N} d_{i,\sigma(i)}.$$

The set of optimal solutions to the assignment problem (AP) is given by

$$ap\,(D) = \left\{ \sigma \in P_n : \bigotimes_{i\in N} d_{i,\sigma(i)} = maper\,(D) \right\}.$$

It is known that

$$ap\,(D) = ap\,(diag\,(v) \otimes D) \tag{2.9}$$

for all $v \in \mathbb{R}^n$. (See [13] for more information on the assignment problem).

Given $A \in \overline{\mathbb{R}}^{m\times n}$, we define the following series:

$$A^+ = A \oplus A^2 \oplus A^3 \oplus \dots$$

and

$$A^* = I \oplus A^+ = I \oplus A \oplus A^2 \oplus \dots .$$

If these series converge to matrices that do not contain $+\infty$, then the matrix A^+ is called the *weak transitive closure of* A and A^* is called the *strong transitive closure of* A. Note that this happens if and only if $\lambda\,(A) \le 0$. In this case, we have

$$A^+ = A \oplus A^2 \oplus \dots \oplus A^k \text{ for all } k \ge n-1$$

and

$$A^* = I \oplus A \oplus A^2 \oplus \cdots \oplus A^k \text{ for all } k \geq n.$$

The matrices A^+ and A^* are of fundamental importance in max-algebra. This follows from the fact that they enable us to efficiently describe all solutions to

$$Ax = \lambda x, \lambda \in \mathbb{R} \tag{2.10}$$

in the case of A^+, and all finite solutions to

$$Ax \leq \lambda x, \lambda \in \mathbb{R} \tag{2.11}$$

in the case of A^*. Note that in (2.10), for all matrices A there are at most n eigenvalues and all eigenvalues can be found in $O(n^3)$ time. All corresponding eigenspaces also can be found in $O(n^3)$ time [25].

3. One-sided parametrised systems - a strongly polynomial algorithm

3.1 Introduction

In this chapter we develop a theory for a generalisation of one-sided max-linear systems by considering a parametrised version of the problem. Briefly, in addition to constant entries, the vector b also has some parameter entries. The work of this chapter identifies the values of the parameter for which a non-trivial solution exists to the one-sided system.

We show that we can identify the full set for the parameter easily and then, using known methods on one-sided systems, describe all solutions for every such parameter value. That is, we can describe all solutions in strongly polynomial time. The results here are interesting in their own right but the real usefulness of these one-sided parametrised systems is in their ability to find solutions to the seemingly more complicated two-sided systems appearing in chapters 7 and 10.

3.2 Problem formulation

Remark 3.1 (Motivation). *In chapter 7 we explore systems* $Ax = By$*, where* B *has exactly two columns and in chapter 10 we explore systems* $Ax = \lambda Bx$*, where* B *is an outer product. The one-sided parametrised systems in this chapter will be used there.*

By one-sided systems we mean systems of the form

$$Ax = b, \tag{3.1}$$

where $A \in \mathbb{R}^{m\times n}, x \in \mathbb{R}^n$ and $b \in \mathbb{R}^m$. It is known (see [25] and [35]) that a solution to (3.1) exists if and only if $\overline{x}$ is a solution, where

$$(\forall j)\, \overline{x}_j := \bigoplus_{i\in M}{}^{\prime} \left(b_i a_{ij}^{-1}\right). \tag{3.2}$$

We are interested here in the similar system

$$Ax = b\,(\alpha)\,, \tag{3.3}$$

where the parameter $\alpha \in [\underline{\alpha}, \overline{\alpha}]$ for some real $\underline{\alpha} < \overline{\alpha}$ and b is of the form

$$b\,(\alpha) = \begin{pmatrix} \alpha \\ \vdots \\ \alpha \\ 0 \\ \vdots \\ 0 \end{pmatrix} \begin{matrix} \left.\begin{matrix} \\ \\ \\ \end{matrix}\right\} k \\ \left.\begin{matrix} \\ \\ \\ \end{matrix}\right\} m-k \end{matrix} .$$

That is, $(\forall i)\,(1 \leq i \leq k)\, b_i = \alpha$ and $(\forall i)\,(k+1 \leq i \leq m)\, b_i = 0$, for some $1 \leq k \leq m$.

We define $K := \{1, 2, \ldots, k\}$. Also, define $A[K]$ to be the matrix A restricted to the rows from K, similarly $b[K]$.

As in the non-parameter version, the vector $\overline{x}$ defined by (3.2) is key but now $\overline{x} = \overline{x}\,(\alpha)$. It is still true that system (3.3) has a solution if and only if $\overline{x}\,(\alpha)$ is a solution. As a result, a key question now is "for what values of α, if any, in the interval $[\underline{\alpha}, \overline{\alpha}]$ is the vector $\overline{x}\,(\alpha)$

a solution to the system $Ax = b(\alpha)$?" Note it is also known (see [25] and [35]) that

$$(\forall \alpha)\, A\overline{x}(\alpha) \leq b(\alpha)\,. \tag{3.4}$$

Clearly, a necessary condition for the existence of a solution to the parametrised system is the existence of a solution to the sub-system $A[K]\,x = b[K]$, or equivalently, the system

$$A[K]\,x = (\alpha, \alpha, \ldots, \alpha)^T\,. \tag{3.5}$$

It will be clear in later work why this obvious point is worth mentioning.

Let $j \in N$, then

$$\begin{aligned}
\overline{x}_j(\alpha) &= \sideset{}{'}\bigoplus_{i\in M} \left(b_i a_{ij}^{-1}\right) \\
&= \sideset{}{'}\bigoplus_{i\in K} \left(\alpha a_{ij}^{-1}\right) \oplus' \sideset{}{'}\bigoplus_{i\in M\backslash K} \left(a_{ij}^{-1}\right) \\
&= \left(\alpha \sideset{}{'}\bigoplus_{i\in K} a_{ij}^{-1}\right) \oplus' \left(\sideset{}{'}\bigoplus_{i\in M\backslash K} a_{ij}^{-1}\right).
\end{aligned}$$

For notational purposes, denote

$$(\forall j \in N)\, \phi_j := \sideset{}{'}\bigoplus_{i\in K} a_{ij}^{-1} \tag{3.6}$$

and

$$(\forall j \in N)\, \pi_j := \sideset{}{'}\bigoplus_{i\in M\backslash K} a_{ij}^{-1}, \tag{3.7}$$

so that

$$(\forall j \in N)\, \overline{x}_j(\alpha) = (\alpha\phi_j) \oplus' \pi_j. \tag{3.8}$$

3.3 The subsystem $A[K]x = (\alpha, \ldots, \alpha)^T$

Recall (3.5) and let $i \in K$. Then, by (3.8)

$$
\begin{aligned}
(A\overline{x}(\alpha))_i &= \bigoplus_{j \in N} (a_{ij}\overline{x}_j(\alpha)) \\
&= \bigoplus_{j \in N} (a_{ij}((\alpha\phi_j) \oplus' \pi_j)) \\
&= \bigoplus_{j \in N} ((a_{ij}(\alpha\phi_j)) \oplus' (a_{ij}\pi_j)) \\
&= \bigoplus_{j \in N} ((\alpha a_{ij}\phi_j) \oplus' (a_{ij}\pi_j)). \qquad (3.9)
\end{aligned}
$$

Recall that if we wish to solve the subsystem (3.5), then we require $(Ax)_i = \alpha$ for $i \in K$. Note also, since $i \in K$, that for each $j \in N$ we have

$$a_{ij}\phi_j = a_{ij} \bigoplus_{u \in K}' a_{uj}^{-1} \leq 0,$$

(since $\bigoplus'_{u \in K} a_{uj}^{-1} \leq a_{ij}^{-1}$) and equality holds if and only if $\bigoplus'_{u \in K} a_{uj}^{-1} = a_{ij}^{-1}$, which holds if and only if

$$\bigoplus_{u \in K} a_{uj} = a_{ij}. \qquad (3.10)$$

Note that if inequality holds strictly for every $j \in N$, then

$$
\begin{aligned}
& (\forall j \in N)\, a_{ij}\phi_j < 0 \\
\Rightarrow & (\forall j \in N)\, \alpha a_{ij}\phi_j < \alpha \\
\Rightarrow & (\forall j \in N)\, (\alpha a_{ij}\phi_j) \oplus' (a_{ij}\pi_j) < \alpha \\
\Rightarrow & \bigoplus_{j\in N} ((\alpha a_{ij}\phi_j) \oplus' (a_{ij}\pi_j)) < \alpha \\
\Rightarrow & (A\overline{x}(\alpha))_i < \alpha.
\end{aligned}
$$

We conclude in this case that no solution to the subsystem (3.5) (and so also to the system (3.3)) exists. This gives us the following Lemma.

Lemma 3.2. *If the subsystem $A[K]x = \alpha$ has a solution, then for each $i \in K$ there exists $j \in N$ such that $\bigoplus_{u\in K} a_{uj} = a_{ij}$.*

In terms of the set covering problem, the above necessary condition is equivalent to saying that the column maxima of A (restricted to the rows from K) cover K.

In view of Lemma 3.2, we assume from now on that for each $i \in K$, there exists $j \in N$ such that (3.10) holds. That is, N_i is non-empty, where:

$$(\forall i \in K)\, N_i := \left\{ j \in N : a_{ij} = \bigoplus_{u\in K} a_{uj} \right\},$$

so that

$$(\forall j \in N_i) \bigoplus_{u\in K} a_{uj} = a_{ij}$$

and

$$(\forall j \in N \backslash N_i) \bigoplus_{u\in K} a_{uj} > a_{ij}.$$

Now let $i \in K$ be fixed again and let $j \in N_i$ (we can do this since we are assuming now that N_i is non-empty). Then $\alpha a_{ij}\phi_j = \alpha$ by (3.6) and (3.3). Note that if $a_{ij}\pi_j \geq \alpha$ also, then $(\alpha a_{ij}\phi_j) \oplus' (a_{ij}\pi_j) = \alpha$ and

$$\begin{aligned}
\alpha &\geq (A\overline{x}(\alpha))_i \quad \text{(by (3.4))} \\
&= \bigoplus_{t\in N} ((\alpha a_{it}\phi_t) \oplus' (a_{it}\pi_t)) \quad \text{(by (3.9))} \\
&\geq (\alpha a_{ij}\phi_j) \oplus' (a_{ij}\pi_j) \\
&= \alpha,
\end{aligned}$$

so $(A\overline{x}(\alpha))_i = \alpha$, as required.

So, we conclude that if there exists $j \in N_i$ such that $a_{ij}\pi_j \geq \alpha^*$ for some $\alpha^* \in [\underline{\alpha}, \overline{\alpha}]$, then $\overline{x}(\alpha^*)$ solves the ith equation. Equivalently, if there exists $j \in N_i$ such that $a_{ij}\pi_j \geq \underline{\alpha}$, then $\overline{x}(\alpha)$ solves the ith equation for all $\alpha \in [\underline{\alpha}, a_{ij}\pi_j] \cap [\underline{\alpha}, \overline{\alpha}] = [\underline{\alpha}, (a_{ij}\pi_j) \oplus' \overline{\alpha}]$.

Conversely, if $(\forall j \in N_i)\, a_{ij}\pi_j < \underline{\alpha}$, then

$$\begin{aligned}
(A\overline{x}(\alpha))_i &= \bigoplus_{j\in N_i} \left((\alpha a_{ij}\phi_j) \oplus' \overbrace{(a_{ij}\pi_j)}^{<\underline{\alpha}\leq\alpha} \right) \oplus \bigoplus_{j\notin N_i} \left(\overbrace{(\alpha a_{ij}\phi_j)}^{<\alpha \text{ since } j\notin N_i} \oplus' (a_{ij}\pi_j) \right) \\
&< \alpha \oplus \alpha \\
&= \alpha,
\end{aligned}$$

and so no solution exists for the ith equation. We conclude with the following Lemma.

Lemma 3.3. *Let $i \in K$. Then there exists a solution to the ith equation if and only if $N_i \neq \emptyset$ and there exists $j \in N_i$ such that $a_{ij}\pi_j \geq \underline{\alpha}$.*

A couple of remarks:

Remark 3.4. *Suppose $j \in N_i$ is such that $a_{ij}\pi_j < \underline{\alpha}$. Then the quantity $(a_{ij}\pi_j) \oplus' \overline{\alpha} = a_{ij}\pi_j < \underline{\alpha}$.*

Remark 3.5. *If there are multiple $j \in N_i$ satisfying $a_{ij}\pi_j \geq \underline{\alpha}$, then we want to choose a j which maximises the size of the solution interval for α.*

By combining Remarks 3.4 and 3.5, we see that the set of α for which $\overline{x}(\alpha)$ solves the ith equation is

$$\left[\underline{\alpha}, \overline{\alpha} \oplus' \bigoplus_{j \in N_i} (a_{ij}\pi_j)\right].$$

Note that the above interval is empty if and only if $(\forall j \in N_i)\, a_{ij}\pi_j < \underline{\alpha}$, as expected.

Finally, we wish to solve the ith equation for all $i \in K$. We should then see that we require

$$\begin{aligned}
&(\forall i \in K)\, \alpha \leq \overline{\alpha} \oplus' \bigoplus_{j \in N_i} (a_{ij}\pi_j) \\
\Leftrightarrow &\alpha \leq \bigoplus_{i \in K}{}' \left(\overline{\alpha} \oplus' \bigoplus_{j \in N_i} (a_{ij}\pi_j)\right) \\
\Leftrightarrow &\alpha \leq \overline{\alpha} \oplus' \bigoplus_{i \in K}{}' \left(\bigoplus_{j \in N_i} (a_{ij}\pi_j)\right).
\end{aligned}$$

We summarise our results in the following Lemma.

Lemma 3.6. *The vector $\overline{x}(\alpha)$ solves the subsystem $A[K]\, x = \alpha$ if and only if $(\forall i \in K)\, N_i \neq \emptyset$ and*

$$\underline{\alpha} \leq \alpha \leq \overline{\alpha} \oplus' \bigoplus_{i \in K}{}' \left(\bigoplus_{j \in N_i} (a_{ij}\pi_j)\right).$$

In fact, since by definition we have $\max \emptyset = \epsilon$ and $\underline{\alpha} < \overline{\alpha}$, we also have

Lemma 3.7. *The subsystem $A[K]\, x = \alpha$ has a solution if and only if*

$$\bigoplus_{i \in K}{}' \left(\bigoplus_{j \in N_i} (a_{ij}\pi_j)\right) \geq \underline{\alpha}.$$

3.4 The subsystem $A[M\backslash K]\,x = b[M\backslash K]$

Let $i \in M\backslash K$. Then, by (3.8),

$$(A\overline{x}(\alpha))_i = \bigoplus_{j\in N} ((\alpha a_{ij}\phi_j) \oplus' (a_{ij}\pi_j)).$$

Recall we require (since $i \in M\backslash K$) that

$$(A\overline{x}(\alpha))_i = 0.$$

Note also, since $i \in M\backslash K$ and using (3.7), that for each $j \in N$,

$$a_{ij}\pi_j = a_{ij} \bigoplus_{u\in M\backslash K}{}' a_{uj}^{-1} \leq a_{ij}a_{ij}^{-1} = 0$$

and that equality holds if and only if $a_{ij}^{-1} = \bigoplus'_{u\in M\backslash K} a_{uj}^{-1}$, which holds if and only if

$$a_{ij} = \bigoplus_{u\in M\backslash K} a_{uj}. \tag{3.11}$$

Note if inequality holds strictly for each $j \in N$, then

$$(\forall j \in N)\,(\alpha a_{ij}\phi_j) \oplus' (a_{ij}\pi_j) \leq a_{ij}\pi_j < 0$$

and so

$$\bigoplus_{j\in N} ((\alpha a_{ij}\phi_j) \oplus' (a_{ij}\pi_j)) < 0.$$

In this case, it follows that the ith equation is not solved for any $\alpha \in [\underline{\alpha}, \overline{\alpha}]$.

As a result, we assume that for each $i \in M\backslash K$ there exists $j \in N$ such that (3.11) holds. That is, N_i is non-empty, where

$$(\forall i \in M\backslash K)\, N_i := \left\{ j \in N : a_{ij} = \bigoplus_{u \in M\backslash K} a_{uj} \right\}.$$

It follows that a necessary condition to solve the ith equation for all $i \in M\backslash K$ is that the column maxima in A (restricted to the rows of $M\backslash K$) cover $M\backslash K$. We state this in a Lemma.

Lemma 3.8. *If the subsystem $A[M\backslash K]\,x = 0$ has a solution, then $N_i \neq \emptyset$ for each $i \in M\backslash K$.*

Let $i \in M\backslash K$ be fixed again. Note that $(\forall j \in N_i)\, a_{ij} = \overbrace{\bigoplus_{u \in M\backslash K} a_{uj}}^{=\pi_j^{-1}} \Leftrightarrow a_{ij}\pi_j = 0$, and $(\forall j \notin N_i)\, a_{ij} < \bigoplus_{u \in M\backslash K} a_{uj} \Leftrightarrow a_{ij}\pi_j < 0$.

Let $j \in N_i$, then

$$a_{ij}\pi_j = 0.$$

Note that if $\alpha a_{ij}\phi_j \geq a_{ij}\pi_j\,(=0)$ which holds if and only if

$$\alpha \geq a_{ij}^{-1}\phi_j^{-1},$$

then

$$\begin{aligned} &(\alpha a_{ij}\phi_j) \oplus' (a_{ij}\pi_j) = a_{ij}\pi_j = 0 \\ \Rightarrow\, &(A\overline{x}\,(\alpha))_i = 0, \end{aligned}$$

as required.

We conclude, if there exists $j \in N_i$ such that $a_{ij}^{-1}\phi_j^{-1} \leq \alpha^*$ for some $\alpha^* \in [\underline{\alpha}, \overline{\alpha}]$, then $\overline{x}\,(\alpha)$ solves the ith equation. Equivalently, if there exists $j \in N_i$ such that $a_{ij}^{-1}\phi_j^{-1} \leq$

$\overline{\alpha}$, then $\overline{x}(\alpha)$ is a solution for the ith equation for all $\alpha \in \left[a_{ij}^{-1}\phi_j^{-1}, \overline{\alpha}\right] \cap [\underline{\alpha}, \overline{\alpha}] = \left[\underline{\alpha} \oplus \left(a_{ij}^{-1}\phi_j^{-1}\right), \overline{\alpha}\right]$.

Conversely, if $(\forall j \in N_i)\, a_{ij}^{-1}\phi_j^{-1} > \overline{\alpha}\, (\Rightarrow \alpha a_{ij}\phi_j < 0)$, then

$$\begin{aligned}(A\overline{x}(\alpha))_i &= \bigoplus_{j \in N_i} \left(\overbrace{(\alpha a_{ij}\phi_j)}^{<0 \text{ by assumption}} \oplus' (a_{ij}\pi_j) \right) \oplus \bigoplus_{j \notin N_i} \left((\alpha a_{ij}\phi_j) \oplus' \overbrace{(a_{ij}\pi_j)}^{<0 \text{ since } j \notin N_i} \right) \\ &< 0 \oplus 0 \\ &= 0\end{aligned}$$

and so no equality holds.

In conclusion, for $i \in M \backslash K$, there exists $\alpha \in [\underline{\alpha}, \overline{\alpha}]$ such that $\overline{x}(\alpha)$ solves the ith equation if and only if $N_i \neq \emptyset$ and there exists $j \in N_i$ such that $a_{ij}^{-1}\phi_j^{-1} \leq \alpha \leq \overline{\alpha}$. We summarise with the following Lemma.

Lemma 3.9. *Let $i \in M \backslash K$. There exists a solution to the ith equation if and only if $N_i \neq \emptyset$ and there exists $j \in N_i$ such that $a_{ij}^{-1}\phi_j^{-1} \leq \overline{\alpha}$.*

We have a couple of remarks.

Remark 3.10. *Suppose $j \in N_i$ such that $a_{ij}^{-1}\phi_j^{-1} > \overline{\alpha}$. Then $\underline{\alpha} \oplus \left(a_{ij}^{-1}\phi_j^{-1}\right) = a_{ij}^{-1}\phi_j^{-1} > \overline{\alpha}$.*

Remark 3.11. *If there are multiple $j \in N_i$ such that $a_{ij}^{-1}\phi_j^{-1} \leq \overline{\alpha}$, then we want to choose a j which maximises the size of the solution interval for α.*

Combining Remarks 3.10 and 3.11 we see that the set of α for which $\overline{x}(\alpha)$ solves the ith equation is

$$\left[\underline{\alpha} \oplus \bigoplus_{j \in N_i}{}' \left(a_{ij}^{-1}\phi_j^{-1}\right), \overline{\alpha}\right].$$

Note that the above interval is empty if and only if $(\forall j \in N_i)\, a_{ij}^{-1}\phi_j^{-1} > \overline{\alpha}$.

Finally, we want to solve the ith equation for all $i \in M \backslash K$. We should see that we require

$$
\begin{aligned}
&(\forall i \in M \backslash K)\, \alpha \geq \underline{\alpha} \oplus \sideset{}{'}\bigoplus_{j \in N_i} \left(a_{ij}^{-1} \phi_j^{-1}\right) \\
\Leftrightarrow &\alpha \geq \bigoplus_{i \in M \backslash K} \left(\underline{\alpha} \oplus \sideset{}{'}\bigoplus_{j \in N_i} \left(a_{ij}^{-1} \phi_j^{-1}\right) \right) \\
\Leftrightarrow &\alpha \geq \underline{\alpha} \oplus \bigoplus_{i \in M \backslash K} \left(\sideset{}{'}\bigoplus_{j \in N_i} \left(a_{ij}^{-1} \phi_j^{-1}\right) \right).
\end{aligned}
$$

We summarise our results in the following Lemma.

Lemma 3.12. *The vector* $\overline{x}(\alpha)$ *solves the subsystem* $A[M \backslash K]\, x = 0$ *if and only if* $(\forall i \in M \backslash K)\, N_i \neq \emptyset$ *and*

$$
\underline{\alpha} \oplus \bigoplus_{i \in M \backslash K} \left(\sideset{}{'}\bigoplus_{j \in N_i} \left(a_{ij}^{-1} \phi_j^{-1}\right) \right) \leq \alpha \leq \overline{\alpha}.
$$

In fact, since by definition we have $\max \emptyset = \epsilon$ and by assumption $\underline{\alpha} < \overline{\alpha}$, we have the following.

Lemma 3.13. *The subsystem* $A[M \backslash K]\, x = 0$ *has a solution if and only if*

$$
\bigoplus_{i \in M \backslash K} \left(\sideset{}{'}\bigoplus_{j \in N_i} \left(a_{ij}^{-1} \phi_j^{-1}\right) \right) \leq \overline{\alpha}.
$$

3.5 The whole system $Ax = b(\alpha)$

We are now at the stage where we can solve the entire system by simply taking the intersection of the intervals from Lemmas 3.7 and 3.13. This intersection is

$$\left[\underline{\alpha} \oplus \bigoplus_{i \in M \backslash K} \left({\bigoplus_{j \in N_i}}' \left(a_{ij}^{-1} \phi_j^{-1}\right) \right), \overline{\alpha} \oplus' {\bigoplus_{i \in K}}' \left(\bigoplus_{j \in N_i} \left(a_{ij} \pi_j\right) \right) \right].$$

We summarise with our Theorem.

Theorem 3.14. *The parametrised system $Ax = b(\alpha)$ has a solution if and only if*

$$\underline{\alpha} \oplus \bigoplus_{i \in M \backslash K} \left({\bigoplus_{j \in N_i}}' \left(a_{ij}^{-1} \phi_j^{-1}\right) \right) \leq \alpha \leq \overline{\alpha} \oplus' {\bigoplus_{i \in K}}' \left(\bigoplus_{j \in N_i} \left(a_{ij} \pi_j\right) \right),$$

which can be checked in $O(mn)$ time.

Example 3.15. *Let*

$$A = \begin{pmatrix} 2 & 1 & 3 & 1 \\ 1 & 3 & 1 & 2 \\ 0 & 2 & 3 & 1 \\ 1 & 1 & 2 & 2 \end{pmatrix}$$

and consider the one-sided parametrised system

$$\begin{pmatrix} 2 & 1 & 3 & 1 \\ 1 & 3 & 1 & 2 \\ 0 & 2 & 3 & 1 \\ 1 & 1 & 2 & 2 \end{pmatrix} x = \begin{pmatrix} \alpha \\ \alpha \\ 0 \\ 0 \end{pmatrix}.$$

We have $K = \{1, 2\}$ and $M \backslash K = \{3, 4\}$. Further, by (3.6) and (3.7), $\phi_1 = -2, \phi_2 = -3, \phi_3 = -3, \phi_4 = -2, \pi_1 = -1, \pi_2 = -2, \pi_3 = -3$ and $\pi_4 = -2$. We calculate the *following so we may apply Theorem 3.14:*

$\max_{j \in N_1} (a_{1j}\pi_j) = \max(1, 0) = 1,$

$\max_{j \in N_2} (a_{2j}\pi_j) = \max(1, 0) = 1,$

$\min_{j \in N_3} \left(a_{3j}^{-1}\phi_j^{-1}\right) = \min(1, 0) = 0,$

$\min_{j \in N_4} \left(a_{4j}^{-1}\phi_j^{-1}\right) = \min(1, 0) = 0.$

Applying Theorem 3.14 we find $Ax = b(\alpha)$ *has a solution if and only if* $\max(0, 0) \leq \alpha \leq \min(1, 1)$ *if and only if* $0 \leq \alpha \leq 1$.

3.6 Summary

We have studied the $m \times n$ parametrised system $Ax = b(\alpha)$ for the finite matrix A and vector b with a special structure (refer to (3.2)). By considering the two arising subsystems separately, we are able (in $O(mn)$ time) to explicitly give the set of α, restricted to the given interval $[\underline{\alpha}, \overline{\alpha}]$, for which the parametrised system $Ax = b(\alpha)$ has a solution. Further, due to known results on one-sided systems, for any such α, we can describe the full set of solutions.

4. Two-sided, homogeneous systems of inequalities - two strongly polynomial solution methods if B has exactly one finite entry per row

4.1 Introduction

We present a strongly polynomial method for solving two-sided systems of inequalities

$$Ax \leq Bx, \tag{4.1}$$

where $A \in \mathbb{R}^{m \times n}$ and $B \in \overline{\mathbb{R}}^{m \times n}$ has exactly one finite entry per row. Necessarily, $m \geq n$. We give an $O(mn + n^3)$ method (based on the sub-eigenvector problem) called the *aggregation method*. We will see also that the aggregation method finds a set of generators for the solution set of (4.1).

The aggregation method essentially takes every inequality and rewrites each as a set of inequalities, each comparing exactly two variables. The problem therefore is transformed to a system of dual inequalities, which is essentially the sub-eigenvector problem and a set of generators for the solution set can be found in strongly polynomial time.

We note here that we may assume without loss of generality that B has no ϵ columns (that is, B is column $\mathbb{R}$-astic). To see this, let $j \in N$ and suppose $B_j = \epsilon$, where B_j denotes column j of B. Note now that the x_j term in any vector x does not affect the vector Bx and the vector Ax is component-wise non-increasing for decreasing x_j. Since we are looking for x such that $Ax \leq Bx$, it follows that we can simply choose x_j to be sufficiently small, or even set $x_j = \epsilon$, effectively removing columns A_j and B_j from our system. A consequence of this note is that we also have $m \geq n$ (this follows by the pigeonhole principle).

4.2 The strongly polynomial "aggregation method" for converting to the sub-eigenvector problem

For all $j \in N$ define

$$R_j := \{i \in M : b_{ij} \in \mathbb{R}\}.$$

That is, R_j is the set of rows i such that b_{ij} is finite.

Note that, by our assumed form of B, that each R_j is non-empty and that $\bigcup_{j \in N} R_j = M$. Let $j \in N$, we have

$$(\forall i \in R_j)\, a_{i1}x_1 \oplus \cdots \oplus a_{in}x_n \leq x_j. \tag{4.2}$$

Since all the right-hand sides of (4.2) are the same, we can apply the method of aggregation to realise that (4.2) holds if and only if

$$\begin{aligned}
&\bigoplus_{i\in R_j} a_{i1}x_1 \oplus \cdots \oplus a_{in}x_n \le x_j \\
\Leftrightarrow &\bigoplus_{i\in R_j}\bigoplus_{t\in N} a_{it}x_t \le x_j \\
\Leftrightarrow &\bigoplus_{t\in N}\bigoplus_{i\in R_j} a_{it}x_t \le x_j. \qquad (4.3)
\end{aligned}$$

We have reduced the system of inequalities $A[R_j]\,x \le B[R_j]\,x, j \in N$ to a single inequality (4.3). It is clear now that the system $Ax \le Bx$ can be reduced (by considering (4.3) for all distinct R_j) to an equivalent $n \times n$ system $A'x \le B'x$. Note that B' still has the property that each row contains exactly one finite entry (since B' is obtained by deleting rows of B) and, in addition, B' has the property that every column contains exactly one finite entry. This new property follows from the reduction of multiple inequalities of the same type to a single one. The matrix B' is, by definition, a permutation matrix and so its inverse B'^{-1} exists. It is now easy to convert system (4.1) to one of the form $Cx \le x$, where $C = B'^{-1}A$ is an $n \times n$ matrix. We have reduced the system (4.1) to the sub-eigenvector problem, for which there exist $O(n^3)$ methods to find a set of generators of the solution set [25].

The main step in the reduction process is converting the system of inequalities (4.2) for $i \in R_j$ (of size $O(m)$) to a single inequality (4.3) and repeating this for all $j \in N$ ($O(n)$), so the reduction process takes $O(mn)$ time. It follows that we can find a set of generators for (4.1) in $O(mn + n^3)$ time.

The following very simple example shows the reduction process working.

Example 4.1. *Consider the system*

$$\begin{pmatrix} -1 & 4 \\ 0 & 1 \\ -6 & -2 \end{pmatrix} x \le \begin{pmatrix} 0 & \epsilon \\ 0 & \epsilon \\ \epsilon & 0 \end{pmatrix} x.$$

The first two inequalities can be combined (both contain an x_1 term on the right-hand side). We get the reduced system

$$\begin{pmatrix} 0 & 4 \\ -6 & -2 \end{pmatrix} x \le \begin{pmatrix} 0 & \epsilon \\ \epsilon & 0 \end{pmatrix} x.$$

In this example, B' is already the identity matrix and we are left with the sub-eigenvector problem

$$\begin{pmatrix} 0 & 4 \\ -6 & -2 \end{pmatrix} x \le x.$$

5. Two-sided homogeneous systems of equations - strongly polynomial method if matrices have exactly two finite entries per row, appearing in the same position

5.1 Introduction

We consider here the two-sided system $Ax = Bx$ where $A, B \in \overline{\mathbb{R}}^{m \times n}$ each have exactly two finite entries per row, appearing in the same position in A and B.

By making use of the Cancellation Rule (Lemma 2.7), we show that the two-sided system can be converted to an equivalent sub-eigenvector problem in strongly polynomial time. It follows that we can find a set of generators for the solution set in strongly polynomial time.

5.2 Problem formulation

We consider two-sided systems of equations in max-algebra of the form

$$Ax = Bx \tag{5.1}$$

for matrices $A, B \in \overline{\mathbb{R}}^{m\times n}$, each having exactly two finite entries per row according to the rule $(\forall i)\,(\forall j)\; a_{ij} \in \mathbb{R}$ if and only if $b_{ij} \in \mathbb{R}$. It should be stressed that we aim to find a non-trivial solution $x \neq \epsilon$ which is not necessarily finite. Indeed, in the method presented we may set some components x_j equal to ϵ. The ith equation is

$$a_{ij_i}x_{j_i} \oplus a_{ik_i}x_{k_i} = b_{ij_i}x_{j_i} \oplus b_{ik_i}x_{k_i}, \tag{5.2}$$

for some $j_i, k_i \in N$ and $a_{ij_i}, a_{ik_i}, b_{ij_i}, b_{ik_i} \in \mathbb{R}$.

An example illustrates the assumed form of matrices A and B.

Example 5.1.

$$\begin{pmatrix} 1 & & 2 & & \\ 2 & & 0 & & \\ & 2 & & 2 & \\ & 3 & & & 1 \end{pmatrix} x = \begin{pmatrix} 0 & & 1 & & \\ 1 & & 2 & & \\ & 0 & & 0 & \\ & 2 & & & 1 \end{pmatrix} x.$$

Solving this small system reveals that the solution set S is given by

$$S = \left\{(\epsilon, \epsilon, \epsilon, \epsilon, c) : c \in \overline{\mathbb{R}}\right\}$$

and so no finite solution exists.

We now make some remarks which allow us to remove certain rows and columns of matrices A and B in (5.1), reducing (5.1) to a desirable special form, from which we can find a solution.

5.3 Remarks

We make some important remarks which will allow us to simplify our problem. Recall the Cancellation Rule 2.7, page 18.

Remark 5.2. *If for any equation, the* i*th equation (5.2), say, there are exactly two cancellations (on opposite sides), then (after cancelling) we obtain an equality of the form*

$$a_{ij}x_j = b_{ik}x_k$$

for some j *and* k*. We may then remove the column* k *(say) from both sides of (5.1) (that is, remove column* k *from both* A *and* B*), writing all instances of* x_k *in terms of* x_j *and incorporating these new coefficients into the* j *columns of* A *and* B*. Having done this we may also remove row* i*, as the* i*th equation (5.2) is now solved by simply re-introducing* x_k *later. (Note we may now have that some rows of* A *(and so also* B*) have less than two finite entries, this point is addressed later).*

Remark 5.3. *If for any equation, the* i*th equation (5.2), say, we have two cancellations happening on the same side, then (after cancelling) we obtain an equality of the form*

$$a_{ij}x_j \oplus a_{ik}x_k = \epsilon$$

for some j *and* k*. It follows (since* a_{ij}, a_{ik} *are finite) that we set* $x_j = x_k = \epsilon$*. Having done this, we may remove the* i*th equation (5.2) as this is now satisfied and remove also columns* j *and* k *from both* A *and* B *in (5.1). (Again, some rows of* A*, and so also of* B*, may now have strictly less than two finite entries).*

Remark 5.4. *If for any equation, the* i*th equation (5.2), say, we have no cancellations* *happening, then we have*

$$a_{ij}x_j \oplus a_{ik}x_k = a_{ij}x_j \oplus a_{ik}x_k$$

for some j and k. We then remove the ith equation (5.2) (that is, row i from both A and B) as this is trivially satisfied.

Remark 5.5. *We assume now that matrices A and B in (5.1) have been reduced according to Remarks 5.2, 5.3 and 5.4. Note that these Remarks only deal with equations in which none or exactly two cancellations happen. Equations for which exactly one cancellation occurs have not yet been mentioned.*

As noted in Remarks 5.2 and 5.3, for each i we now have that row i of matrix A (and so also matrix B) has at most two finite entries in (5.1). (Note that by Remarks 5.2, 5.3 and 5.4, for each i matrices A and B will have exactly the same number of finite entries in row i, this number may be zero, one or two).

If row i has no finite entries we may simply remove row i, since $\epsilon = \epsilon$ is trivially satisfied.

If row i has exactly one finite entry (in both A and B), then the ith equation will read

$$a_{ij}x_j = b_{ij}x_j$$

for some j. If $a_{ij} = b_{ij}$, then again we may remove row i (from A and B) since the ith equation is trivially satisfied. Otherwise, we set $x_j = \epsilon$ and remove row i and column j (from A and B).

Finally, if row i has exactly two finite entries, then it must be that exactly one cancellation is due to happen here.

Remark 5.6 (conclusion)**.** *Based on Remarks 5.2, 5.3, 5.4 and 5.5, we may assume without loss of generality that (5.1) is a system in which every row of A and B has exactly* *two finite entries appearing in the same position and exactly one cancellation happens per equation.*

An example helps make clear the reduction process.

Example 5.7. *Consider the* 5×4 *system*

$$\begin{pmatrix} 5 & 1 & & \\ & 2 & & 1 \\ 1 & & 0 & \\ & & 3 & 0 \\ & 3 & & 4 \end{pmatrix} x = \begin{pmatrix} 0 & 2 & & \\ & 1 & & 1 \\ 0 & & 0 & \\ & & 2 & 4 \\ & 2 & & 4 \end{pmatrix} x.$$

In the first equation, we have two cancellations, happening on opposite sides. We have the equality $5x_1 = 2x_2 \Leftrightarrow x_2 = 3x_1$*. Making this substitution throughout and removing the first equation and second column yields the following system, where columns correspond to original variables* x_1, x_3, x_4*.*

$$\begin{pmatrix} 5 & & 1 \\ 1 & 0 & \\ & 3 & 0 \\ 6 & & 4 \end{pmatrix} x' = \begin{pmatrix} 4 & & 1 \\ 0 & 0 & \\ & 2 & 4 \\ 5 & & 4 \end{pmatrix} x'.$$

From the third equation of the reduced system, we have two cancellations happening on opposite sides. We have

$$3x_3 = 4x_4 \Leftrightarrow x_4 = (-1)x_3.$$

Applying the same procedure as in the previous step we now reach the final reduced form

$$\begin{pmatrix} 5 & 0 \\ 1 & 0 \\ 6 & 3 \end{pmatrix} x'' = \begin{pmatrix} 4 & 0 \\ 0 & 0 \\ 5 & 3 \end{pmatrix} x'',$$

where columns correspond to original variables x_1, x_3. Notice how each row has exactly two finite entries and exactly one cancellation occurs in each row. This is our final form, we assume this form without loss of generality for the remainder of this chapter.

Below is an example of a system in its reduced form, showing that our final form in general will not necessarily have only two $\mathbb{R}$-astic columns in A and B.

Example 5.8.

$$\begin{pmatrix} 1 & & 0 \\ 0 & & 1 \\ 1 & 1 & \end{pmatrix} x = \begin{pmatrix} 1 & & 3 \\ 3 & & 1 \\ 0 & 1 & \end{pmatrix} x.$$

5.4 Systems of inequalities and the sub-eigenvector problem

From this point we assume without loss of generality that (5.1) is a system for which in each row there are exactly two finite entries in the matrices A and B, appearing in the same position and with the extra property that exactly one cancellation occurs per row (per equation).

5.4.1 Algorithm A1

Algorithm A1 :

Input : Matrices $A, B \in \overline{\mathbb{R}}^{m\times n}$, where each row of A and B have exactly two finite entries, appearing in the same position. Also, in each row there is exactly one cancellation.

Output : Answer to the question of finite solvability of the system $Ax = Bx$ and a set of generators of the solution set in the affirmative case.

(**I**) Define $(\forall i \in N)(\forall j \in N), i \neq j$

$$I_{ij} := \{k \in M : a_{ki}, b_{ki}, a_{kj}, b_{kj} \in \mathbb{R}, a_{ki} = b_{ki}, a_{kj} \neq b_{kj}\}.$$

(**II**) Define $(\forall i \in N)(\forall j \in N), i \neq j$

$$c_{ij} := \bigoplus_{k \in I_{ij}} (\max(a_{kj}, b_{kj}) - a_{ki}).$$

(**III**) Define $(\forall i \in N)\, c_{ii} := \epsilon$.

(**IV**) If $\lambda(C) > 0$, then identify a positive cycle $(i_0, i_1, \ldots, i_r = i_0)$, some r, in C and go to (**V**). Else, $\lambda(C) \leq 0$ and go to (**VI**).

(**V**) Set $x_{i_s} := \epsilon$ for $s = 0, \ldots, r-1$ and update C by deleting rows/columns i_s of C. If this deletes all rows/columns of C, then STOP and return "there is no non-trivial solution", else, go to (**IV**).

(**VI**) STOP and return "x is a solution if and only if $x = C^* u, u \in \mathbb{R}^{m \times n}$".

Theorem 5.9. *The algorithm A1 is correct and terminates in $O(mn^2 + n^4)$ time.*

Proof. Let $A, B \in \overline{\mathbb{R}}^{m \times n}$ be as in the input of Algorithm A1. Let $i, j \in N, i \neq j$. Define

$$I_{ij} := \{k \in M : a_{ki}, b_{ki}, a_{kj}, b_{kj} \in \mathbb{R}, a_{ki} = b_{ki}, a_{kj} \neq b_{kj}\}.$$

I_{ij} is the set of equations in which variables i and j appear together with finite coefficients and the cancellation happens with the coefficients of x_j.

Now, let $k \in I_{ij}$ for some $i \neq j$. Then equation k:

$$a_{ki}x_i \oplus \max(a_{kj}, b_{kj})\, x_j = a_{ki}x_i.$$

Now, equation k is satisfied if and only if

$$\max(a_{kj}, b_{kj})\, x_j \le a_{ki}x_i,$$

which holds if and only if

$$x_i - x_j \ge \max(a_{kj}, b_{kj}) - a_{ki}.$$

Since this is true for all $k \in I_{ij}$, we have

$$x_i - x_j \ge \bigoplus_{k \in I_{ij}} (\max(a_{kj}, b_{kj}) - a_{ki}) = c_{ij},$$

as defined in A1.

Note that for all $i, x_i - x_i \ge \epsilon$ is trivially satisfied.

The two-sided system $Ax = Bx$ is therefore equivalent to

$$(\forall i)\,(\forall j)\, x_i - x_j \ge c_{ij},$$

which (in max-algebraic notation) is equivalent to

$$(\forall i \in N) \bigoplus_{j \in N} (c_{ij}x_j) \le x_i,$$

or, in the compact form,

$$Cx \le x.$$

This is the subeigenvector problem for $\lambda = 0$.

It is known ([25], Theorem 1.6.18(b)) that if $\lambda(C) \leq 0$, then $Cx \leq x, x \in \mathbb{R}^n$ if and only if $x = C^* u, u \in \mathbb{R}^n$.

If $\lambda(C) > 0$, it is then possible to find such a positive cycle in the associated digraph D_C. Suppose we have the positive cycle $(i_0, i_1, \ldots, i_{r-1}, i_r = i_0)$, for some r. Then we have

$$\begin{cases} x_{i_0} \geq c_{i_0 i_1} x_{i_1} \\ x_{i_1} \geq c_{i_1 i_2} x_{i_2} \\ \vdots \\ x_{i_{r-1}} \geq c_{i_{r-1} i_0} x_{i_0}. \end{cases} \tag{5.3}$$

Putting together the inequalities from (5.3) we have

$$x_{i_0} \geq c_{i_0 i_1} c_{i_1 i_2} \ldots c_{i_{r-1} i_0} x_{i_0} > x_{i_0},$$

a contradiction (if one assumes that $x_{i_0}, x_{i_1}, \ldots, x_{i_{r-1}} \in \mathbb{R}$). It follows that at least one of $x_{i_0}, x_{i_1}, \ldots, x_{i_{r-1}}$, say x_{i_s}, is equal to ϵ. But then we have that

$$\epsilon = x_{i_s} \geq c_{i_s i_{s+1}} x_{i_{s+1}} \Rightarrow x_{i_{s+1}} = \epsilon$$

and so on. We conclude that $x_{i_0} = x_{i_1} = \cdots = x_{i_{r-1}} = \epsilon$, and so we may remove these variables, reducing the size of the matrix C and eliminating the positive cycle. Repeating this process we either destroy the matrix C completely, or reach a matrix C' such that $\lambda(C') \leq 0$.

To see the complexity of the algorithm observe the following.

Step (**I**) defines I_{ij} for all $i, j \in N, i \neq j$. For every such I_{ij}, we consider each $k \in M$. Thus, this step takes $O(mn^2)$ time. Similarly, it takes $O(mn^2)$ time to calculate c_{ij} for all $i, j \in N, i \neq j$. Step (**III**) is trivial and takes $O(1)$ time.

There will be at most n repetitions of steps (**IV**) and (**V**). In each, we calculate $\lambda(C)$ ($O(n^3)$ time) and identify a positive cycle ($O(n^2)$ time). Together, these steps run in $O(n^4)$ time.

Finally, in step (**VI**) we calculate C^* in $O(n^3)$ time.

In total, the computational complexity is $O(mn^2) + O(n^4) = O(mn^2 + n^4)$ time.

□

We conclude with a non-trivial example to show that our method works. (Blank spaces denote ϵ).

Example 5.10. *Consider the system*

$$\begin{pmatrix} 1 & & & & \epsilon & 0 & & \\ & & 1 & & \epsilon & & 1 & \\ & 1 & & 2 & \epsilon & & & \\ 3 & & & & \epsilon & 4 & & \\ & 2 & & & \epsilon & & & 6 \end{pmatrix} x = \begin{pmatrix} 2 & & & & \epsilon & 0 & & \\ & & 1 & & \epsilon & & 0 & \\ & 0 & & 2 & \epsilon & & & \\ 5 & & & & \epsilon & 4 & & \\ & 2 & & & \epsilon & & & 4 \end{pmatrix} x.$$

We calculate

$$I_{61} = \{1, 4\},$$
$$I_{37} = \{2\},$$
$$I_{42} = \{3\},$$
$$I_{28} = \{5\}.$$

For all other $i \neq j$, $I_{ij} = \emptyset$. Note that $\bigcup_{i \neq j} I_{ij} = M$.

Now,

$$\begin{aligned} c_{61} &= \bigoplus_{k=1,4} \Big(\bigoplus (a_{k1}, b_{k1}) - a_{k6} \Big) \\ &= \bigoplus \Big(\bigoplus (1,2) - 0, \bigoplus (3,5) - 4 \Big) \\ &= \bigoplus (2-0, 5-4) \\ &= 2. \end{aligned}$$

$$c_{37} = \bigoplus (a_{27}, b_{27}) - a_{23} = 1 - 1 = 0.$$

$$c_{42} = \bigoplus (a_{32}, b_{32}) - a_{34} = 1 - 2 = -1.$$

$$c_{28} = \bigoplus (a_{58}, b_{58}) - a_{52} = 6 - 2 = 4.$$

For all other $i \neq j$, $c_{ij} = \epsilon$. *Also,* $(\forall i)\, c_{ii} = \epsilon$.

We have

$$C = \begin{pmatrix} \epsilon & & & & & & & \\ & \epsilon & & & & & & 4 \\ & & \epsilon & & & & 0 & \\ & -1 & & \epsilon & & & & \\ & & & & \epsilon & & & \\ 2 & & & & & \epsilon & & \\ & & & & & & \epsilon & \\ & & & & & & & \epsilon \end{pmatrix}.$$

We can check that $\lambda(C) < 0$ *and so we compute*

$$C^* = \begin{pmatrix} \epsilon & & & & & & & \\ & \epsilon & & & & & & 4 \\ & & \epsilon & & & & 0 & \\ & -1 & & \epsilon & & & & 3 \\ & & & & \epsilon & & & \\ 2 & & & & & \epsilon & & \\ & & & & & & \epsilon & \\ & & & & & & & \epsilon \end{pmatrix}.$$

Now let us choose some vector u*, say the zero vector, then*

$$\begin{pmatrix} \epsilon & & & & & & & \\ & \epsilon & & & & & & 4 \\ & & \epsilon & & & & 0 & \\ & -1 & & \epsilon & & & & 3 \\ & & & & \epsilon & & & \\ 2 & & & & & \epsilon & & \\ & & & & & & \epsilon & \\ & & & & & & & \epsilon \end{pmatrix} u = \begin{pmatrix} \epsilon \\ 4 \\ 0 \\ 3 \\ \epsilon \\ 2 \\ \epsilon \\ \epsilon \end{pmatrix}.$$

It is easy to check that this is a solution to our original system!

5.5 Summary

By making some natural, simplifying assumptions, we were able to assume without loss of generality the extra condition that in each row there is exactly one entry in A with the same value as the corresponding entry in B. This allowed us to use the Cancellation

Rule - reducing (5.1) to the sub-eigenvector problem, which is easily solved. We finished with a few examples.

6. Two-dimensional, two-sided homogeneous systems of equations and two-dimensional GEP - strongly polynomial solution methods

6.1 Introduction

In this chapter all matrices are finite, have exactly two columns and all vectors have exactly two components.

We approach this chapter in stages. First, we give a complete description of the solution space for two-dimensional two-sided systems $Ax = Bx$. This will help us to solve the 2×2 generalised eigenproblem $Ax = \lambda Bx$, which in turn is used to give the explicit set of solutions for the two-dimensional generalised eigenproblem. We extensively use the Cancellation Rule (Lemma 2.7) to obtain our results.

The contents of this chapter have been published in [17].

6.2 Problem formulation

The case when $x_1 = \varepsilon$ reduces the generalized eigenproblem

$$Ax = \lambda Bx \tag{6.1}$$

to the question of whether or not the second columns of A and B are proportional (and the coefficient of proportionality is then the unique generalized eigenvalue). Similarly for $x_2 = \varepsilon$, so we will restrict our attention to the task of finding finite x satisfying (6.1). By homogeneity of $V(A, B, \lambda)$ we can assume that $x_1 = 0$. We will therefore study the problem of finding $x_2 \in \mathbb{R}$ such that

$$a_{i1} \oplus a_{i2}x_2 = \lambda b_{i1} \oplus \lambda b_{i2}x_2; \quad i \in M. \tag{6.2}$$

Before we discuss the generalised eigenproblem, we will show in Subsection 6.3 how to find all solutions of two-sided systems

$$Ax = Bx, x \in \mathbb{R}^2 \tag{6.3}$$

for $A, B \in \mathbb{R}^{m \times 2}$ and then in Subsection 6.4 we show how to solve (6.2) for $m = 2$.

Section 6.5 can be seen as an independent generalisation of the results in [12], where the set of solutions to (6.2) in the 2×2 case is described, to the system (6.2) in the $m \times 2$ case.

6.3 Two-dimensional two-sided systems

In system (6.3) we can assume without loss of generality $x_1 = 0$:

$$a_{i1} \oplus a_{i2}x_2 = b_{i1} \oplus b_{i2}x_2; \quad i \in M. \tag{6.4}$$

Let us denote

$$V_i = \{x_2 \in \mathbb{R}; a_{i1} \oplus a_{i2}x_2 = b_{i1} \oplus b_{i2}x_2\}; \quad i \in M$$

and $V = \bigcap_{i \in M} V_i$.

We will explicitly describe each V_i. Let us apply the Cancellation Rule of Lemma 2.7 to (6.4). For every $i \in M$ there are either two, or one or no cancellations.

(i) If there is no cancellation then

$$a_{i1} = b_{i1} \text{ and } a_{i2} = b_{i2}$$

and so $V_i = \mathbb{R}$.

(ii) If there is exactly one cancellation then it can be assumed without loss of generality to take place on the left-hand side and we consider two cases.

Either $a_{i1} < b_{i1}$ and $a_{i2} = b_{i2}$ so that (6.4) reduces to

$$a_{i2}x_2 = b_{i1} \oplus a_{i2}x_2$$

yielding $V_i = [a_{i2}^{-1}b_{i1}, +\infty)$.

Or $a_{i1} = b_{i1}$ and $a_{i2} < b_{i2}$ so that (6.4) reduces to

$$a_{i1} = a_{i1} \oplus b_{i2}x_2$$

yielding $V_i = (-\infty, a_{i1}b_{i2}^{-1}]$.

(iii) If there are two cancellations and they take place on the same side then this side becomes ε yielding $V_i = \emptyset$. If the two cancellations take place on different sides then either $a_{i1} > b_{i1}$ and $a_{i2} < b_{i2}$ so that (6.4) reduces to

$$a_{i1} = b_{i2}x_2$$

yielding $V_i = \left\{a_{i1}b_{i2}^{-1}\right\}$ or, $a_{i1} < b_{i1}$ and $a_{i2} > b_{i2}$ yielding similarly $V_i = \left\{a_{i2}^{-1}b_{i1}\right\}$.

Since each V_i obtained above is a closed interval (including possibly a singleton or empty set) and can be found in a constant number of operations, the intersection $V = \bigcap_{i \in M} V_i$ is also a closed interval (including possibly a singleton or empty set) and can be found in $O(m)$ time.

We conclude:

Proposition 6.1. *The solution set to (6.3) is of the form*

$$\left\{\alpha\,(0, x_2)^T; \alpha \in \mathbb{R}, x_2 \in V\right\}$$

where V is a closed interval (including possibly a singleton or empty set), which can be found in $O(m)$ time as described above.

6.4 Generalized eigenproblem for 2×2 matrices

Our aim in this subsection is to describe the whole spectrum for the 2×2 generalized eigenproblem (6.1) which, without loss of generality, can be written

$$\begin{aligned} a_{11} \oplus a_{12}x_2 &= \lambda b_{11} \oplus \lambda b_{12}x_2 \\ a_{21} \oplus a_{22}x_2 &= \lambda b_{21} \oplus \lambda b_{22}x_2, \end{aligned} \tag{6.5}$$

where all a_{ij} and b_{ij} are real numbers.

This is a very special case, already solved in [12] but will be of key importance for solving the general two-dimensional case in the next subsection for which we use a different methodology from that in [12] for solving the general case in Subsection 6.5.

Recall (Proposition 2.6), $\Lambda(A,B) \subseteq [L,U]$. Both L and U can easily be found (in $O(mn)$ time). We will therefore assume that $L < U$ since otherwise we have $\Lambda(A,B) = \emptyset$ (if $L > U$) or $L = U$ is the unique candidate for a value in $\Lambda(A,B)$ and this can be verified easily for instance using the tools of Subsection 6.3. We will distinguish four cases and describe the spectrum in each of them. Recall

$$C = (c_{ij}) = \left(a_{ij}b_{ij}^{-1}\right).$$

The feasibility interval $[L,U]$ has exactly one of the forms below:

$$\begin{gathered}
[\max(c_{11},c_{21}), \min(c_{12},c_{22})], \\
[\max(c_{12},c_{22}), \min(c_{11},c_{21})], \\
[\max(c_{11},c_{22}), \min(c_{12},c_{21})], \\
[\max(c_{12},c_{21}), \min(c_{11},c_{22})].
\end{gathered}$$

Note that the first case can be equivalently described by inequalities $c_{11} < c_{12}, c_{21} < c_{22}$, similarly the other cases. The first two cases can be transformed to each other by swapping the variables x_1 and x_2. Similarly the last two cases. So we essentially have only two cases. In fact we will only deal with the third (and thus also with the fourth) case as the first (two) will be covered by the discussion in Subsection 6.5.

In what follows we denote

$$\gamma^2 = a_{12}a_{21}b_{11}^{-1}b_{22}^{-1}. \tag{6.6}$$

Proposition 6.2. *If* $c_{11} < c_{12}, c_{22} < c_{21}$ *and* $L < U$ *then* $\Lambda(A,B) = \{\widehat{\gamma}\}$, *where* $\widehat{\gamma}$ *is the* *unique projection of* γ *onto* $[L, U]$, *that is*

$$\widehat{\gamma} = \begin{cases} L & \text{if } \gamma \leq L, \\ \gamma & \text{if } \gamma \in (L, U), \\ U & \text{if } \gamma \geq U. \end{cases}$$

Proof. Note first that by the assumptions we have $L = \max(c_{11}, c_{22})$ and $U = \min(c_{12}, c_{21})$. Let us denote

$$S = (L, U) \cap \Lambda(A, B).$$

It is sufficient to prove the following statements:

(i) $S \neq \emptyset \Longrightarrow S = \{\gamma\}$,

(ii) $\gamma \in (L, U) \Longrightarrow \gamma \in S$,

(iii) $\gamma \in (L, U) \Longrightarrow L, U \notin \Lambda(A, B)$,

(iv) $\gamma \leq L \Longrightarrow \Lambda(A, B) = \{L\}$ and

(v) $\gamma \geq U \Longrightarrow \Lambda(A, B) = \{U\}$.

In order to prove (i) suppose $\lambda \in S$. Hence we have

$$c_{11}, c_{22} < \lambda < c_{12}, c_{21}$$

and thus (using $c_{ij} = a_{ij} b_{ij}^{-1}$)

$$\begin{aligned} a_{11} &< \lambda b_{11}, \quad a_{12} > \lambda b_{12}, \\ a_{22} &< \lambda b_{22}, \quad a_{21} > \lambda b_{21}. \end{aligned}$$

System (6.5) reduces by the Cancellation Rule (Lemma 2.7) in this case to

$$\left.\begin{array}{ccc} & a_{12}x_2 = \lambda b_{11} & \\ a_{21} & = & \lambda b_{22}x_2. \end{array}\right\}$$

So $x_2 = \lambda b_{11} a_{12}^{-1}$ and $x_2 = \lambda^{-1} a_{21} b_{22}^{-1}$, from which $\lambda = \gamma$ follows.

(ii) Suppose $\gamma \in (L, U)$ and put $\lambda = \gamma$. By taking $x_2 = \lambda b_{11} a_{12}^{-1} = \lambda^{-1} a_{21} b_{22}^{-1}$ we see that $\lambda \in \Lambda(A, B)$.

(iii) Suppose that $\gamma \in (L, U)$ and $\lambda = L \in \Lambda(A, B)$. If $c_{11} < c_{22}$ then

$$c_{11} < c_{22} = \lambda < c_{12}, c_{21}$$

and thus

$$a_{11} < \lambda b_{11}, \quad a_{12} > \lambda b_{12},$$
$$a_{22} = \lambda b_{22}, \quad a_{21} > \lambda b_{21}.$$

By cancellations and substituting λb_{22} for a_{22} system (6.5) reduces to

$$\left.\begin{array}{rl} a_{12} x_2 = \lambda b_{11} & \\ a_{21} \oplus \lambda b_{22} x_2 = & \lambda b_{22} x_2. \end{array}\right\}$$

So $x_2 = \lambda b_{11} a_{12}^{-1}$ and $x_2 \geq \lambda^{-1} a_{21} b_{22}^{-1}$, from which $\lambda^2 \geq \gamma^2$, a contradiction.

A contradiction is obtained in a similar way when $c_{11} > c_{22}$ or $c_{11} = c_{22}$.

The case of $\lambda = U \in \Lambda(A, B)$ is dealt with in a similar way.

(iv) Suppose $\gamma \leq L$. Due to (i) it is sufficient to prove that $L \in \Lambda(A, B)$ and $U \notin \Lambda(A, B)$. Let $\lambda = L$. It is easily verified that x_2 is a solution to (6.5) where

$$x_2 = \begin{cases} \lambda b_{11} a_{12}^{-1}, & \text{if } c_{11} < c_{22}; \\ \lambda^{-1} a_{21} b_{22}^{-1}, & \text{if } c_{11} > c_{22} \text{ and} \\ \text{any value in } \left[a_{11} a_{22}^{-1}, a_{11} a_{12}^{-1}\right] & \text{if } c_{11} = c_{22}. \end{cases}$$

Let $\lambda = U$ and suppose $c_{12} < c_{21}$. Then $\lambda > \gamma$ and

$$c_{11}, c_{22} < \lambda = c_{12} < c_{21}$$

and thus

$$\begin{aligned} a_{11} &< \lambda b_{11}, \quad a_{12} = \lambda b_{12}, \\ a_{22} &< \lambda b_{22}, \quad a_{21} > \lambda b_{21}. \end{aligned}$$

By cancellations and substituting λb_{12} for a_{12} system (6.5) reduces to

$$\left.\begin{array}{rcl} a_{12}x_2 & = & \lambda b_{11} \oplus a_{12}x_2 \\ a_{21} & = & \lambda b_{22}x_2. \end{array}\right\}$$

So $x_2 \geq \lambda a_{12}^{-1} b_{11}$ and $x_2 = \lambda^{-1} b_{22}^{-1} a_{12}$, from which $\lambda^2 \leq \gamma^2$, a contradiction.

A contradiction can similarlly be obtained when $c_{12} > c_{21}$ or $c_{12} = c_{21}$.

(v) The proof of this part is similar to that of (iv) and is omitted here. □

6.5 Generalized eigenproblem: the two-dimensional case

As before, due to the finiteness of x and homogeneity we assume that $x_1 = 0$ and we therefore study system (6.2).

We will distinguish three cases.

Case 1: If $c_{i1} = c_{i2}$ for some $i \in M$ then by Proposition 2.6 this value is the unique candidate for the generalized eigenvalue. Using the method of Subsection 6.3 it can be readily checked whether this is indeed the case.

Case 2: If $c_{i_1 1} < c_{i_1 2}$ and $c_{i_2 1} > c_{i_2 2}$ for some $i_1, i_2 \in M$ then the 2×2 system

$$\begin{pmatrix} a_{i_1 1} & a_{i_1 2} \\ a_{i_2 1} & a_{i_2 2} \end{pmatrix} x = \lambda \begin{pmatrix} b_{i_1 1} & b_{i_1 2} \\ b_{i_2 1} & b_{i_2 2} \end{pmatrix} x$$

has a unique eigenvalue by Proposition 6.2 and is therefore a unique candidate for an eigenvalue of the whole system. This can easily be checked by the method of Subsection 6.3.

Case 3: If $c_{i1} < c_{i2}$ for all $i \in M$ (the case when $c_{i1} > c_{i2}$ for all $i \in M$ can be discussed similarly) then for any $i \in M$ the feasibility interval for the i^{th} equation alone is $[c_{i1}, c_{i2}]$. Suppose $\lambda \in \Lambda(A, B) \cap (c_{i1}, c_{i2})$. Then $a_{i1} < \lambda b_{i1}$ and $a_{i2} > \lambda b_{i2}$ and the equation

$$a_{i1} \oplus a_{i2} x_2 = \lambda b_{i1} \oplus \lambda b_{i2} x_2 \tag{6.7}$$

reduces using cancellations to

$$a_{i2} x_2 = \lambda b_{i1}.$$

Hence

$$x_2 = \lambda b_{i1} a_{i2}^{-1} \tag{6.8}$$

and thus the dependence of x_2 on λ over (c_{i1}, c_{i2}) is expressed by a linear function (with slope 1). This concludes the case when λ is strictly between c_{i1} and c_{i2}. To finish Case 3 suppose now that $\lambda = c_{i1} \in \Lambda(A, B)$. Then $a_{i1} = \lambda b_{i1}$ and $a_{i2} > \lambda b_{i2}$ and equation (6.7) reduces using cancellations to

$$a_{i1} \oplus a_{i2} x_2 = a_{i1}.$$

Hence $x_2 \leq a_{i1} a_{i2}^{-1}$. Similarly if $\lambda = c_{i2} \in \Lambda(A, B)$ then $x_2 \geq b_{i1} b_{i2}^{-1}$. Note that

$$\lim_{\lambda \to c_{i1}} \lambda b_{i1} a_{i2}^{-1} = a_{i1} a_{i2}^{-1} \quad \text{and} \quad \lim_{\lambda \to c_{i2}} \lambda b_{i1} a_{i2}^{-1} = b_{i1} b_{i2}^{-1}$$

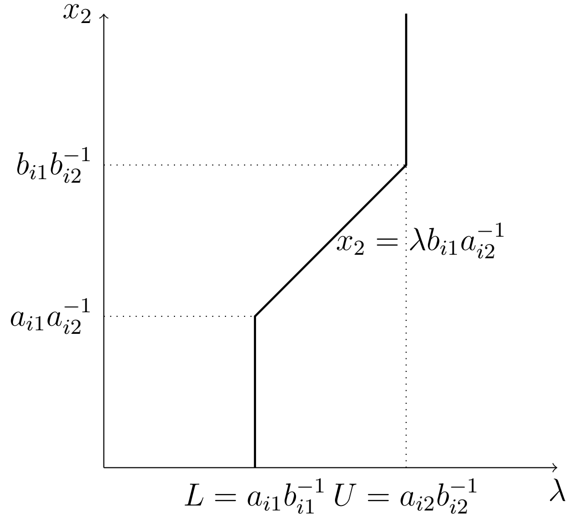

Figure 6.1: dependence of x_2 on λ over $[c_{i1}, c_{i2}]$

and so the graph of dependence of x_2 on λ over $[c_{i1}, c_{i2}]$ is a continuous, piece-wise linear map, see Figure 6.1. This result is consistent with the fact that $a_{i1}a_{i2}^{-1} < b_{i1}b_{i2}^{-1}$, since this is equivalent to $c_{i1} < c_{i2}$.

Finally we note that for the whole system we have

$$(L, U) = \bigcap_{i \in M} (c_{i1}, c_{i2}) .$$

Thus if $\lambda \in \Lambda(A, B) \cap (L, U)$ then $\lambda \in \Lambda(A, B) \cap (c_{i1}, c_{i2})$ for every $i \in M$ and so x_2 is the *common* value of all $\lambda b_{i1}a_{i2}^{-1}, i \in M$ (6.8). This implies $(L, U) \subseteq \Lambda(A, B)$. We have proved:

Proposition 6.3. *If $A, B \in \mathbb{R}^{m \times 2}$ and $c_{i1} < c_{i2}$ for every $i \in M$ then a generalized eigenvalue in (L, U) exists if and only if all values in (L, U) are generalized eigenvalues. This is equivalent to the requirement that all values $b_{i1}a_{i2}^{-1}$ for $i \in M$ coincide.*

If the condition in Proposition 6.3 is satisfied then by continuity also $L, U \in \Lambda(A, B)$ and in this case $\Lambda(A, B) = [L, U]$. If not then L, U have to be examined separately for being generalized eigenvalues. Figures 6.2-6.6 indicate that all possibilities may occur

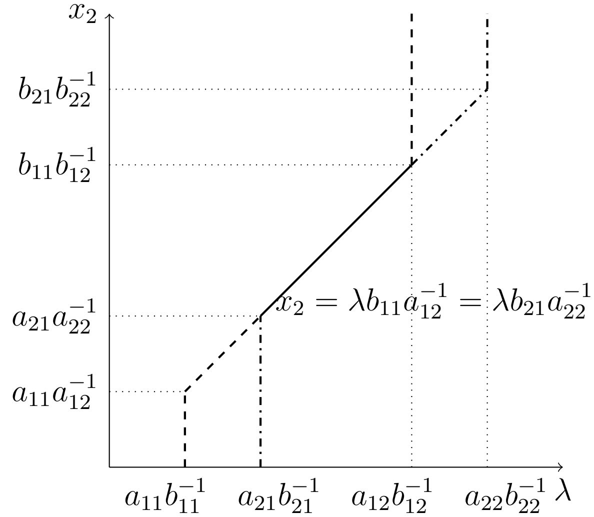

Figure 6.2: A continuum of solutions: $\lambda \in [L, U] = \left[a_{21}b_{21}^{-1}, a_{12}b_{12}^{-1}\right], x_2 = \lambda b_{11}a_{12}^{-1} = \lambda b_{21}a_{22}^{-1}$

(both of L, U, exactly one, or none of them in $\Lambda(A, B)$).

Summarizing all cases we have that if $A, B \in \mathbb{R}^{m\times 2}$ then $\Lambda(A, B)$ can be found in $O(m)$ time and has one of the following forms (the illustrating figures are drawn for $m = 2$):

- $[L, U]$, see Figure 6.2,
- $\{L, U\}$, see Figure 6.3,
- $\{\lambda\}$, where $\lambda \in [L, U]$, see Case 2 and Figures 6.4 and 6.5,
- $\emptyset$, see Figure 6.6 and Case 1.

In all cases the eigenspace associated with a fixed generalized eigenvalue is described in Proposition 6.1.

6.6 Summary

We saw that the Cancellation Rule can be a powerful tool. Using it carefully allowed us to give explicit solutions for two-dimensional two-sided systems and the two-dimensional

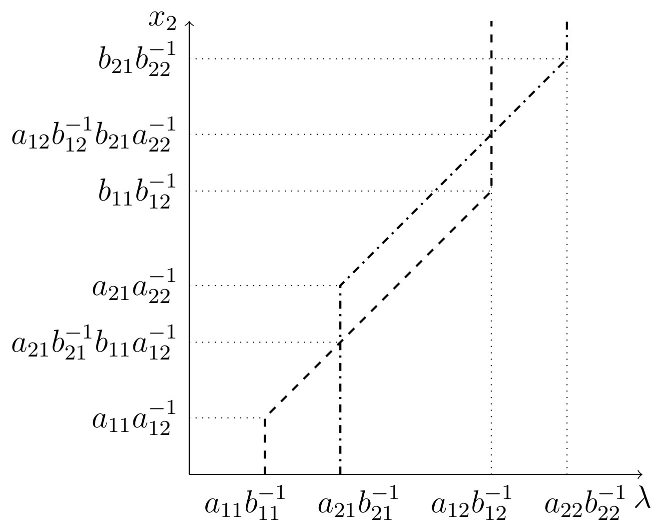

Figure 6.3: Two solutions: $(\lambda, x_2) = \left(L = a_{21}b_{21}^{-1}, a_{21}b_{21}^{-1}b_{11}a_{12}^{-1}\right)$ and $(\lambda, x_2) = (U = a_{12}b_{12}^{-1}, a_{12}b_{12}^{-1}b_{21}a_{22}^{-1})$

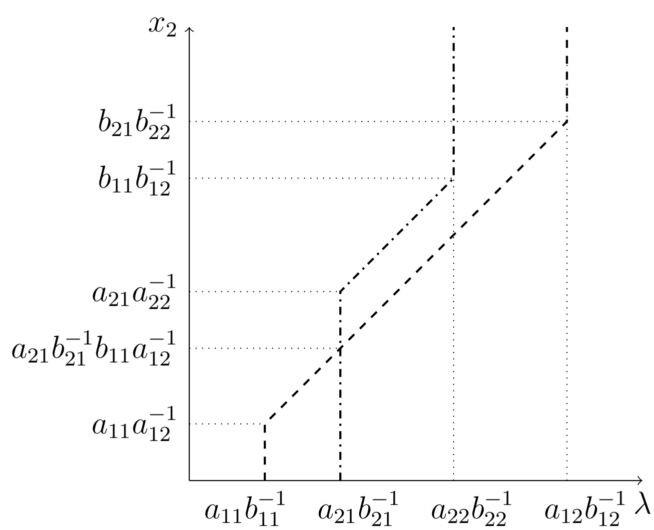

Figure 6.4: One solution: $(\lambda, x_2) = (L = a_{21}b_{21}^{-1}, a_{21}b_{21}^{-1}b_{11}a_{12}^{-1})$

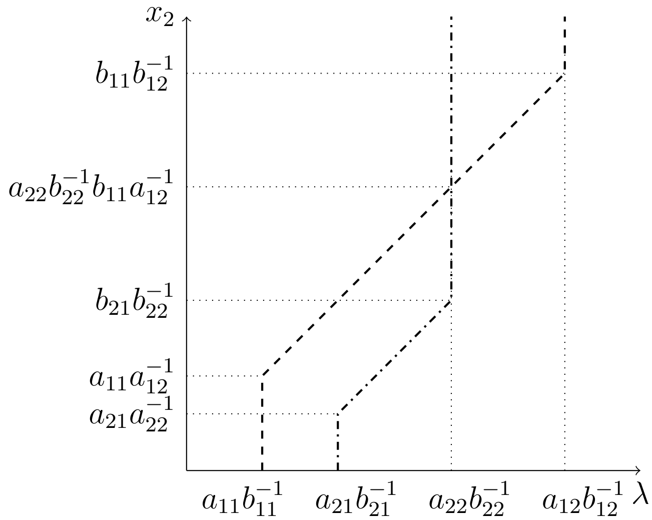

Figure 6.5: One solution: $(\lambda, x_2) = (U = a_{22}b_{22}^{-1}, a_{22}b_{22}^{-1}b_{11}a_{12}^{-1})$

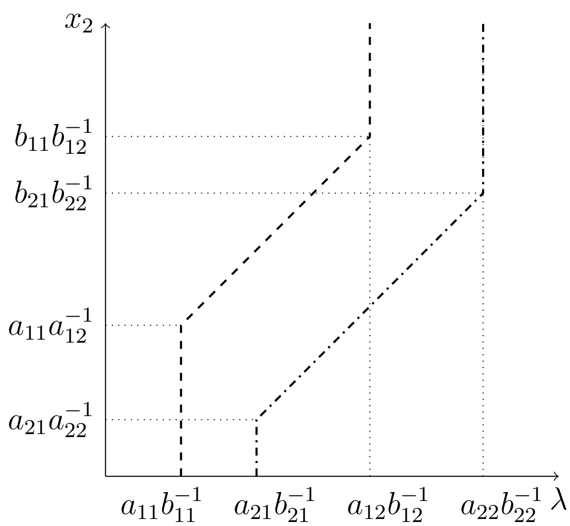

Figure 6.6: No solutions

generalised eigenproblem.

7. Two-sided systems of equations with separated variables - strongly polynomial solution method if B has exactly two columns

7.1 Introduction

In this chapter, the matrices A and B are finite and the matrix B (say) has exactly two columns.

We show that such two-sided systems with separated variables $Ax = By$ can be solved in strongly polynomial time, in that we can find a finite solution. We do this by considering this system as a sequence of one-sided parametrised systems, of the form studied in chapter 3. It is possible, by considering different solution types, to fully describe the solution set for each type.

7.2 Problem formulation

We are concerned here with a special case of the two-sided systems of max-linear equations with separated variables; namely

$$Ax = By, \tag{7.1}$$

where $A \in \mathbb{R}^{m\times n}, B \in \mathbb{R}^{m\times 2}, x \in \overline{\mathbb{R}}^{n}, y \in \overline{\mathbb{R}}^{2}$. Our aim is to find a non-trivial solution if it exists and identify the case when it does not. (We will see that we can in fact look for a finite solution). Throughout this work we assume (x, y) is a non-trivial solution, unless otherwise stated. Note that the algorithm in this chapter can be adapted to find all solutions in polynomial time.

We note that due to the finite nature of A and B, for any solution we have $y = \epsilon \Leftrightarrow x = \epsilon$, therefore we have a non-trivial solution if and only if we have a solution for which y is non-trivial. Throughout this work we will therefore assume that $y \neq \epsilon$.

Note that if $y_1 = \epsilon$, say, then we can set $y_2 = 0$ and then (7.1) is simply equivalent to a one-sided system of the form $Ax = b$ for some vector b; such systems are easily solved (see [25]). Similarly if $y_2 = \epsilon$. That is, we can always start by checking solvability of $Ax = B_1 y_1$ and $Ax = B_2 y_2$, where B_j denotes column j of the matrix B for $j = 1, 2$. For the rest of this chapter, we assume that these two one-sided systems have been checked and any solutions that have been found have been added to the solution set. As such, we assume from now on that y is finite and so, without loss of generality, $y_1 = 0$. It follows now that we may also assume without loss of generality that x is finite.

A preliminary observation simplifies our work. Firstly, we may assume, without loss of generality, that $(\forall i)\, b_{i1} = 0$ (by scaling rows of the matrices A and B appropriately). We will also assume without loss of generality, by rearranging rows, that the sequence $b_{12}, b_{22}, \dots, b_{m2}$ is non-increasing. It is then possible for us to conveniently partition the set M according to the values of b_{i2}.

Rigorously, we have

$$b_{12} = \cdots = b_{k_1 2} > b_{k_1+1,2} = \cdots = b_{k_2,2} > \cdots > b_{k_{p-1}+1,2} = \cdots = b_{k_p,2}. \qquad (7.2)$$

We denote

$$K_1 = \{1, \ldots, k_1\}, K_2 = \{k_1 + 1, \ldots, k_2\}, \ldots, K_p = \{k_{p-1} + 1, \ldots, k_p\}. \tag{7.3}$$

So if we let $1 \leq t \leq p$, then $(\forall i \in K_t)\, b_{i2} = b_{k_t 2}$. We then have that the sequence $b_{k_1 2}, b_{k_2 2}, \ldots, b_{k_p 2}$ is strictly decreasing, equivalently, the sequence $b_{k_1 2}^{-1}, b_{k_2 2}^{-1}, \ldots, b_{k_p 2}^{-1}$ is strictly increasing. Let us also define

$$(\forall t)\, \gamma_t := b_{k_1 2} b_{k_t 2}^{-1}, \tag{7.4}$$

so $\gamma_1 = 0$. Note also that (γ) is a strictly increasing sequence.

An example illustrates the partitioning process.

Example 7.1.

$$\begin{pmatrix} -1 & 0 & 1 \\ -1 & -2 & 0 \\ 2 & 0 & 0 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & -1 \\ 0 & -1 \\ 0 & -3 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}.$$

We set $K_1 = \{1\}, K_2 = \{2, 3\}, K_3 = \{4\}$.

In Example 7.1 above we therefore have $\gamma_1 = 1 \otimes 1^{-1} = 0, \gamma_2 = 1 \otimes (-1)^{-1} = 2, \gamma_3 = 1 \otimes (-3)^{-1} = 4$. Note also in Example 7.1 that for $y_2 \in (-\infty, -1]$, our system is equivalent to

$$\begin{pmatrix} -1 & 0 & 1 \\ -1 & -2 & 0 \\ 2 & 0 & 0 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix},$$

a one-sided system which is easily solvable (see chapter 2).

Similarly, if $y_2 \in [-1, 1]$, then our system is equivalent to

$$\begin{pmatrix} -1 & 0 & 1 \\ -1 & -2 & 0 \\ 2 & 0 & 0 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \alpha \\ 0 \\ 0 \\ 0 \end{pmatrix},$$

where $\alpha \in [0, 2]$. This is a one sided system where b is now a parametrised vector. We have seen in chapter 3 that such systems, whilst more involved, are also easily solvable.

Again, if $y_2 \in [1, 3]$, then our system is equivalent to

$$\begin{pmatrix} -1 & 0 & 1 \\ -1 & -2 & 0 \\ 2 & 0 & 0 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \alpha \\ \alpha - 2 \\ \alpha - 2 \\ 0 \end{pmatrix},$$

where $\alpha \in [2, 4]$.

Finally, if $y_2 \in [3, \infty]$, then our system is equivalent to

$$\begin{pmatrix} -1 & 0 & 1 \\ -1 & -2 & 0 \\ 2 & 0 & 0 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \alpha \\ \alpha - 2 \\ \alpha - 2 \\ \alpha - 4 \end{pmatrix},$$

where $\alpha \in [4, \infty)$.

There is a pattern emerging and we summarise it in the following Lemma. The reader should refer to the definitions and results of chapter 3 and also to (7.2), (7.3) and (7.4) for definitions of terms used in the following Lemma.

Lemma 7.2. *Consider the system $Ax = By$ for finite matrices A and B, where B has exactly two columns.*

1. *For $y_2 \in \left(-\infty, b_{k_1 2}^{-1}\right]$, (7.1) is equivalent to $Ax = 0$, a one-sided max-linear system of equations.*

2. *For $y_2 \in \left[b_{k_p 2}^{-1}, \infty\right)$, (7.1) is equivalent to $Ax = b(\alpha)$, where* $(\forall 1 \leq t \leq p)\,(\forall i \in K_t)\, b_i = \alpha\gamma_t^{-1}, \alpha \in [\gamma_p, \infty)$.

3. *For all $1 \leq r \leq p-1$, for $y_2 \in \left[b_{k_r 2}^{-1}, b_{k_{r+1} 2}^{-1}\right]$, (7.1) is equivalent to $Ax = b(\alpha)$, where* $(\forall 1 \leq t \leq r)\,(\forall i \in K_t)\, b_i = \alpha\gamma_t^{-1}, (\forall r+1 \leq t \leq p)\,(\forall i \in K_t)\, b_i = 0, \alpha \in [\gamma_r, \gamma_{r+1}]$.

Proof. The proof is in three parts.

1. First assume $y_2 \in \left(\infty, b_{k_1 2}^{-1}\right]$. It follows $(\forall 1 \leq t \leq p)\,(\forall i \in K_t)$ that $-\infty \leq b_{i2}y_2 \leq b_{i2}b_{k_1 2}^{-1} \leq b_{k_1 2}b_{k_1 2}^{-1} = 0 = b_{i1}y_1$, ($b_{i2} \leq b_{k_1 2}$ due to the monotonicity of the column vector B_2). It follows that $(\forall i \in M)\, b_{i1}y_1 \oplus b_{i2}y_2 = b_{i1}y_1 = 0$, the result follows.

2. Next, assume that $y_2 \in \left[b_{k_p 2}^{-1}, \infty\right)$. Let $1 \leq t \leq p$ and $i \in K_t$. When $y_2 = b_{k_p 2}^{-1}$ we have that $b_{i2}y_2 = b_{k_t 2}b_{k_p 2}^{-1} = \left(b_{k_1 2}b_{k_p 2}^{-1}\right)\left(b_{k_1 2}b_{k_t 2}^{-1}\right)^{-1} = \gamma_p\gamma_t^{-1} \geq 0$ (due to the sequence (γ) being strictly increasing).
   It follows that $b_{i1}y_1 \oplus b_{i2}y_2 \geq \gamma_p\gamma_t^{-1}$ and equality holds if and only if $y_2 = b_{k_p 2}^{-1}$. The result follows.

3. Now let $1 \le r \le p-1$ and let $y_2 \in \left[b_{k_r 2}^{-1}, b_{k_{r+1} 2}^{-1}\right]$. Let $1 \le t \le r$ and $i \in K_t$. Then firstly, $b_{i2} y_2 = b_{k_t 2} y_2 \ge b_{k_t 2} b_{k_r 2}^{-1} = \gamma_r \gamma_t^{-1}$ and equality holds if and only if $y_2 = b_{k_r 2}^{-1}$. (Note also the inequality $\gamma_r \gamma_t^{-1} \ge 0$).

   Secondly, $b_{i2} y_2 = b_{k_t 2} y_2 \le b_{k_t 2} b_{k_{r+1} 2}^{-1} = \gamma_{r+1} \gamma_t^{-1}$ and equality holds if and only if $y_2 = b_{k_{r+1} 2}^{-1}$.

   It follows that $b_i = \alpha \gamma_t^{-1}$, where $\alpha \in [\gamma_r, \gamma_{r+1}]$.

   Now let $r+1 \le t \le p$ and $i \in K_t$. Then $b_{i2} y_2 = b_{k_t 2} y_2 \le b_{k_t 2} b_{k_{r+1} 2}^{-1} \le 0$ (due to monotonicity of column vector B_2 and since $t \ge r+1$). It follows that

   $b_{i1} y_1 \oplus b_{i2} y_2 = 0 \oplus b_{i2} y_2 = 0$. The result follows.

□

7.3 Outline of work

The following is a brief outline of how the work in this chapter will continue. In the above we have shown that the two-sided system (7.1) with separated variables can be viewed as a sequence of one-sided systems of the form $Ax = b(\alpha)$, where $\alpha \in [\underline{\alpha}, \overline{\alpha}]$ for some $\underline{\alpha}, \overline{\alpha} \in \overline{\mathbb{R}}$. In fact, the first system, i.e. for $y_2 \in \left(-\infty, b_{k_1 2}^{-1}\right]$, is a one-sided system without a parameter, which we will call S_0. It is easy to find all solutions (if any) to S_0 since it is essentially a one-sided system. We are now left with $p \le m$ one-sided parametrised systems. The system corresponding to the case $y_2 \in \left[b_{k_p 2}^{-1}, \infty\right)$ will be called S_p and $(\forall r)\,(1 \le r \le p-1)$ the system corresponding to the case $y_2 \in \left[b_{k_r 2}^{-1}, b_{k_{r+1} 2}\right]$ will be called S_r.

7.4 Theory

Consider the one-sided parametrised system S_r for some $r \ge 1$. By scaling the equations appropriately we may assume without loss of generality that S_r is of the form discussed in chapter 3. As a consequence, the results there are available to us here. The system S_r has a very distinct form. More precisely, it can be split into two parts: the top half (where the parameter α appears in the vector b); and the bottom half (where the parameter does not appear in b). In the language of chapter 3, we call the top half K and the bottom half $M \backslash K$. Define for all r, $K'_r := \bigcup_{1 \le t \le r} K_t$. An important result for cutting down the complexity time for a solution is the following.

Lemma 7.3. *Consider the system S_r for some $1 \le r \le p-1$. If there does not exist $\alpha \in [\gamma_r, \gamma_{r+1}]$ such that $\overline{x}^{(r)}(\alpha)$ solves S_r restricted to K'_r, namely the system $A[K'_r]\, x = b[K'_r]$, then there is no non-trivial solution to $S_{r'}$ for any $r' \ge r$.*

Proof. This possibly surprising result is essentially a consequence of the homogeneity of our system.

Let $1 \le r \le p-1$ and consider the system S_r. Let $1 \le t \le r$ and $i \in K_t$. Before we continue, note that the vector $\overline{x}$ depends not only on α but also the system we are considering. So in system S_r we denote $\overline{x}(\alpha)$ by $\overline{x}^{(r)}(\alpha)$. We aim to show that if there does not exist $\alpha \in [\gamma_r, \gamma_{r+1}]$ such that $\overline{x}^{(r)}(\alpha)$ solves the ith equation in S_r, then for all $r' > r$, there does not exist $\alpha \in [\gamma_{r'}, \gamma_{r'+1}]$ such that $\overline{x}^{(r')}(\alpha)$ solves the ith equation in $S_{r'}$.

So, assume that for all $\alpha \in [\gamma_r, \gamma_{r+1}]$, we have

$$\left(A\overline{x}^{(r)}(\alpha)\right)_i < \alpha\gamma_t^{-1}.$$

Here we have used the fact that $(\forall i)\,(A\overline{x}(\alpha))_i \le (By)_i$ and we are assuming in this case that equality does not hold. Also, we saw in Lemma 7.2 the form of the right hand side.

Now let $r' > r$. Let $\alpha_r \in [\gamma_r, \gamma_{r+1}]$ and $\alpha_{r'} \in [\gamma_{r'}, \gamma_{r'+1}]$ be fixed. We have that

$$\left(A\overline{x}^{(r)}(\alpha_r)\right)_i = \bigoplus_{j \in N} \left(a_{ij}\overline{x}_j^{(r)}(\alpha_r)\right)$$

and

$$\left(A\overline{x}^{(r')}\left(\alpha_{r'}\right)\right)_i = \bigoplus_{j\in N}\left(a_{ij}\overline{x}_j^{(r')}\left(\alpha_{r'}\right)\right).$$

We wish to examine the change in these two quantities, in order to do that, we must examine the change in the quantities $\overline{x}^{(r)}\left(\alpha_r\right)$ and $\overline{x}^{(r')}\left(\alpha_{r'}\right)$. In particular, we wish to find the maximum possible size of the increase from the quantity $\overline{x}^{(r)}\left(\alpha_r\right)$ to the quantity $\overline{x}^{(r')}\left(\alpha_{r'}\right)$.

Let $j \in N$. Then (by (3.2))

$$\begin{aligned}\overline{x}_j^{(r)}\left(\alpha_r\right) &:= \sideset{}{'}\bigoplus_{i\in M}\left(b_i a_{ij}^{-1}\right)\\ &= \sideset{}{'}\bigoplus_{i\in\cup_{1\leq t\leq r}K_t}\left(\alpha_r\gamma_t^{-1}a_{ij}^{-1}\right)\oplus'\sideset{}{'}\bigoplus_{i\in\cup_{r+1\leq t\leq r'}K_t}\left(a_{ij}^{-1}\right)\oplus'\sideset{}{'}\bigoplus_{i\in\cup_{r'+1\leq t\leq p}K_t}\left(a_{ij}^{-1}\right)\\ &= \alpha_r\sideset{}{'}\bigoplus_{i\in\cup_{1\leq t\leq r}K_t}\left(\gamma_t^{-1}a_{ij}^{-1}\right)\oplus'\sideset{}{'}\bigoplus_{i\in\cup_{r+1\leq t\leq r'}K_t}\left(a_{ij}^{-1}\right)\oplus'\sideset{}{'}\bigoplus_{i\in\cup_{r'+1\leq t\leq p}K_t}\left(a_{ij}^{-1}\right).\end{aligned}$$

Define

$$\begin{aligned}a &:= \alpha_r\sideset{}{'}\bigoplus_{i\in\cup_{1\leq t\leq r}K_t}\left(\gamma_t^{-1}a_{ij}^{-1}\right)\\ b &:= \sideset{}{'}\bigoplus_{i\in\cup_{r+1\leq t\leq r'}K_t}\left(a_{ij}^{-1}\right)\\ c &:= \sideset{}{'}\bigoplus_{i\in\cup_{r'+1\leq t\leq p}K_t}\left(a_{ij}^{-1}\right),\end{aligned}$$

so that $\overline{x}_j^{(r)}\left(\alpha_r\right) = a\oplus' b\oplus' c$.

Similarly,

$$\begin{aligned}\overline{x}_j^{(r')}\left(\alpha_{r'}\right) =&\alpha_{r'} \bigoplus_{i\in\cup_{1\leq t\leq r}K_t}^{\prime} \left(\gamma_t^{-1}a_{ij}^{-1}\right)\\ &\oplus' \alpha_{r'} \bigoplus_{i\in\cup_{r+1\leq t\leq r'}K_t}^{\prime} \left(\gamma_t^{-1}a_{ij}^{-1}\right)\\ &\oplus' \bigoplus_{i\in\cup_{r'+1\leq t\leq p}K_t}^{\prime} \left(a_{ij}^{-1}\right).\end{aligned}$$

Define

$$\begin{aligned}a' &:= \alpha_{r'} \bigoplus_{i\in\cup_{1\leq t\leq r}K_t}^{\prime} \left(\gamma_t^{-1}a_{ij}^{-1}\right)\\ b' &:= \bigoplus_{i\in\cup_{r+1\leq t\leq r'}K_t}^{\prime} \left(\gamma_t^{-1}a_{ij}^{-1}\right)\\ c' &:= \bigoplus_{i\in\cup_{r'+1\leq t\leq p}K_t}^{\prime} \left(a_{ij}^{-1}\right),\end{aligned}$$

so that $\overline{x}_j^{(r')}\left(\alpha_{r'}\right) = a' \oplus' \alpha_{r'}b' \oplus' c'$.

Notice that $a' = \left(\alpha_{r'}\alpha_r^{-1}\right) a$ and so the size of the increase from a to a' is $\alpha_{r'}\alpha_r^{-1}$. Also, $c' = c$ and so there is no increase here. Examining the increase from b to b' is where the difficulty lies, this is due to the γ_t^{-1} term appearing in b'. Note that here $r+1 \leq t \leq r'$ and that $0 > \gamma_{r+1}^{-1} > \gamma_{r+2}^{-1} > \cdots > \gamma_{r'}^{-1}$ due to the monotonicity of γ. As a result, there is actually a decrease from the quantity b to b'. We are trying to maximise increase and so equivalently, we are trying to minimise decrease. This happens when the minimum in b is attained for $t = r+1$ and so the smallest decrease possible is γ_{r+1}^{-1}. It follows that the maximum possible increase from b to $\alpha_{r'}b'$ is $\alpha_{r'}\gamma_{r+1}^{-1}$. But

$$\alpha_r \leq \gamma_{r+1} \Rightarrow \alpha_r^{-1} \geq \gamma_{r+1}^{-1} \Rightarrow \alpha_{r'}\gamma_{r+1}^{-1} \leq \alpha_{r'}\alpha_r^{-1}.$$

We conclude that the maximum possible increase from $\overline{x}^{(r)}(\alpha_r)$ to $\overline{x}^{(r')}(\alpha_{r'})$, given by $\overline{x}^{(r')}(\alpha_{r'})\left(\overline{x}^{(r)}(\alpha_r)\right)^{-1}$ is

$$\begin{aligned}\overline{x}^{(r')}(\alpha_{r'})\left(\overline{x}^{(r)}(\alpha_r)\right)^{-1} &= \left(a' \oplus' (\alpha_{r'}b') \oplus' c'\right)(a \oplus' b \oplus' c)^{-1} \\ &\leq \left(\left(a\left(\alpha_{r'}\alpha_r^{-1}\right)\right) \oplus' \left(b\left(\alpha_{r'}\alpha_r^{-1}\right)\right) \oplus' c\right)(a \oplus' b \oplus' c)^{-1} \\ &\leq \alpha_{r'}\alpha_r^{-1}.\end{aligned}$$

It follows, finally, that since the maximum possible increase from $\overline{x}^{(r)}(\alpha_r)$ to $\overline{x}^{(r')}(\alpha_{r'})$ is $\alpha_{r'}\alpha_r^{-1}$, that the maximum possible increase from

$$\left(A\overline{x}^{(r)}(\alpha_r)\right)_i \text{ to } \left(A\overline{x}^{(r')}(\alpha_{r'})\right)_i \text{ is } \alpha_{r'}\alpha_r^{-1}.$$

We have examined the left hand side of the ith equation in systems S_r and S_{r+1} and seen that the increase is no more than $\alpha_{r'}\alpha_r^{-1}$. We now examine the corresponding right hand sides.

Recall $i \in K_t$, where $1 \leq t \leq r$. So

$$\left(A\overline{x}^{(r)}(\alpha_r)\right)_i = \alpha_r\gamma_t^{-1}$$

and

$$\left(A\overline{x}^{(r')}(\alpha_{r'})\right)_i = \alpha_{r'}\gamma_t^{-1}.$$

It is clear that the increase from $\alpha_r\gamma_t^{-1}$ to $\alpha_{r'}\gamma_t^{-1}$ is exactly $\alpha_{r'}\alpha_r^{-1}$.

We now have everything we need.

We start with the strict inequality in system S_r:

$$\left(A\overline{x}^{(r)}\left(\alpha_r\right)\right)_i < \alpha_r \gamma_t^{-1}.$$

From S_r to $S_{r'}$, the left hand side increases by at most $\alpha_{r'}\alpha_r^{-1}$ and the right hand side increases by exactly $\alpha_{r'}\alpha_r^{-1}$. It follows that the inequality still holds strictly! Since this argument is true for any $\alpha_r \in [\gamma_r, \gamma_{r+1}]$ and any $\alpha_{r'} \in [\gamma_{r'}, \gamma_{r'+1}]$, it follows that for all $\alpha \in [\gamma_{r'}, \gamma_{r'+1}]$, the vector $\overline{x}^{(r)}(\alpha)$ does not solve the ith equation. □

We now have everything we need to create an algorithm for finding a solution (x, y) for the system $Ax = By$; we proceed as follows. Refer to (7.2), (7.3) and (7.4) for definitions of quantities used.

7.5 Algorithm A2

Algorithm A2:

Input: Finite matrices A and B where B has exactly two columns.

Output: Answer to the question of solvability of the system $Ax = By$. A non-trivial solution if one exists.

(**I**) Identify sets $K_1, K_2, \ldots, K_p$ and quantities $\gamma_1, \gamma_2, \ldots, \gamma_p$.

(**II**) Identify one-sided parametrised systems $S_0, S_1, \ldots, S_p$ and scale each so of the form in previous Section.

(**III**) Solve the system S_0. If there is a solution, then return it and terminate algorithm. If not, go to (**IV**).

(**IV**)$r := 1$.

(**V**) Consider the system S_r. Find the interval from Theorem 3.14. If the interval is non-empty, then we have found a solution (there is a one-to-one correspondence between α and y_2). Return the solution and terminate the algorithm. If the interval is empty and

$r \leq p-1$, go to (**VI**). If the interval is empty and $r = p$; then terminate the algorithm, no solution exists.

(**VI**) Check the inequality from Lemma 3.7. If the interval is non-empty, then $r := r+1$ and go back to (**V**). If the interval is empty, then terminate the algorithm as all the subsequent systems have no solution (Lemma 7.3).

Theorem 7.4. *The algorithm A2 is correct and terminates in $O(m^2 n)$ time.*

Proof. Correctness follows immediately from the work in this chapter.

To see the complexity, observe the following.

In step (**I**), the quantities K_t and γ_t are found in $O(m)$ time.

In step (**II**) we identify $O(m)$ parametrised systems and scale each (where each such system has $O(m)$ equations). This takes $O(m^2)$ time.

Finding the principal solution for the system S_0 takes $O(mn)$ time - this is the complexity for step (**III**).

Steps (**V**) and (**VI**) iterate $O(m)$ times. For each iteration, the interval in (**V**) is found in $O(mn)$ time, the inequality in (**VI**) is checked in the same time. These iterations therefore take $O(m^2 n)$ time.

In total, we conclude the algorithm A2 runs in $O(m^2 n)$ time. □

Example 7.5. *Recall Example 7.1. Let us apply the algorithm to that example. We had*

$$S_0 : \begin{pmatrix} -1 & 0 & 1 \\ -1 & -2 & 0 \\ 2 & 0 & 0 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix},$$

but here the principal solution $\overline{x} = (-2,-1,-1)^T$ is not a solution and so we consider S_1.

Recall

$$S_1 : \begin{pmatrix} -1 & 0 & 1 \\ -1 & -2 & 0 \\ 2 & 0 & 0 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \alpha \\ 0 \\ 0 \\ 0 \end{pmatrix},$$

where $\alpha \in [0, 2]$. Applying Theorem 3.14 we see that the unique solution is $\alpha = 1$. Making the substitution we obtain the one-sided system

$$\begin{pmatrix} -1 & 0 & 1 \\ -1 & -2 & 0 \\ 2 & 0 & 0 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

with principal solution $\overline{x} = (-2, -1, 0)^T$. Relating back to the original system

$$\begin{pmatrix} -1 & 0 & 1 \\ -1 & -2 & 0 \\ 2 & 0 & 0 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & -1 \\ 0 & -1 \\ 0 & -3 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

we get the corresponding solution $x = (-2, -1, 0)^T, y = (0, 0)^T$.

7.6 Summary

We saw that we could consider the two-sided system with separated variables as a sequence of one-sided parametrised systems, as studied in chapter 3. We could construct a strongly polynomial algorithm for finding a solution in an iterative process. Further, the algorithm could be modified depending if the user wishes to find a solution or describe all solutions. This follows since we can partition the problem into a set of parametrised systems, for each of which we can describe the full solution set as in chapter 3.

8. Two-sided systems of homogeneous equations - minimally active and essential systems

8.1 Introduction

The focus in this chapter is in demonstrating a pivotal role of the matrix $A \oplus B$ (denoted C throughout this chapter) and its max-algebraic permanent for solving two-sided linear systems $Ax = Bx$ where A and B are finite, square matrices. Note that the results of this chapter can be extended to the non-finite case, provided that the permanent of C is finite. We study only the finite case here.

We study two special types of the system $Ax = Bx$. The first type, called *minimally active*, is defined by the requirement that for every non-trivial solution x, the maximum on each side of every equation is attained exactly once (see Definition 8.4). For the second type, called *essential systems*, we require that every component of any non-trivial solution is *active* on at least one side of at least one equation. By active component x_j we mean that there exists $i \in M$ such that in the matrix A, say, we have $\bigoplus_{t \in N} a_{it} x_t = a_{ij} x_j$. So "active" essentially means "attains a maximum". Equivalently, all non-trivial solutions are finite (see Definition 8.17).

We prove that in every solvable two-sided max-linear system of minimally active or essential type, all positions in $C := A \oplus B$ active in any optimal permutation for the assignment problem for C, are also active for some non-trivial solution (any non-trivial solution in the minimally active case). This enables us to deduce conditions on a solution x for which it is possible in some cases to find x in polynomial time. It is also proved that any essential system can be transformed to a minimally active system in polynomial time. Theorem 8.6 is the key theorem of this chapter.

The contents of this chapter have been submitted to the journal Discrete Applied Mathematics.

8.2 Notes

Let $D = (d_{ij}) \in \overline{\mathbb{R}}^{m\times n}$ and $x \in \overline{\mathbb{R}}^{n}$. Then

$$\mathcal{A}(x, D) := \left\{ (i,j) \in M \times N : d_{ij}x_j = \bigoplus_{t \in N} d_{it}x_t \right\}, \tag{8.1}$$

where "$\mathcal{A}$" is for "active". Let $i, j \in N$ and $x \in V(A, B)$ (see (2.7)). If $(i,j) \in \mathcal{A}(x, A)$ ($(i,j) \in \mathcal{A}(x, B)$), then we say (i,j) is $x - active$ in A (B). It follows that $\mathcal{A}(x, A)$ ($\mathcal{A}(x, B)$) is the set of positions which are $x - active$ in A (in B).

Let $i \in M$ and define

$$av_{x,D}(i) := \{k \in N : (i,k) \in \mathcal{A}(x, D)\}, \tag{8.2}$$

where "*av*" stands for "*active variable*".

Finally, let $j \in N$ and define

$$ae_{x,D}(j) := \{i \in M : (i,j) \in \mathcal{A}(x, D)\}, \tag{8.3}$$

where "*ae*" stands for "*active equation*".

Suppose that $A = (a_{ij}) \in \overline{\mathbb{R}}^{m \times n}$ and $B = (b_{ij}) \in \overline{\mathbb{R}}^{m \times n}$ are given. The problem of solving the two-sided linear system (A, B) is the task of finding $x \in \overline{\mathbb{R}}^{n}, x \neq \epsilon$ (a non-trivial solution) such that

$$Ax = Bx. \tag{8.4}$$

Due to the finiteness of A and B, it can be shown (see chapter 2) that a non-trivial solution exists if and only if a finite solution exists. As such, we restrict our attention to finding finite solutions to (8.4).

Recall

$$V(A, B) = \{x \in \mathbb{R}^n : Ax = Bx\}. \tag{8.5}$$

We will assume in the rest of this section that $m = n$ and $C := A \oplus B$. Recall from Chapter 2, page 19 that $ap(C) = ap(diag(v)\, C)$ for all $v \in \mathbb{R}^n$.

Note that

$$(\forall x \in V(A, B))\, (\forall i) \bigoplus_{t \in N} a_{it} x_t = \bigoplus_{t \in N} b_{it} x_t = \bigoplus_{t \in N} c_{it} x_t. \tag{8.6}$$

Hence $\mathcal{A}(x, C) = \mathcal{A}(x, A) \cup \mathcal{A}(x, B)$ and for all i we have $av_{x,C}(i) = av_{x,A}(i) \cup av_{x,B}(i)$.

For $x \in V(A, B)$ and $i \in N$, we have $(\exists j_1, j_2)$ such that $(i, j_1) \in \mathcal{A}(x, A)$ and $(i, j_2) \in \mathcal{A}(x, B)$. Note that j_1 and j_2 are not necessarily distinct.

We have $(\forall x \in V(A, B))\, (\forall i)\, |av_{x,A}(i)|, |av_{x,B}(i)| \geq 1$.

It is easily shown that if $V(A, B) \neq \emptyset$, then there exists $x \in V(A, B)$ such that for all $j, ae_{x,C}(j) \neq \emptyset$. As such, we define

$$\tilde{V}(A, B) := \{x \in V(A, B) : (\forall j)\, ae_{x,C}(j) \neq \emptyset\}. \tag{8.7}$$

In the rest of this chapter, we are interested in finding solutions $x \in \tilde{V}(A, B)$.

Definition 8.1. *Let $x \in \tilde{V}(A, B)$ and let $\sigma \in ap(C)$, σ is called "$x - optimal$" if*

$(\forall i)\,(i,\sigma(i)) \in \mathcal{A}(x,C)$.

Example 8.2.

$$A = \begin{pmatrix} 3 & 8 & 2 \\ 7 & 1 & 4 \\ 0 & 5 & 3 \end{pmatrix}, \quad B = \begin{pmatrix} 5 & 5 & 5 \\ 3 & 4 & 5 \\ 5 & 3 & 2 \end{pmatrix}, \quad C = \begin{pmatrix} 5 & 8 & 5 \\ 7 & 4 & 5 \\ 5 & 5 & 3 \end{pmatrix} = A \oplus B.$$

Note that

$$ap(C) = \{(1,2)(3), (1,2,3)\}.$$

Let $x = (0,-1,2)^T$. *Then,* $\sigma = (1,2)(3) \in ap(C)$ *is* x*-optimal since* $(1,2), (2,1), (3,3) \in \mathcal{A}(x)$.

Similarly, $\sigma' = (1,2,3) \in ap(C)$ *is* x*-optimal since* $(1,2), (2,3), (3,1) \in \mathcal{A}(x)$.

Lemma 8.3. *If* $\left(\exists x \in \tilde{V}(A,B)\right)(\exists \sigma \in ap(C))\,\sigma$ *is* $x - optimal$*, then* $(\forall \sigma' \in ap(C))\,\sigma'$ *is* $x - optimal$.

Proof. Let $\sigma \in ap(C)$ be $x - optimal$, so that $(\forall i)\,c_{i,\sigma(i)}x_{\sigma(i)} = \bigoplus_{t\in N} c_{it}x_t$. Now consider $\sigma' \in ap(C), \sigma' \neq \sigma$.

Define $I := \{i \in N : \sigma'(i) = \sigma(i)\}, \overline{I} := N \backslash I$.

Clearly, $(\forall i \in I)\,(i,\sigma'(i)) \in \mathcal{A}(x)$, since $(i,\sigma'(i)) = (i,\sigma(i))$. If for all $i \in \overline{I}$, $c_{i,\sigma'(i)}x_{\sigma'(i)} = c_{i,\sigma(i)}x_{\sigma(i)}$, then $\left(\forall i \in \overline{I}\right)(i,\sigma'(i)) \in \mathcal{A}(x)$. It follows in this case that σ' is $x - optimal$. So suppose there exists $i \in \overline{I}$ such that

$$c_{i,\sigma'(i)}x_{\sigma'(i)} \neq c_{i,\sigma(i)}x_{\sigma(i)}. \tag{8.8}$$

Now note that $\sigma, \sigma' \in ap(C)$, which implies $\sigma, \sigma' \in ap(C_x)$. Therefore

$$\sum_{i\in N} c_{i,\sigma'(i)}x_{\sigma'(i)} = \sum_{i\in N} c_{i,\sigma(i)}x_{\sigma(i)}, \tag{8.9}$$

which is equivalent to

$$\sum_{i\in I} c_{i,\sigma'(i)}x_{\sigma'(i)} + \sum_{i\in \overline{I}} c_{i,\sigma'(i)}x_{\sigma'(i)} = \sum_{i\in I} c_{i,\sigma(i)}x_{\sigma(i)} + \sum_{i\in \overline{I}} c_{i,\sigma(i)}x_{\sigma(i)}. \tag{8.10}$$

Hence

$$\sum_{i\in \overline{I}} c_{i,\sigma'(i)}x_{\sigma'(i)} = \sum_{i\in \overline{I}} c_{i,\sigma(i)}x_{\sigma(i)}. \tag{8.11}$$

It follows from (8.8) and (8.11) that $\left(\exists u\in \overline{I}\right) c_{u,\sigma'(u)}x_{\sigma'(u)} > c_{u,\sigma(u)}x_{\sigma(u)}$, but this contradicts the assumption that $(u,\sigma(u))\in \mathcal{A}(x)$. We conclude that σ' is $x-optimal$. □

8.3 Minimally active systems

Definition 8.4. *The system* (A,B) *is called minimally active if* $\left(\forall x\in \tilde{V}(A,B)\right)(\forall i\in N)\,|av_{x,A}(i)| = |av_{x,B}(i)| = 1.$

Interestingly, we have the following property for minimally active systems. Recall (2.7) and (8.7) for the defnitions of $V(A,B)$ and $\tilde{V}(A,B)$ respectively.

Lemma 8.5. *Let* (A,B) *be a minimally active system. Then* $\tilde{V}(A,B) = V(A,B)$.

Proof. We only need to show that for the minimally active system (A,B), we have $V(A,B)\subseteq \tilde{V}(A,B)$.

Suppose for a contradiction that $\left(\exists x\in V(A,B)\setminus \tilde{V}(A,B)\right)$. Let $j\in N$ such that $ae_x(j)=\emptyset$. We increase x_j until x_j becomes active in some equation i (this will happen due to the finiteness of A and B), producing a new solution x'. But, since $|av_{x,A}(i)| = |av_{x,B}(i)| = 1$, it follows that, say, $|av_{x',A}(i)| = 2$, contradicting the assumption of minimal activity. □

It is known ([25], Lemma 7.1.1) that V is a subspace. Lemma 8.5 confirms that $\tilde{V}$ is a subspace also when (A,B) is minimally active. Note that in the remainder of the section, we will always have $V=\tilde{V}$.

We state now the main result of this section.

Theorem 8.6. *Let* (A, B) *be a minimally active system. Then* $V(A, B) \neq \emptyset$ *if and only if* $(\exists x \in V(A, B))\,(\forall \sigma \in ap(A \oplus B))\,\sigma$ *is* $x-$*optimal.*

This Theorem may be very useful. Such a result would allow us to deduce important information about a solution, without any a priori knowledge of what such a solution might be. Further, this information is obtained by finding $ap(C)$, something which is easily done (in polynomial time) with the help of, say, the Hungarian algorithm [56].

Remark 8.7. *The 'if' statement of Theorem 8.6 is trivial, we need the proof of the 'only if' part only. Also, due to Lemma 8.3, we only need to show there exists* $x \in V(A, B)$ *such that* σ *is* $x-$*optimal for some sigma* $\in ap(C)$.

Note that we have not given any way to check that a system is minimally active in general. Example 8.8, however, is easily shown to be minimally active. To see this, first apply the Cancellation Rule, yielding the system

$$\begin{cases} 8x_2 = 5x_1 \oplus 5x_3 \\ 7x_1 = 4x_2 \oplus 5x_3 \\ 5x_1 = 5x_2 \oplus 3x_3. \end{cases}$$

Equivalently,

$$\begin{cases} x_2 = (-3)\,x_1 \oplus (-3)\,x_3 \\ x_1 = (-3)\,x_2 \oplus (-2)\,x_3 \\ x_1 = x_2 \oplus (-2)\,x_3. \end{cases}$$

Without loss of generality, $x_1 = 0$. We see also that $x_1 \geq x_2 > (-3)\,x_2$. Therefore, $x_1 = (-2)\,x_3$ and so $x_3 = 2$. It follows that $x_2 = (-3) \oplus (-1) = -1$. We have then, after scaling, the unique solution is $x^T = (0, -1, 2)$. Note that for x as defined above we have $(\forall i = 1, 2, 3)\, |av_{x,A}(i)| = |av_{x,B}(B)| = 1$.

Before the proof of Theorem 8.6, we give some examples to show the power of the theorem. Example 8.9 is not minimally active but the application of Theorem 8.6 still yields a solution - we will see why in Section 8.4.

Example 8.8.

$$A = \begin{pmatrix} 3 & 8 & 2 \\ 7 & 1 & 4 \\ 0 & 5 & 3 \end{pmatrix}, \quad B = \begin{pmatrix} 5 & 5 & 5 \\ 3 & 4 & 5 \\ 5 & 3 & 2 \end{pmatrix}, \quad C = \begin{pmatrix} 5 & \textcircled{8} & 5 \\ \textcircled{7} & 4 & \textcircled{5} \\ \textcircled{5} & 5 & \textcircled{3} \end{pmatrix}.$$

We see that $maper(C) = 18$ *and* $ap(C) = \{(1,2)(3), (1,2,3)\}$. *So if* $V(A,B) \neq \emptyset$*, then* $(\exists x \in V(A,B))\,(1,2), (2,1), (3,3), (2,3), (3,1) \in \mathcal{A}(x)$*, in which case we have*

$$\begin{cases} 8x_2 = 5x_1 \oplus 5x_3 \\ 7x_1 = 5x_3 \\ 5x_1 = 3x_3. \end{cases}$$

We set without loss of generality $x_1 = 0$ *and deduce* $x_3 = 2 \Rightarrow x_2 = -1$*, hence* $x^T = (0, -1, 2)$*. Indeed,* $x^T = (0, -1, 2)$ *is a solution.*

Example 8.9.

$$A = \begin{pmatrix} -4 & 3 & 0 & 2 \\ 5 & -1 & 6 & 3 \\ 7 & 3 & 0 & 4 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 2 & 6 & 1 \\ 3 & 5 & 0 & 7 \\ 2 & 12 & 6 & 3 \end{pmatrix}, \quad C = \begin{pmatrix} 0 & 3 & 6 & 2 \\ 5 & 5 & 6 & 7 \\ 7 & 12 & 6 & 4 \end{pmatrix}.$$

By simply repeating the last rows, we can convert this to an equivalent (identical solution *set) square system as follows.*

$$A' = \begin{pmatrix} -4 & 3 & 0 & 2 \\ 5 & -1 & 6 & 3 \\ 7 & 3 & 0 & 4 \\ 7 & 3 & 0 & 4 \end{pmatrix}, \quad B' = \begin{pmatrix} 0 & 2 & 6 & 1 \\ 3 & 5 & 0 & 7 \\ 2 & 12 & 6 & 3 \\ 2 & 12 & 6 & 3 \end{pmatrix}, \quad C' = \begin{pmatrix} 0 & 3 & 6 & 2 \\ 5 & 5 & 6 & 7 \\ 7 & 12 & 6 & 4 \\ 7 & 12 & 6 & 4 \end{pmatrix}.$$

We apply the Hungarian method for finding $ap\,(C)$.

$$C' = \begin{pmatrix} 0 & 3 & 6 & 2 \\ 5 & 5 & 6 & 7 \\ 7 & 12 & 6 & 4 \\ 7 & 12 & 6 & 4 \end{pmatrix} \to \begin{pmatrix} -6 & -3 & 0 & -4 \\ -2 & -2 & -1 & 0 \\ -5 & 0 & -6 & -8 \\ -5 & 0 & -6 & -8 \end{pmatrix}$$

$$\to \begin{pmatrix} -4 & -3 & 0 & -4 \\ 0 & -2 & -1 & 0 \\ -3 & 0 & -6 & -8 \\ -3 & 0 & -6 & -8 \end{pmatrix} \to \begin{pmatrix} -4 & -6 & \textcircled{0} & -4 \\ 0 & -5 & -1 & \textcircled{0} \\ \textcircled{0} & \textcircled{0} & -3 & -5 \\ \textcircled{0} & \textcircled{0} & -3 & -5 \end{pmatrix}.$$

We have highlighted the elements of all optimal permutations from $ap\,(C')$. *In the original matrix* C' *(before Hungarian method is applied), this corresponds to*

$$\begin{pmatrix} 0 & 3 & \textcircled{6} & 2 \\ 5 & 5 & 6 & \textcircled{7} \\ \textcircled{7} & \textcircled{12} & 6 & 4 \\ \textcircled{7} & \textcircled{12} & 6 & 4 \end{pmatrix}.$$

Now, without loss of generality $x_1 = 0$. *From the third (fourth) row:* $7x_1 = 12x_2 \Rightarrow x_2 = -5$. *From the first row:* $6x_3 = (0x_1) \oplus (3x_2) \oplus (2x_4) = 0 \oplus (2x_4) \Rightarrow x_3 = -6 \oplus (-4)\,x_4$.

(Note we are using the Cancellation Rule by no longer distinguishing between matrices A and B, considering only entries of the matrix C as possible active entries).
From row two: $7x_4 = 5x_1 \oplus 5x_2 \oplus 6x_3 = 5 \oplus 2x_4$. *So* $7x_4 = 5 \Leftrightarrow x_4 = -2$.
Then, $x_3 = -6 \oplus (-4 - 2) = -6$.
Finally then, we have $x^T = (0, -5, -6, -2)$, *which we can confirm is a solution.*

It is not at all obvious that we will always have enough information about x to find it exactly, as we did in Examples 8.8 and 8.9 (since finding $ap\,(C)$ does not necessarily highlight all elements of $\mathcal{A}\,(x)$). This leads to some interesting open questions, which we discuss in Section 8.6. For now though, we focus on the task of proving Theorem 8.6.

Lemma 8.10. *If for some $x \in V\,(A, B)$, there exists a permutation $\sigma \in P_n$ such that for all $i \in N$ we have $(i, \sigma\,(i)) \in \mathcal{A}\,(x)$, then $\sigma \in ap\,(A \oplus B)$. Further, σ is $x - optimal$.*

Proof. We prove this by showing $\sigma \in ap\,(diag\,(x)\,C)$, by contradiction.

Suppose not, then $(\exists i \in N)\, c_{i,\sigma(i)} x_{\sigma(i)} < \bigoplus_{t \in N} c_{it} x_t$, which contradicts $(i, \sigma\,(i)) \in \mathcal{A}\,(x)$.

Further, σ is $x - optimal$ by definition. □

Our task is clear. As a consequence of Lemma 8.10, it suffices to show that if $V\,(A, B) \neq \emptyset$, then there exists $x \in V\,(A, B)$ such that the elements of $\mathcal{A}\,(x)$ admit a permutation σ, in which case $\sigma \in ap\,(C)$ by Lemma 8.10. Theorem 8.6 then follows as a result of Lemma 8.3. See also Remark 8.7.

The problem of finding a permutation is equivalent to the 1-factor problem in bipartite graphs, which in turn is equivalent to the problem of finding a perfect matching in a corresponding bipartite graph.

For clarity of exposition, we will use different symbols for row/column indices.

Definition 8.11. *Let $A, B \in \mathbb{R}^{n \times n}$ such that $V\,(A, B) \neq \emptyset$. Let $x \in V\,(A, B)$ and define the bipartite graph $G_x\,(A, B)$ (or simply G_x when it is clear to which system of* *matrices we are referring) with vertex sets S_x and T_x as follows. $S_x = \{s_1, \ldots, s_n\}, T_x = \{t_1, \ldots t_n\}, (\forall s_i \in S)(\forall t_j \in T)\, s_i t_j \in E(G_x)$ if and only if $(i,j) \in \mathcal{A}(x)$, (effectively, $S_x = T_x = N$). We also define a $3-$colouring of the edges of G_x as follows:*

$$c(s_i t_j) = \begin{cases} c_1, if\, (i,j) \in \mathcal{A}(x,A) \setminus \mathcal{A}(x,B) \\ c_2, if\, (i,j) \in \mathcal{A}(x,B) \setminus \mathcal{A}(x,A) \\ c_3, if\, (i,j) \in \mathcal{A}(x,A) \cap \mathcal{A}(x,B). \end{cases}$$

We call c the activity colouring. We denote by $c(G_x)$ the graph G_x edge-coloured with c.

Definition 8.12.

- *For $s \in S_x$, define $N(s) := \{t \in T_x : st \in E(G_x)\}$.*
- *For $s \in S_x$, for $r \in \{1,2,3\}$, define*
  $N_{c_r}(s) := \{t \in T_x : t \in N(s) \text{ and } c(st) = c_r\}$.
- *For $S' \subseteq S_x$, define $N(S') := \cup_{s \in S'} N(s)$.*
- *For $t \in T_x$, define $N(t) := \{s \in S_x : st \in E(G_x)\}$.*
- *For $T' \subseteq T_x$, define $N(T') := \cup_{t \in T'} N(t)$.*

Remark 8.13. *We may refer to vertex sets S or T when there is no ambiguity for which vector x we are considering.*

Essentially,

$$N_{c_1}(s) = av_{x,A}(s) \setminus av_{x,B}(s),$$
$$N_{c_2}(s) = av_{x,B}(s) \setminus av_{x,A}(s),$$
$$N_{c_3}(s) = av_{x,A}(s) \cap av_{x,B}(s).$$

For $x \in V(A,B)$, vertex $s_i \in S_x$ corresponds to equation i in the system $Ax = Bx$ and so we may talk about $i \in S_x$ without any confusion. Similarly, $t_j \in T_x$ corresponds to x_j, so we may talk about $j \in T_x$, or even $x_j \in T_x$.

From now we assume that $V(A,B) \neq \emptyset$. Our goal then, is to show there is an $x \in V(A,B)$ such that G_x has a perfect matching. Equivalently, we show there is an $x \in V(A,B)$ such that the size of the minimum vertex cover in G_x is n, due to the following Lemma, which follows from König-Egervary Theorem [55].

Lemma 8.14. *Let G_x be a bipartite graph with vertex sets S, T such that $|S| = |T| = n$. For any x, a perfect matching in G_x exists if and only if the size of a minimum vertex cover is n.*

We are ready now for the proof of the main result, Theorem 8.6. We complete the proof via the following equivalent Lemma.

Lemma 8.15. *If (A,B) is minimally active, then $(\exists x \in V(A,B))$ such that the size of the minimum vertex cover in G_x is n.*

Proof. Let (A,B) be minimally active and $x \in V(A,B)$. Consider the activity colouring; $c(G_x)$. Note that S_x covers all edges in G_x and so the size of the minimum vertex cover is always less than or equal to n. Let W be a minimum vertex covering of G_x. If $|W| = n$, then we are done, so suppose $|W| \leq n-1$.

Also note that $(\forall x' \in V(A,B))(\forall s \in S_{x'})\, 1 \leq |N(s)| \leq 2$, due to $x' \in V(A,B)$ (≥ 1) and the minimal activity property (≤ 2). Note $|N(s)| = 2$ corresponds to the case when $|N_{c_1}(s)| = |N_{c_2}(s)| = 1$ and $|N_{c_3}(s)| = 0$. Also, $|N(s)| = 1$ corresponds to the case when $|N_{c_1}(s)| = |N_{c_2}(s)| = 0$ and $|N_{c_3}(s)| = 1$.

Define $W_S := W \cap S, W_T := W \cap T$. Define $\overline{W_S} := S \backslash W_S, \overline{W_T} := T \backslash W_T$. If we have $|\overline{W_T}| \leq |W_S|$, then it follows that $|W| \geq n$, a contradiction. So $|\overline{W_T}| > |W_S|$.

Since $(\forall s \in S_x) |N(s)| \geq 1$ $(x \in V(A,B))$ and $(\forall t \in T) |N(t)| \geq 1$ (definition of $\tilde{V}(A,B)$ and $V = \tilde{V}$ due to (A,B) being minimally active), it follows that $|W_S|, |W_T| \geq 1$. In fact, from the definition of W, we have that

$$\left(\forall j \in \overline{W_T}\right)(\exists s \in W_S)(s,j) \in E(G_x).$$

Every $s \in W_S$ has a neighbour in $\overline{W_T}$ (else $N(s)$ covered by W_T and by removing s from W we obtain a smaller vertex cover). Also, for every $s \in \overline{W_S}, s$ has no neighbours in $\overline{W_T}$ (by definition of W). It follows that $N\left(\overline{W_T}\right) \subseteq W_S$. In fact, it follows that $N\left(\overline{W_T}\right) = W_S$, though we only need that $N\left(\overline{W_T}\right) \subseteq W_S$. Now:

If $\left(\forall s \in N\left(\overline{W_T}\right)\right) |N_{c_1}(s) \cap \overline{W_T}| = |N_{c_2}(s) \cap \overline{W_T}| = 1$ (recall then that $\left(\forall s \in N\left(\overline{T}\right)\right) N(s) \cap W_T = \emptyset$), then we can define the solution x' by:

$$x'_k := \begin{cases} \alpha x_k, \text{ if } x_k \in \overline{W_T} \\ x_k, \text{ otherwise,} \end{cases}$$

where

$$\alpha := \bigoplus_{i \in \overline{W_S}}' \left\{ \bigoplus_{j \in \overline{W_T}}' \left[\left(\bigoplus_{t \in N} c_{it} x_t \right) (c_{ij} x_j)^{-1} \right] \right\}.$$

The vector x' is a solution to equations corresponding to $\overline{W_S}$ since $\left(\forall j \in \overline{W_T}\right)\left(\forall i \in \overline{W_S}\right)(i,j) \notin \mathcal{A}(x)$. The constant α is defined so that the variables of $\overline{W_T}$ are increased to exactly the first point where (u,t) becomes $x-\textit{active}$ for some $u \in \overline{W_S}$ and some $t \in \overline{W_T}$.

It follows that $x' \in V(A,B)$ and we then have that $\left(\exists u \in \overline{W_S}\right)$ at least one of the following holds:

- $|av_{x',A}(u) \cap av_{x',B}(u)| \geq 2$;

- $|av_{x',A}(u)| \geq 2$;
- $|av_{x',B}(u)| \geq 2$; or
- $|av_{x',A}(u) \cap av_{x',B}(u)| = 1$ and $|av_{x',A}(u)| + |av_{x',B}(u)| \geq 1$.

In any case, we contradict the assumption that (A, B) is minimally active.

It follows then that (for original solution x) $\left(\exists s_1 \in N\left(\overline{W_T}\right) = W_S\right)$ with exactly one neighbour in $\overline{W_T}$ (say t_1, with $c(s_1 t_1) = c_{r_1}, r_1 \in \{1, 2, 3\}$), and at most one neighbour in W_T (no such neighbour in the case $r_1 = 3$, exactly one such neighbour otherwise).

Consider now $\overline{W_T} \setminus \{t_1\}$. Note that

$$\emptyset \neq N\left(\overline{W_T} \setminus \{t_1\}\right) \subseteq W_S \setminus \{s_1\}.$$

As before, if $\left(\forall s \in N\left(\overline{W_T} \setminus \{t_1\}\right)\right) |N_{c_1}(s) \cap \left(\overline{W_T} \setminus \{t_1\}\right)| = |N_{c_2}(s) \cap \left(\overline{W_T} \setminus \{t_1\}\right)| = 1$, then we can define a solution x' which contradicts the assumption of minimal activity of (A, B).

Again then, we conclude $\left(\exists s_2 \in N\left(\overline{W_T} \setminus \{t_1\}\right) \subseteq W_S \setminus \{s_1\}\right)$ with exactly one neighbour in $\overline{W_T} \setminus \{t_1\}$ (say t_2 with $c(s_2 t_2) = c_{r_2}, r_2 \in \{1, 2, 3\}$), and at most one neighbour in $T \setminus \left(\overline{W_T} \setminus \{t_1\}\right)$.

We continue in this way, pairing off elements of W_S and $\overline{W_T}$. Eventually, since $|\overline{W_T}| > |W_S|$, we run out of vertices in W_S. We have defined $t_1, \ldots, t_{|W_S|}$ and $s_1, \ldots, s_{|W_S|}$. Let $T' := \overline{W_T} \setminus \left\{t_1, \ldots, t_{|W_S|}\right\} \neq \emptyset$. It follows that $N(T') \subseteq W_S \setminus \left\{s_1, \ldots, s_{|W_S|}\right\} = \emptyset$. This contradicts the assumption that $(\forall t \in T) |N(t)| \geq 1$.

We conclude that it was our initial assumption, namely that $|\overline{W_T}| > |W_S|$, which was wrong. It follows that for our original x, the size of the minimum vertex cover in G_x is n. In fact, since $x \in V(A, B)$ was arbitrary, the result holds true for all $x \in V(A, B)$. $\square$

In fact, we proved a stronger result than Theorem 8.6. To be exact, we showed that the conditions of Theorem 8.6 hold for all solutions, not just one.

Theorem 8.16. *Let (A, B) be a minimally active system. Then $V(A, B) \neq \emptyset$ if and only if $(\forall x \in V(A, B))\,(\forall \sigma \in ap(A \oplus B))\,\sigma$ is $x-$optimal.*

8.4 Essential systems

In this section we show that we can generalise the results of Section 8.3 to a wider class of systems, which we call *essential systems*.

Definition 8.17. *Let $A, B \in \mathbb{R}^{n\times n}$. We say that (A, B) is essential if $V(A, B) = \tilde{V}(A, B) \neq \emptyset$, where the sets $V(A, B)$ and $\tilde{V}(A, B)$ are defined in (2.7) and (8.7) respectively.*

We use the term essential for the following reason. If $V(A, B) \neq \tilde{V}(A, B)$, then there exists $x \in V(A, B) \backslash \tilde{V}(A, B), x \neq \epsilon$, such that $Ax = Bx$. By definition of V and $\tilde{V}$, there exists $j \in N$ such that $x_j = \epsilon$. It follows that the system (A', B') (where A' and B' are obtained from A and B respectively by removing column j) has a non-trivial solution. Variable j is *inessential* in the original system.

It can be shown that (A, B) in Example 8.28 is an essential system which is not minimally active. To see that (A, B) is not minimally active, consider the solution $x^T = (1, 0, 2)$. To see that the system is essential, note that it is equivalent to show for $i = 1, 2, 3$ that there is no non-trivial solution x for which $x_i = \epsilon$. It is easy to check that no such solution exists. As an example, suppose for a contradiction there exists $x \in \overline{\mathbb{R}}^n, x_1 = \epsilon, x \neq \epsilon$ such that $Ax = Bx$. After applying the Cancellation Rule, it follows that

$$\begin{cases} 3x_3 = 5x_2 \oplus 3x_3 \\ 3x_2 = 1x_3 \\ \epsilon = 4x_2 \oplus 2x_3, \end{cases}$$

contradicting $x \neq \epsilon$.

For the remainder of this Section, (A, B) is an essential system. As in Section 8.3, we will only use the notation $V(A, B)$ (or simply V where no confusion can arise) but it should be remembered that $V = \tilde{V}$.

We generalise the results of Section 8.3 by showing that essential systems are related to minimally active systems. The following Lemma is key to showing that this can be done.

Lemma 8.18. *Let $A, B \in \mathbb{R}^{n\times n}$ such that (A, B) is essential and not minimally active. Let $z \in V(A, B), r \in N$ such that, say, $|av_{z,A}(r)| \geq 2$ (the case for $|av_{z,B}(r)| \geq 2$ is similar). Let $s \in av_{z,A}(r)$. Then there exists $\delta^* > 0$ sufficiently small such that for all $0 < \delta \leq \delta^*$ the matrices $A^{(\delta)}, B^{(\delta)}$ defined by*

$$a_{uv}^{(\delta)} = \begin{cases} a_{rs}\delta^{-1}, & \text{if } u{=}r, v{=}s \\ a_{uv}, & o.w. \end{cases}$$

$$B^{(\delta)} = B$$

satisfy the following:

1. $z \in V\left(A^{(\delta)}, B^{(\delta)}\right)$.

2. $av_{z,A^{(\delta)}}(r) = av_{z,A}(r) \setminus \{s\}$.

3. $\emptyset \neq V\left(A^{(\delta)}, B^{(\delta)}\right) \subseteq V(A, B)$.

4. $\left(\forall x \in V\left(A^{(\delta)}, B^{(\delta)}\right)\right)(\forall i \in N)\, av_{x,A^{(\delta)}}(i) \subseteq av_{x,A}(i)$,
$av_{x,B^{(\delta)}}(i) \subseteq av_{x,B}(i)$.

5. *For all* $x \in V\left(A^{(\delta)}, B^{(\delta)}\right), (r, s) \notin \mathcal{A}\left(x, A^{(\delta)}\right)$.

6. $\left(A^{(\delta)}, B^{(\delta)}\right)$ *is essential.*

Before we prove Lemma 8.18, we have some comments.

Remark 8.19. *For* $x \in V(A, B), i \in N$ *and* $s_i \in S_x$*, consider the activity colouring* $c(G_x)$*. Then for* s_i*, at least one of the following holds:*

- s_i *is incident with an edge of colour* c_3*;*
- s_i *is incident with an edge of colour* c_1 *and an edge of colour* c_2*.*

Note that if s_i *is incident with only one edge, then that edge is coloured* c_3 *(the converse is not true in general).*

Definition 8.20. *If for all* $j_1, j_2 \in T$*, there is a path from* j_1 *to* j_2 *in* G_x*, then we say* G_x *is variable connected.*

Remark 8.21. *Since for all* $s \in S, |N(s)| \geq 1$*, it follows that* G_x *is variable connected if and only if* G_x *is connected. From now, we say only connected.*

Clearly, if $x = \alpha x'$ *for some* $x, x' \in V(A, B), \alpha \in \mathbb{R}$*, then* $G_x = G_{x'}$*. In the next Lemma we show that the converse is also true (that this cannot happen otherwise).*

Lemma 8.22. *If* $G \in \mathcal{G}$ *is connected, then for all* $x, x' \in V(A, B)$ *such that* $G_x = G_{x'} = G$*, there exists* $\alpha \in \mathbb{R}$ *such that* $x' = \alpha x$*. That is,* G *corresponds to exactly one solution (up to scaling).*

Proof. Let $G \in \mathcal{G}$ be connected and let $x \in V(A,B)$ such that $G_x = G$. Let $t_1, t_2 \in T, t_1 \neq t_2$ and let P be a path from t_1 to t_2 in G_x. Using the definition of $E(G_x)$, we see that $x_{t_1} x_{t_2}^{-1}$ is a fixed constant. That is

$$(\forall x \in V(A,B)) (\forall t_1, t_2) \, x_{t_1} x_{t_2}^{-1} = \Delta_{t_1,t_2} \text{ (constant).}$$

Since t_1, t_2 were arbitrary, the result follows.

Note that it doesn't matter which path we choose if many are available. If path P_1 yields $x_{t_1} x_{t_2}^{-1} = \alpha_1$, and path P_2 yields $x_{t_1} x_{t_2}^{-1} = \alpha_2, \alpha_1 \neq \alpha_2$, then $x \notin V(A,B)$, a contradiction.

□

Definition 8.23. *$x \in V(A,B)$ is called a connected solution if G_x is connected.*

Definition 8.24. *Let $x \in V(A,B)$. Denote by $com(x)$ the number of components of G_x.*

The following Lemma is given without proof but it should be noted that the ideas of the proof are similar to those used in the proof of Lemma 8.15.

Lemma 8.25. *Let $x \in V(A,B), x$ not connected and consider a component of G_x with the set of nodes X. Define $S' := X \cap S$ and $T' := X \cap T$. Then for all $s \in S'$ we have at least one of the following (by Remark 8.19):*

- $|N_{c_3}(s) \cap T'| \geq 1,$ *or*
- $|N_{c_1}(s) \cap T'|, |N_{c_2}(s) \cap T'| \geq 1.$

Define a new vector x' using

$$\alpha := \bigoplus_{i \notin S'}' \left\{ \bigoplus_{j \in T'}' \left[\left(\bigoplus_{t \in N} c_{it} x_t \right) (c_{ij} x_j)^{-1} \right] \right\},$$

and

$$x'_k := \begin{cases} \alpha x_k, & \text{if } x_k \in T' \\ x_k, & \text{if } x_k \notin T'. \end{cases}$$

Then $x' \in V(A,B)$ *and* $com(x') < com(x)$.

Let $x \in V(A,B), x$ not connected. By applying Lemma 8.25 repeatedly, we can transform x to a vector $\overline{x}$ such that $\overline{x}$ is connected and $\overline{x} \in V(A,B)$.

Note $\overline{x}$ may not be unique (the connected solution $\overline{x}$ depends on which component we use in Lemma 8.25). We denote by *connect* (x) the set of connected $\overline{x} \in V(A,B)$ that can be obtained from x in this way.

We are now ready for the proof of Lemma 8.18.

Proof. [Proof of Lemma 8.18] We are essentially reducing exactly one element in the system (A,B). Let $x \in V(A,B)$. Immediately, we can see that for all $\delta > 0$, $x \in V\left(A^{(\delta)}, B^{(\delta)}\right)$ (property (1)), which in turn means $V\left(A^{(\delta)}, B^{(\delta)}\right) \neq \emptyset$ (first part of property (3)). It is also clear for all $\delta > 0$ that $av_{x,A^{(\delta)}}(i) = av_{x,A}(i) \setminus \{j\}$ (property (2)).

- We show next that for all $\delta > 0$ sufficiently small $V\left(A^{(\delta)}, B^{(\delta)}\right) \subseteq V(A,B)$ (second part of property (3)). First note that if $a_{ij} = b_{ij} = c_{ij}$, then this follows immediately. To see this, let $\delta > 0$ and let
  $x \in V\left(A^{(\delta)}, B^{(\delta)}\right) \backslash V(A,B)$. It follows that $a_{ij}x_j > \bigoplus_{t \in N} b_{it}x_t \geq b_{ij}x_j = a_{ij}x_j$, a contradiction. So assume $a_{ij} \neq b_{ij}$.
  Let us start with a fixed $\delta_0 > 0$. If $V\left(A^{(\delta_0)}, B^{(\delta_0)}\right) \subseteq V(A,B)$, then we are done, so suppose not.
  Let $\Gamma(\delta)$ be the set of connected solutions in $V\left(A^{(\delta)}, B^{(\delta)}\right) \backslash V(A,B)$ for any $\delta > 0$. Since the number of connected bipartite graphs with the set of nodes S, T is finite and each corresponds to only one solution (up to multiples), it follows that $\Gamma(\delta_0)$ is finite (up to multiples).

Now, let $w \in \Gamma(\delta_0)$. We have $(\forall u)\, u \neq i, a_u w = b_u w$, and so $a_i w \neq b_i w$. Also, since $(\forall v \neq j)\, a_{iv}^{(\delta_0)} = a_{iv}$, it follows that $a_{ij} w_j > b_i w$ and $\bigoplus_{t \in N} a_{it} w_t = a_{ij} w_j$. (Recall that b_i denotes row i of B).

We have then that

$$\begin{cases} a_{ij}^{(\delta_0)} w_j \leq b_i w \Leftrightarrow a_{ij} \delta_0^{-1} w_j \leq b_i w \\ a_{ij} w_j > b_i w. \end{cases}$$

Therefore, $(\exists \delta_0')\, 0 < \delta_0' \leq \delta_0$ such that $a_{ij} (\delta_0')^{-1} w_j = b_i w$, thus for any $\delta_1, 0 < \delta_1 < \delta_0'$, we have

$$a_{ij}^{(\delta_1)} w_j > b_i w$$

and so $w \notin V\left(A^{(\delta_1)}, B^{(\delta_1)}\right) \backslash V(A, B)$. Note then that we also have for all multiples of w, namely $\alpha w, \alpha \in \mathbb{R}$, that $\alpha w \notin V\left(A^{(\delta_1)}, B^{(\delta_1)}\right) \backslash V(A, B)$.

Let $\delta_1 = \frac{1}{2}\delta_0'$ (for instance). Define $\delta(w) := \delta_1$, and define $\delta(w')$ in the same way for all $w' \in \Gamma(\delta_0)$. Since $\Gamma(\delta_0)$ is finite (up to multiples) and for all $w' \in \Gamma(\delta_0)$, for all $\alpha \in \mathbb{R}$ we have $\delta(w') = \delta(\alpha w')$, we can define

$$\delta^* := \min_{w \in \Gamma(\delta_0)} \delta(w) > 0.$$

We see that $\Gamma(\delta^*) = \emptyset$.

We now show that δ^* is sufficiently small so that $V\left(A^{(\delta^*)}, B^{(\delta^*)}\right) \subseteq V(A, B)$, as desired.

Let δ^* be as defined and suppose for a contradiction that there exists

$w \in V\left(A^{(\delta^*)}, B^{(\delta^*)}\right) \backslash V(A, B)$.

Vector w is not connected because $\Gamma(\delta^*) = \emptyset$ and since $w \notin V(A, B)$, we have $a_{ij} w_j > b_i w$. Let X_1 be the set of nodes of the component of G_w that contains w_j and define $S' := X_1 \cap S, T' = X_1 \cap T$. Note that $N(T') = S'$. Since w is not connected it follows that $(\forall u \in N(T') = S')$ either:

- $|N_{c_1}(u) \cap T'|, |N_{c_2}(u) \cap T'| \geq 1$; or
- $|N_{c_3}(u) \cap T'| \geq 1$.

As such, we may increase $w_k, k \in T'$ (similarly as in the proof of Lemma 8.15) by γ_1, say, until there is a new edge between T' and $S \backslash S'$. Call this new solution w'. We have $a_{ij} w'_j = a_{ij} w_j \gamma_1 > b_i w \gamma_1 \geq b_i w'$, and so $w' \in V\left(A^{(\delta^*)}, B^{(\delta^*)}\right) \backslash V(A, B)$ also. Repeat the procedure with the component of $G_{w'}$ containing w'_j, until we obtain a connected solution $\overline{w} \in V\left(A^{(\delta^*)}, B^{(\delta^*)}\right) \backslash V(A, B)$, contradicting $\Gamma(\delta^*) = \emptyset$.

- Note that property (6) follows immediately from property (3).

- Next, we show property (4) holds. Let $x \in V\left(A^{(\delta^*)}, B^{(\delta^*)}\right) \subseteq V(A, B)$ and let $u \in N, u \neq i$. Then $av_{x,A^{(\delta^*)}}(u) = av_{x,A}(u)$ and $av_{x,B}(u) = av_{x,B}(u)$, since $a_u^{(\delta^*)} = a_u$ and $b_u^{(\delta^*)} = b_u$. Also, since $a_{it} x_t \leq b_i x \; (\forall t)$ and $b_i^{(\delta^*)} = b_i$, it follows that $av_{x,A^{(\delta^*)}}(i) \subseteq av_{x,A}(i)$ and $av_{x,B^{(\delta^*)}}(i) = av_{x,B}(i)$.

- Finally, property (5). We show that for all $x' \in V\left(A^{(\delta^*)}, B^{(\delta^*)}\right)$ we have $j \notin av_{x',A^{(\delta^*)}}(i)$. Suppose for a contradiction $\left(\exists x' \in V\left(A^{(\delta^*)}, B^{(\delta^*)}\right)\right) j \in av_{x',A^{(\delta^*)}}(i)$, that is

$$a_{ij}^{(\delta^*)} x'_j = b_i^{(\delta^*)} x' \Leftrightarrow a_{ij} (\delta^*)^{-1} x'_j = b_i x'.$$

But then $a_{ij} x'_j > b_i x'$, and so $x' \notin V(A, B)$, a contradiction since $V\left(A^{(\delta^*)}, B^{(\delta^*)}\right) \subseteq V(A, B)$ (property (3)). (The entry $a_{ij}^{(\delta^*)}$ has essentially become a "dead element").

□

Remark 8.26. *Property (6) of Lemma 8.18 should serve to clarify that we are safe to refer only to V in the statement of Lemma 8.18 and that for both systems in the statement* *of Lemma 8.18, we still have $V = \tilde{V}$.*

The following Theorem is the final step to convert an essential system to a minimally active one.

Theorem 8.27. *Let $A, B \in \mathbb{R}^{n\times n}$, (A, B) essential and not minimally active. Then there is a sequence of systems $(A, B), \left(A^{(1)}, B^{(1)}\right), \ldots, \left(A^{(k)}, B^{(k)}\right)$, such that*

$$\emptyset \neq V\left(A^{(k)}, B^{(k)}\right) \subseteq V\left(A^{(k-1)}, B^{(k-1)}\right) \subseteq \cdots \subseteq V\left(A^{(1)}, B^{(1)}\right) \subseteq V(A, B),$$

and $\left(A^{(k)}, B^{(k)}\right)$ is minimally active, for some $k \in \mathbb{N}$.

Proof.

We construct a sequence of systems by repeated use of Lemma 8.18. We call this process “reduction”. It is not clear immediately that reduction terminates in finite time but if it does terminate in a finite number of steps with system $\left(A^{(k)}, B^{(k)}\right)$, then, since reduction has terminated, we have

$$\left(\forall x \in V\left(A^{(k)}, B^{(k)}\right)\right)(\forall i \in N)\, |av_{x,A^{(k)}}(i)| = |av_{x,B^{(k)}}(i)| = 1$$

and so $\left(A^{(k)}, B^{(k)}\right)$ is minimally active by definition. Then, from repeated use of Lemma 8.18, property (3), we have

$$\emptyset \neq V\left(A^{(k)}, B^{(k)}\right) \subseteq V\left(A^{(k-1)}, B^{(k-1)}\right) \subseteq \cdots \subseteq V\left(A^{(1)}, B^{(1)}\right) \subseteq V(A, B).$$

It remains to show that reduction does indeed terminate in a finite number of steps. In fact, we will show that reduction terminates in less than $2n^2$ iterations.

Suppose not for a contradiction and so we define systems
$\left(A^{(1)}, B^{(1)}\right), \ldots, \left(A^{\left(2n^2\right)}, B^{\left(2n^2\right)}\right)$ using Lemma 8.18. Define $\left(A^{(0)}, B^{(0)}\right) := (A, B)$. Now,

$(\forall r)\, 1 \leq r \leq 2n^2$, the transition from $\left(A^{(r-1)}, B^{(r-1)}\right)$ to $\left(A^{(r)}, B^{(r)}\right)$ is based on the reduction of exactly one entry of A or B. That is, either a_{ij} or b_{ij} for some i, j. We define $(i(r), j(r)) := (i, j)$.

Consider $(i(s), j(s))$, some $1 \leq s \leq 2n^2 - 1$. By Lemma 8.18, property (5), we have $\left(\forall x \in V\left(A^{(s)}, B^{(s)}\right)\right)(i(s), j(s)) \notin \mathcal{A}\left(x, A^{(s)}\right)$. Now, let $s + 1 \leq r \leq 2n^2$ and let $x' \in V\left(A^{(r)}, B^{(r)}\right)$. We claim that $(i(s), j(s)) \notin \mathcal{A}\left(x', A^{(r)}\right)$. To see this, note that (by repeated use of Lemma 8.18, property (4))

$$av_{x',A^{(r)}}(i(s)) \subseteq av_{x',A^{(r-1)}}(i(s)) \subseteq \cdots \subseteq av_{x',A^{(s)}}(i(s)).$$

We have essentially shown that once we reduce an element in the matrix A (in the matrix B) in the reduction process, then that element is now a "dead element" for all subsequent systems in the reduction process. Since there are n^2 elements in matrix A (in matrix B), there are a total of $2n^2$ elements in total which may be eliminated. This, with the fact that every system in the reduction process has non-empty solution set (Lemma 8.18, property (3)), leads us to conclude that reduction must terminate in at most $2n^2$ iterations. □

We give here an example of the process of reduction.

Example 8.28. *Let* $A = \begin{pmatrix} 4 & 0 & 3 \\ 0 & 3 & 0 \\ 3 & 0 & 0 \end{pmatrix}, B = \begin{pmatrix} 0 & 5 & 3 \\ 1 & 0 & 1 \\ 0 & 4 & 2 \end{pmatrix}$ *and consider the system* $Ax = Bx$.

It is easily checked that the unique solution (after scaling and making the smallest component equal to zero) is $x^T = (1, 0, 2)$. *Note also that* x *is a connected solution and so* $V(A, B) = \left\{\alpha (1, 0, 2)^T : \alpha \in \mathbb{R}\right\}$.

Now, $av_{x,A}(1) = \{1, 3\}$ *and so we reduce, say, component* a_{13}. *Reducing* a_{13} *by* 1 *is* *sufficient since then* $A' = \begin{pmatrix} 4 & 0 & 2 \\ 0 & 3 & 0 \\ 3 & 0 & 0 \end{pmatrix}, B = \begin{pmatrix} 0 & 5 & 3 \\ 1 & 0 & 1 \\ 0 & 4 & 2 \end{pmatrix}$ *and we have:*

1. $x \in V(A', B') \Rightarrow V(A', B') \neq \emptyset$.

2. $av_{x,A'}(1) = \{1\} = \{1,3\} \setminus \{3\} = av_{x,A}(1) \setminus \{3\}$.

3. *It is easily checked that* $x^T = (1,0,2)$ *is the unique solution for the system* $A'x = B'x$ *and so indeed we have* $V(A', B') = \{\alpha(1,0,2) : \alpha \in \mathbb{R}\} \subseteq V(A, B)$.

4. *It is clear that* $(\forall u \in N)\, av_{x,A'}(u) \subseteq av_{x,A}(u)$ *and* $av_{x,B'}(u) \subseteq av_{x,B}(u)$.

5. *The entry* $(1,3)$ *is not* $(x, A') -$ *active.*

Now, note that $av_{x,B'}(1) = \{2,3\}$ *and so we reduce, say, component* b_{12}. *Reducing* b_{12} *by* 1 *is sufficient, since then* $A'' = \begin{pmatrix} 4 & 0 & 2 \\ 0 & 3 & 0 \\ 3 & 0 & 0 \end{pmatrix}, B'' = \begin{pmatrix} 0 & 4 & 3 \\ 1 & 0 & 1 \\ 0 & 4 & 2 \end{pmatrix}$ *and we have*

1. $x \in V(A'', B'') \Rightarrow V(A'', B'') \neq \emptyset$.

2. $av_{x,B''}(1) = \{3\} = \{2,3\} \setminus \{2\} = av_{x,B'}(1) \setminus \{2\}$.

3. *It is easily checked that* $x^T = (1,0,2)$ *is the unique solution for the system* $A''x = B''x$ *and so indeed we have* $V(A'', B'') = \{\alpha(1,0,2) : \alpha \in \mathbb{R}\} \subseteq V(A', B')$.

4. *It is clear that* $(\forall u \in N)\, av_{x,A''}(u) \subseteq av_{x,A'}(u)$ *and* $av_{x,B''}(u) \subseteq av_{x,B'}(u)$.

5. *The entry* $(1,2)$ *is not* $(x, B'') -$ *active.*

Now, note that $av_{x,B''}(3) = \{2,3\}$ *and so we reduce, say, component* b_{33}. *Reducing* b_{33} *by* 1 *is sufficient, since then* $A''' = \begin{pmatrix} 4 & 0 & 2 \\ 0 & 3 & 0 \\ 3 & 0 & 0 \end{pmatrix}, B''' = \begin{pmatrix} 0 & 4 & 3 \\ 1 & 0 & 1 \\ 0 & 4 & 1 \end{pmatrix}$ *and we have*

1. $x \in V(A''', B''') \Rightarrow V(A''', B''') \neq \emptyset$.

2. $av_{x,B'''}(3) = \{2\} = \{2,3\} \setminus \{3\} = av_{x,B''}(3) \setminus \{3\}$.

3. It is easily checked that $x^T = (1,0,2)$ *is the unique solution for the system* $A'''x = B'''x$ *and so indeed we have* $V(A''', B''') = \{\alpha(1,0,2) : \alpha \in \mathbb{R}\} \subseteq V(A'', B'')$.

4. It is clear that $(\forall u \in N)\, av_{x,A'''}(u) \subseteq av_{x,A''}(u)$ *and* $av_{x,B'''}(u) \subseteq av_{x,B''}(u)$.

5. The entry $(3,3)$ *is not* (x, B''') *− active.*

We see that (A''', B''') *is minimally active.*

Now that we have shown that every essential system can be reduced to a minimally active system, we are able to show that Theorem 8.6 from Section 8.3 for minimally active systems, holds also for essential systems. (Note that the stronger result of Theorem 8.16 from Section 8.3 does not hold).

Theorem 8.29. *Let* $A, B \in \mathbb{R}^{n\times n}$ *such that* (A, B) *is an essential system. Then* $V(A, B) \neq \emptyset$ *if and only if* $(\exists x \in V(A, B))\, (\forall \sigma \in ap(C))\, \sigma$ *is* $x -$ *optimal.*

Proof. If (A, B) is minimally active then the result follows immediately from Theorem 8.6. So suppose (A, B) is not minimally active. We have seen in Lemma 8.18 and Theorem 8.27 that there is a sequence of systems $(A, B), \left(A^{(1)}, B^{(1)}\right), \dots, \left(A^{(k)}, B^{(k)}\right)$, such that

$$\emptyset \neq V\left(A^{(k)}, B^{(k)}\right) \subseteq V\left(A^{(k-1)}, B^{(k-1)}\right) \subseteq \cdots \subseteq V\left(A^{(1)}, B^{(1)}\right) \subseteq V(A, B),$$

and $\left(A^{(k)}, B^{(k)}\right)$ is minimally active, for some $k \in \mathbb{N}$. Define $C^{(r)} = A^{(r)} \oplus B^{(r)}$ and let $x \in V\left(A^{(k)}, B^{(k)}\right)$. We have seen in Lemma 8.15 that $\left(\forall \sigma \in ap\left(C^{(k)}\right)\right)\sigma$ is $x -$ *optimal* for the system $\left(A^{(k)}, B^{(k)}\right)$. That is to say, there is a perfect matching M in

$G_x\left(A^{(k)}, B^{(k)}\right)$. In moving from $\left(A^{(k)}, B^{(k)}\right)$ to $\left(A^{(k-1)}, B^{(k-1)}\right)$ we are not losing any edges in $G_x\left(A^{(k)}, B^{(k)}\right)$. That is to say $E\left(G_x\left(A^{(k)}, B^{(k)}\right)\right) \subseteq E\left(G_x\left(A^{(k-1)}, B^{(k-1)}\right)\right)$. (To see this, recall by construction of $\left(A^{(k)}, B^{(k)}\right)$ that we reduced exactly one element in the system $\left(A^{(k-1)}, B^{(k-1)}\right)$). It follows that M is a perfect matching in $G_x\left(A^{(k-1)}, B^{(k-1)}\right)$ also. Continuing, we see that M is a perfect matching in $G_x(A, B)$, as desired.

It follows that $(\exists x \in V(A, B))(\forall \sigma \in ap(C))\, \sigma$ is $x - optimal$.

□

8.5 Next steps

The ideas in this section allow us to take a square, finite and essential system (A, B) and find, for each equation, an active entry (in A without loss of generality). It is known that when the system (A, B) is regular (that is to say $(\forall i)(\forall j)\, a_{ij} \neq b_{ij}$), then if (A, B) is solvable, then $|ap(C)| \geq 2$ and so it follows that for some equations we know multiple active entries (not necessarily any in B though). We have obtained important information about a solution but we have not completed the job of finding a solution to the two-sided system.

8.5.1 Minimally active systems

Let (A, B) be minimally active. Then, without loss of generality, by finding $ap(C)$, we identify active entries in A and B. Note that $maper(C) = 0$ without loss of generality. In the following, blocks with $\tilde{0}$ entries denote a block containing a zero cycle.

$$A = \begin{pmatrix} 0 & & & & & & \\ & 0 & & & & & \\ & & \ddots & & & & \\ & & & \ddots & & & \\ & & & & \ddots & & \\ & & & & & 0 & \\ & & & & & & 0 \end{pmatrix}, B = \left(\begin{array}{c:c:cc|c} 0 & & & & \\ \hdashline & 0 & & & \\ \hdashline & & \tilde{0} & \tilde{0} & \\ & & \tilde{0} & \tilde{0} & \\ \hline & & & & \\ & & & & \end{array}\right).$$

Due to the minimal activity property, there will be no other active entries in A and the zero cycles/loops identified in B will be non-intersecting - yielding the block-diagonal form shown above.

Now, denote by K the set of indices which do not lie on a zero cycle in B. For $k \in K$, we have the substitution

$$x_k = \bigoplus_{t \in N, t \neq k} c_{kt} x_t.$$

With these substitutions, we can reduce to a necessary system of size $(n - |K|) \times (n - |K|)$ which is essentially a system of dual inequalities (since for each equation we have identified exactly one active entry in each of A, B). In fact, the set of solutions to the necessary system of dual inequalities is of the form Gu, where $G \in \mathbb{R}^{r \times r}$, r is the number of zero cycles in B. By backtracking our substitutions, we may convert our original system to a an r-dimensional two-sided system, with $|K|$ equations - that is, a TSLS of size $|K| \times r$.

8.5.2 Other cases

One important special case to consider is when $maper\,(A) = maper\,(B) = maper\,(C)$, since in this case our methods will identify at least one active element in both A and B for every equation - thus making the two-sided system essentially a system of dual inequalities (solvable in polynomial time). In general though, this will not be the case. One interesting possibility is to notice the necessary condition

$$(\forall i \in R)\, a_{ii}x_i = \bigoplus_{t \neq i} c_{it}x_t,$$

where $R := \{r \in N : a_{rr} \neq b_{rr}\}$. Such a subsitution allows us to eliminate variables until we are left with a system of dual inequalities (importantly with strictly less variables than the original system). We can solve such systems in polynomial time but this set of solutions to the reduced system of dual inequalities would only lead to a superset of the set of solutions to the original system.

Remark 8.30. *The ideas of this section work also when instead of finite systems (A, B), we take A and B over $\overline{\mathbb{R}}$, with the condition that maper (C) is finite.*

8.6 Open questions

Question 1 How can we recognise if a square system is essential?

Question 2 How can we recognise if a square system is minimally active?

Question 2 Can the ideas of this section be adapted for the case of non-square systems?

Question 4 What can we say about two-sided systems for which there is no finite permutation in C?

8.7 Summary

We have shown that under reasonable assumptions (the matrices A and B are square and finite and every variable is finite in a non-trivial solution to (8.4)), two-sided systems (8.4) have a strong connection with the assignment problem. Further, if a solution exists, then by solving the assignment problem we identify active elements of a solution. The drawbacks, of course, are that it is not clear how to identify such systems (though an easily identifiable subset of such systems has been discussed here) and that after identifying active elements, it is not always clear how we can ‘fill in the gaps” to identify the rest of the active elements.

9. The generalised eigenproblem - strongly polynomial algorithm if matrices are circulant

9.1 Problem formulation

We begin with a definition.

Definition 9.1 (Circulant Matrix)**.** *A Circulant matrix, denoted C, is a matrix in which each column is a circular shift of its preceding column.*

In this chapter A and B are finite, circulant matrices. We are interested in the generalised eigenproblem for such matrices

$$Ax = \lambda Bx. \tag{9.1}$$

9.2 The strongly polynomial "aggregation method"

By using the method of aggregation we deduce some necessary conditions on the value of the (unique) generalised eigenvalue. Specifically, if the spectrum is non-empty, then the unique eigenvalue is the value of the greatest entry in A, less the value of the greatest entry in B. The result is that the generalised eigenproblem for circulant matrices is easily *converted to the problem of solving two-sided systems for circulant matrices.*

$$T = \begin{pmatrix} t_0 & t_{n-1} & \dots & t_2 & t_1 \\ t_1 & t_0 & t_{n-1} & & t_2 \\ \vdots & t_1 & t_0 & \ddots & \vdots \\ t_{n-2} & & \ddots & \ddots & t_{n-1} \\ t_{n-1} & t_{n-2} & \dots & t_1 & t_0 \end{pmatrix}. \tag{9.2}$$

Let $A, B \in \mathbb{R}^{n \times n}$ be circulant matrices. Then $C = A - B$ is circulant too. Consider the generalised eigenproblem (9.1) and the feasible interval $[L, U]$, where $L := \max_i \min_j c_{ij}$ and $U := \min_i \max_j c_{ij}$. Since C is circulant, $(\forall i) \min_j c_{ij} = \min_k t_k$. Similarly, $(\forall i) \max_j c_{ij} = \max_k t_k$ and so we have

$$[L, U] = [\min_k t_k, \max_k t_k]. \tag{9.3}$$

Clearly the feasible interval is non-empty.

Suppose that $(\lambda; x)$ is a solution for some $\lambda \in [L, U]$ and $x \in \overline{\mathbb{R}}^n, x \neq \epsilon$. Then we have

$$\begin{cases} a_0x_1 \oplus a_{n-1}x_2 \oplus \dots \oplus a_1x_n = \lambda b_0x_1 \oplus \lambda b_{n-1}x_2 \oplus \dots \oplus \lambda b_1x_n \\ a_1x_1 \oplus a_0x_2 \oplus \dots \oplus a_2x_n = \lambda b_1x_1 \oplus \lambda b_0x_2 \oplus \dots \oplus \lambda b_2x_n \\ \vdots \\ a_{n-1}x_1 \oplus a_{n-2}x_2 \oplus \dots \oplus a_0x_n = \lambda b_{n-1}x_1 \oplus \lambda b_{n-2}x_2 \oplus \dots \oplus \lambda b_0x_n. \end{cases}$$

By aggregation using max-plus we then have that

$$\begin{aligned}
&\bigoplus_{i=0}^{n-1}\left(\bigoplus_{j=1}^{n}(a_i x_j)\right) = \bigoplus_{i=0}^{n-1}\left(\bigoplus_{j=1}^{n}(\lambda b_i x_j)\right)\\
\Rightarrow &\bigoplus_{i=0}^{n-1}\left(a_i \bigoplus_{j=1}^{n} x_j\right) = \lambda \bigoplus_{i=0}^{n-1}\left(b_i \bigoplus_{j=1}^{n} x_j\right)\\
\Rightarrow &\bigoplus_{j=1}^{n} x_j \bigoplus_{i=0}^{n-1} a_i = \lambda \bigoplus_{j=1}^{n} x_j \bigoplus_{i=0}^{n-1} b_i\\
\Rightarrow &\lambda \bigoplus_{i=0}^{n-1} b_i = \bigoplus_{i=0}^{n-1} a_i\\
\Rightarrow &\lambda = \bigoplus a_i \left(\bigoplus b_i\right)^{-1}.
\end{aligned}$$

In fact, it is not too hard to show that this choice of λ with the vector $x = 0$ is always a solution for circulant matrices! A less than obvious corollary of this is that this choice of λ must always lie within the feasible interval.

Corollary 9.2. *Let A, B be circulant matrices. Then*

$$\bigoplus' \left(a_i b_i^{-1}\right) \leq \bigoplus a_i \left(\bigoplus b_i\right)^{-1} \leq \bigoplus \left(a_i b_i^{-1}\right).$$

9.3 Summary

By a simple application of the method of aggregation, we showed that the unique eigenvalue can be described as the greatest entry in A, less the greatest entry in B, if A and B are circulant matrices. This also allowed us to deduce an inequality for circulant matrices which is not obvious at first sight.

10. The generalized eigenproblem - strongly polynomial algorithm if B is an outer-product

10.1 Introduction

In this chapter, B (say) is the max-algebraic outer-product of two vectors and we consider the generalised eigenproblem

$$Ax = \lambda Bx. \tag{10.1}$$

We show that without loss of generality we can assume B is the all zeros matrix.

By using the one-sided parametrised systems from chapter 3, we convert (10.1) to an equivalent one-sided parametrised system. The one-sided systems appearing here are slightly different to those in chapter 3 but are easily converted. It follows that we can describe all solutions to (10.1) in strongly polynomial time.

The contents of this chapter have been published in [17]

10.2 Converting to a one-sided parametrised system

Let us consider the generalized eigenproblem (10.1) where B is an outer product of two vectors, say $v = (v_1, ..., v_m)^T \in \mathbb{R}^m$ and $w = (w_1, ..., w_n)^T \in \mathbb{R}^n$. Thus we can write

$$B = (b_{ij}) = vw^T = (v_i + w_j) .$$

Let $V = diag\,(v)$ and $W = diag\,(w)$. Then (10.1) reads

$$Ax = \lambda v w^T x$$

and is equivalent to

$$V^{-1}Ax = \lambda V^{-1}Bx,$$

or,

$$V^{-1}Ax = \lambda \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} w^T x.$$

Set $x = W^{-1}y$, where $y = (y_1, ..., y_n)^T$ and we obtain

$$V^{-1}AW^{-1}y = \lambda \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} (y_1 \oplus y_2 \oplus ... \oplus y_n) .$$

Here the right-hand side is actually equal to

$$\lambda \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{pmatrix} y$$

and so in the system (10.1) with B being an outer product, it can be assumed without loss of generality that B is a zero matrix (a matrix with all entries equal to 0). In such a system the right-hand side of each equation is $\lambda\left(x_1 \oplus x_2 \oplus ... \oplus x_n\right)$ and so an $x \neq \varepsilon$ satisfying $Ax = \lambda 0 x$ exists if and only if there is a $z \in \mathbb{R}$ satisfying

$$A'x = \begin{pmatrix} z \\ \vdots \\ z \end{pmatrix}, \tag{10.2}$$

where A' is obtained from A by adding an extra row whose every entry is λ, that is

$$A' = \begin{pmatrix} A \\ \lambda \; ... \; \lambda \end{pmatrix} = \left(a'_{ij}\right).$$

Clearly, A' is an $(m+1) \times n$ matrix. Theorem 2.3 and Proposition 2.5 enable us to solve such systems and we use them in the next two propositions. It will be useful to denote

$$\begin{aligned} M' &= M \cup \{m+1\}, \\ M'_j &= \left\{ r \in M'; a'_{rj} = \max_{i \in M'} a'_{ij} \right\}, \quad j \in N \end{aligned}$$

and

$$N'_i = \left\{ j \in N; i \in M'_j \right\}, \quad i \in M'.$$

Proposition 10.1. *Let $A \in \mathbb{R}^{m \times n}$. Then $\lambda \in \Lambda(A, 0)$ if and only if $\bigcup_{j \in N} M'_j = M'$.*

Proof. Follows straightforwardly from the previous discussion, Theorem 2.3 and Proposition 2.5. □

Proposition 10.2. $\Lambda(A,0) \subseteq [\lambda_0, \lambda_1]$ *holds for every* $A = (a_{ij}) \in \mathbb{R}^{m \times n}$, *where*

$$\lambda_0 = \min_{j \in N} \max_{i \in M} a_{ij} \tag{10.3}$$

and

$$\lambda_1 = \max_{j \in N} \max_{i \in M} a_{ij}. \tag{10.4}$$

Proof. A real number λ is in $\Lambda(A,0)$ if and only if the system (10.2) has a solution for some $z \in \mathbb{R}$. This is a one-sided system whose solvability does not depend on z because the right-hand side is a constant vector. The solvability criterion is given in Theorem 2.3. We apply this condition to A' using Proposition 2.5, separately to the first m rows and to the last row:

$$(\forall i \in M)\,(\exists j \in N)\,(\forall r \in M)\, a_{ij} \geq a_{rj} \oplus \lambda \tag{10.5}$$

and

$$(\exists j \in N)\,(\forall i \in M)\, \lambda \geq a_{ij}. \tag{10.6}$$

The latter is equivalent to $\lambda \geq \min_{j \in N} \max_{i \in M} a_{ij}$, which proves the lower bound. The first is equivalent to the requirement that for every $i \in M$ there is a $j \in N$ satisfying

$$a_{ij} \geq \max_{r \in M} a_{rj} \oplus \lambda.$$

Since

$$\max_{r \in M} a_{rj} \oplus \lambda \geq \max_{r \in M} a_{rj} \geq a_{ij}$$

it follows that

$$a_{ij} = \max_{r \in M} a_{rj} \oplus \lambda = \max_{r \in M} a_{rj}$$

and $\lambda \leq \max_{r \in M} a_{rj} \leq \max_{j \in N} \max_{i \in M} a_{ij}$, which proves the upper bound. □

Proposition 10.2 does not provide any tool for checking whether $\Lambda(A,0)$ is non-empty. We give this answer next.

Proposition 10.3. *The following statements are equivalent for every* $A = (a_{ij}) \in \mathbb{R}^{m\times n}$*:*

(a) $\Lambda(A,0) \neq \emptyset$,

(b) $\bigcup_{j\in N} M_j = M$,

(c) $N_i(A) \neq \emptyset$ *for every* $i \in M$ *and*

(d) $\lambda_0 \in \Lambda(A,0)$.

Proof. The equivalence of (b) and (c) has been shown in Proposition 2.1. We prove (b) $\Rightarrow$ (d) $\Rightarrow$ (a) $\Rightarrow$ (b). Suppose (b) is true. There is an index $k \in N$ such that

$$\lambda_0 = \max_{i\in M} a_{ik} \leq \max_{i\in M} a_{ij}$$

holds for every $j \in N$. Hence for $\lambda = \lambda_0$ and for every $j \in N$ we have $M_j \subseteq M_j'$ and $m+1 \in M_k'$. It follows that

$$\bigcup_{j\in N} M_j' \supseteq \bigcup_{j\in N} M_j \cup \{m+1\} = M \cup \{m+1\} = M'$$

thus

$$\bigcup_{j\in N} M_j' = M'$$

and so (10.2) has a solution for $\lambda = \lambda_0$, which proves (d) by Proposition 10.1.

The second implication is trivial, so suppose now that $\lambda \in \Lambda(A,0)$. Hence (10.2) has a solution with this value of λ and thus $\bigcup_{j\in N} M_j' = M'$. Let $i \in M$ then $i \in M_j'$ for some $j \in N$ and therefore also $i \in M_j$ because for any $j \in N$ the set M_j' either coincides with M_j or is $M_j \cup \{m+1\}$ or is just $\{m+1\}$. Statement (b) now follows. $\square$

Corollary 10.4. *If $A \in \mathbb{R}^{n\times n}$ is symmetric then $\Lambda(A,0)$ is either empty or $\{\lambda_0\}$.*

Proposition 10.3 shows that the lower bound for the spectrum specified in Proposition 10.2 is tight for every matrix with non-empty spectrum. In general this is not true about the upper bound. For instance if

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}, B = 0$$

then $C = A$ is symmetric and so $\Lambda(A,B) = \{1\}$ by Proposition 10.3 and Corollary 10.4. However, $\lambda_1 = 2 > 1$.

In order to give exact bounds we now introduce the following:

$$\underline{\lambda} = \min_{i\in M} \min_{j\in N_i} a_{ij}$$

and

$$\overline{\lambda} = \min_{i\in M} \max_{j\in N_i} a_{ij}.$$

Proposition 10.5. $\underline{\lambda} = \lambda_0$ *for any* $A \in \mathbb{R}^{m\times n}$.

Proof. There exist $r \in M$ and $s \in N$ such that

$$\lambda_0 = \min_{j\in N} \max_{i\in M} a_{ij} = \max_{i\in M} a_{is} = a_{rs}.$$

So a_{rs} is the smallest column maximum in A. Let $k \in M$. The quantity $\min_{l\in N_k} a_{kl}$ is the smallest of all column maxima appearing in row k of A (recall that this value is $+\infty$ if $N_k = \emptyset$). Hence $\min_{l\in N_k} a_{kl} \geq a_{rs}$ and therefore also $\underline{\lambda} = \min_{k\in M} \min_{l\in N_k} a_{kl} \geq a_{rs} = \lambda_0$.

On the other hand $\underline{\lambda} = \min_{k\in M} \min_{l\in N_k} a_{kl} \leq \min_{l\in N_r} a_{rl} = a_{rs} = \lambda_0$. □

Proposition 10.6. $\Lambda(A,0) \subseteq [\underline{\lambda}, \overline{\lambda}]$ *holds for every* $A = (a_{ij}) \in \mathbb{R}^{m\times n}$ *and* $\{\underline{\lambda}, \overline{\lambda}\} \subseteq \Lambda(A,0)$ *whenever* $\Lambda(A,0) \neq \emptyset$.

Proof. The lower bounds in Propositions 10.2 and 10.6 coincide by Proposition 10.5 so we only need to prove the upper bound and its tightness. Suppose without loss of generality that $\Lambda(A,0) \neq \emptyset$ and so $N_i \neq \emptyset$ for every $i \in M$. Let $\lambda = \overline{\lambda} = \min_{i\in M} \max_{j\in N_i} a_{ij}$. Then $\lambda = \max_{i\in M} a_{is}$ for some $s \in N$ and thus $m+1 \in M'_s$. At the same time if $r \in M$ then $\max_{j\in N_r} a_{rj} \geq \min_{i\in M} \max_{j\in N_i} a_{ij} = \lambda$. Therefore $r \in M'_j$ for some $j \in N_r$. This shows that $\overline{\lambda} \in \Lambda(A,0)$ by Proposition 10.1.

Suppose now that $\lambda > \overline{\lambda}$. Then $\lambda > \max_{j\in N_i} a_{ij}$ for some $i \in M$. Hence $\lambda > a_{ij}$ for all $j \in N_i$ and so $i \notin M'_j$ for all $j \in N_i$. Since also $i \notin M_j$ for all $j \notin N_i$ and $M_j \subseteq M'_j$ for every $j \in N_i$ we have that $i \notin M_j$ for all $j \in N$, thus $\lambda \notin \Lambda(A,0)$ by Proposition 10.3. □

Theorem 10.7. $\Lambda(A,0) = [\underline{\lambda}, \overline{\lambda}]$ *holds for every* $A = (a_{ij}) \in \mathbb{R}^{m\times n}$.

Proof. If $\Lambda(A,0) = \emptyset$ then by Proposition 10.3 $N_i = \emptyset$ for some $i \in M$ and so $\overline{\lambda} = -\infty$, $\underline{\lambda} = +\infty$ and $[\underline{\lambda}, \overline{\lambda}] = \emptyset$.

Suppose now $\Lambda(A,0) \neq \emptyset$. Due to Proposition 10.6 we may also assume that $\underline{\lambda} < \overline{\lambda}$ and we only need to prove that $(\underline{\lambda}, \overline{\lambda}) \subseteq \Lambda(A,0)$. Let $\lambda \in (\underline{\lambda}, \overline{\lambda})$. If $i \in M$, then $\lambda < \overline{\lambda} \leq \max_{j\in N_i} a_{ij} = a_{it}$ for some $t \in N_i$. Hence $i \in M_t = M'_t$ and so $i \in \bigcup_{j\in N} M'_j$. On the other hand $\lambda > \underline{\lambda} \geq a_{rs}$, where $r \in M, s \in N_r$, hence $\lambda = \max_{i\in M'} a'_{is}$, thus $m+1 \in M'_s$. We conclude that $\bigcup_{j\in N} M'_j = M'$ and so $\lambda \in \Lambda(A,0)$ by Proposition 10.1.

□

10.3 Summary

The results of this chapter show a connection between two-sided systems and the set-covering problem if B (say) is an outer-product. In this case, GEP can be solved in

$O\left(\max\left(m,n\right)^3\right)$ time (the whole spectrum can be found). To see the complexity, we assume we know the vectors v and w such that $B = vw^T$, then it takes $O\left(\max\left(m,n\right)^3\right)$ time to calculate $V^{-1}AW^{-1}$. Once the conversion has happened (B is the zero matrix), we calculate $\underline{\lambda}$ and $\overline{\lambda}$ in $O\left(mn\right)$ time. In total, the complexity is $O\left(\max\left(m,n\right)^3\right) + O\left(mn\right) = O\left(\max\left(m,n\right)^3\right)$ time.

11. The generalized eigenproblem - unique candidate if difference of matrices is symmetric and has a saddle point

11.1 Introduction

In this chapter the matrices A and B are square and finite and the matrix $C := A - B$ is symmetric with a saddle point.

We show that the value of the saddle point in C is the unique candidate for a generalised eigenvalue. We give necessary conditions for the saddle point to be an eigenvalue in the general $n \times n$ case.

The results on the saddle point being the unique candidate for the eigenvalue can be found in [17].

11.2 Saddle point and unique generalised eigenvalue

Definition 11.1. *The position (r, s) is a saddle point of a matrix C if c_{rs} is both a column maximum and row minimum.*

Note that C is a symmetric matrix if we have the stronger assumption that both A and B are symmetric matrices. The following [12] is given as a reminder.

Proposition 11.2. *If $A, B \in \mathbb{R}^{n\times n}$ are symmetric matrices, then $|\Lambda(A,B)| \leq 1$.*

Note that in general even if the premise of Proposition 11.2 is satisfied, it is not clear what the unique candidate for the generalized eigenvalue is and even if such a candidate is known it is not clear how to check in polynomial time whether it is such an eigenvalue.

Recall from chapter 2 that given a matrix C,

$$L(C) = \max_{i\in M} \min_{j\in N} c_{ij}$$

and

$$U(C) = \min_{i\in M} \max_{j\in N} c_{ij}.$$

In game theory, the following quantities are known as the *gain-floor* and *loss-ceiling* functions respectively:

$$v_1(C) = \max_{i\in M} \min_{j\in N} c_{ij},$$

$$v_2(C) = \min_{j\in N} \max_{i\in M} c_{ij}.$$

It is known that C has a saddle point if and only if $v_1(C) = v_2(C)$.

Let $C \in \mathbb{R}^{m\times n}$ be any matrix. Clearly, $L(C)$ coincides with $v_1(C)$. Although in general $U(C)$ is different from $v_2(C)$, we observe that for $C^T = (c_{ji}) = \left(c'_{ij}\right)$ we have

$$\begin{aligned} U(C) &= \min_{i\in M} \max_{j\in N} c_{ij} \\ &= \min_{j\in M} \max_{i\in N} c_{ji} \\ &= \min_{j\in M} \max_{i\in N} c'_{ij} \\ &= v_2\left(C^T\right). \end{aligned}$$

Hence $L(C) = v_1(C)$ and $U(C) = v_2(C)$ if C is symmetric, in particular when $C = A - B$ and both A and B are symmetric. Thus if C is symmetric and has a saddle point, say (r, s), then $L(C) = a_{rs} = L(C)$ and so either a_{rs} is the unique generalized eigenvalue or there is no such eigenvalue, see Proposition 2.6. The next two examples confirm that both cases are possible.

Example 11.3. *If*

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix}, B = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

then $C = A$ is symmetric and has saddle point $(1, 2)$ of value 1, which is therefore the unique candidate for a generalized eigenvalue. If there is an associated eigenvector $x = (x_1, x_2)^T$, we may assume without loss of generality $x_1 = 0$ and the individual terms in Ax and λBx are as described in the following matrices:

$$\begin{pmatrix} 2 & 1 + x_2 \\ 1 & x_2 \end{pmatrix}, \begin{pmatrix} 1 & 1 + x_2 \\ 1 & 1 + x_2 \end{pmatrix}.$$

The first equation implies $1 + x_2 \geq 2$ and the second $1 + x_2 \leq 1$, thus $\Lambda(A, B) = \emptyset$.

Example 11.4. *If*

$$A = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, B = \begin{pmatrix} -2 & -1 \\ 0 & 0 \end{pmatrix}$$

then C is the same as in the previous example and so $\lambda = 1$ is the unique candidate for a generalized eigenvalue. Now the zero vector $x = (0, 0)^T$ is an associated eigenvector and thus $\Lambda(A, B) = \{1\}$.

The following example shows that $\Lambda(A, B)$ may be non-empty even if C is symmetric and has no saddle point.

Example 11.5. *If*

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix}, B = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}$$

then

$$C = \begin{pmatrix} 0 & 2 \\ 2 & 1 \end{pmatrix}$$

is a symmetric matrix without a saddle point but $\lambda = 2$ is a generalized eigenvalue with associated eigenvector $x = (0,1)^T$.

We summarize:

Theorem 11.6. *If $A, B \in \mathbb{R}^{n \times n}, C = A - B$ is a symmetric matrix with a saddle point (r, s) and $\Lambda(A, B) \neq \emptyset$ then c_{rs} is the unique generalized eigenvalue for (A, B).*

The author is not aware of any polynomial method for checking that the saddle point value of $A - B$ is a generalized eigenvalue of (A, B). In the next two sections we address conditions under which a saddle point is actually an eigenvalue.

11.3 Necessary condition for the saddle point to be an eigenvalue for square matrices

Definition 11.7. *Let $C \in \mathbb{R}^{n \times n}$. We say c_{kl} is a strict saddle point of C if $(\forall i)\, c_{kl} > c_{il}$ and $(\forall j)\, c_{kl} < c_{kj}$.*

For the remainder of this chapter we use the notation A_{kl} to denote the 2×2 submatrix of A obtained by taking the rows and columns k and l of A. Recall also (2.1), page 15.

Theorem 11.8. *Let $C = A - B \in \mathbb{R}^{n \times n}$ where C is symmetric and has a strict saddle point (k, l) for some $k \neq l$. Then $\Lambda(A, B) = \{c_{kl}\} \Rightarrow d(A_{kl}) \leq c_{ll} - c_{kl}$. That is, the above* *is a necessary condition for the unique candidate c_{kl} to be a solution to the max-algebraic generalised eigenproblem $Ax = \lambda Bx$.*

Proof. Let $A = (a_{ij}) \in \mathbb{R}^{n\times n}$ and $C = (c_{ij}) \in \mathbb{S}^{n\times n}$, that is $c_{ij} = c_{ji}$ for all $i, j \in N$. Let $B = (b_{ij}) \in \mathbb{R}^{n\times n}$ satisfy $A - B = C$, that is $b_{ij} = a_{ij} - c_{ij}$ for all $i, j \in N$. Let (k, l), $k \neq l$ be a strict saddle point of C. Now c_{kl} is a strict column maxima and so

$$(\forall i \neq k)\, c_{kl} - c_{il} > 0. \tag{11.1}$$

Similarly, since c_{kl} is a strict row minima we have

$$(\forall j \neq l)\, c_{kl} - c_{kj} < 0. \tag{11.2}$$

We wish to solve the system $Ax = c_{kl}Bx$, that is, $Ax = B'x$ where $B' = \left(b'_{ij}\right) = (a_{ij} + c_{kl} - c_{ij})$. In what follows $x_k = 0$ without loss of generality. Consider equation k of this system:

$$\begin{aligned} &\max\left(a_{k1} + x_1, \ldots, a_{kl} + x_l, \ldots, a_{kn} + x_n\right) \\ =&\max\left(a_{k1} + c_{kl} - c_{k1} + x_1, \ldots, a_{kl} + c_{kl} - c_{kl} + x_l, \ldots, a_{kn} + c_{kl} - c_{kn} + x_n\right). \end{aligned}$$

By inequality (11.2) and the Cancellation Rule, the right hand side becomes $a_{kl} + x_l$. In particular, then, we must have

$$a_{kl} + x_l \geq a_{kk} + x_k = a_{kk}. \tag{11.3}$$

Similarly, if we consider equation l of the system, we obtain

$$\begin{aligned}&\max\left(a_{l1}+x_1,\ldots,a_{lk}+x_k,\ldots,a_{ln}+x_n\right)\\=&\max\left(a_{l1}+c_{kl}-c_{l1}+x_1,\ldots,a_{lk}+c_{kl}-c_{lk}+x_k,\ldots,a_{ln}+c_{kl}-c_{ln}+x_n\right).\end{aligned}$$

Equivalently, using the symmetry of C, we have

$$\begin{aligned}&\max\left(a_{l1}+x_1,\ldots,a_{lk}+x_k,\ldots,a_{ln}+x_n\right)\\=&\max\left(a_{l1}+c_{kl}-c_{1l}+x_1,\ldots,a_{lk}+c_{kl}-c_{kl}+x_k,\ldots,a_{ln}+c_{kl}-c_{nl}+x_n\right).\end{aligned}$$

By inequality (11.1) and the Cancellation Rule, the left hand side becomes $a_{lk}+x_k=a_{lk}$. In particular, then, we must have

$$a_{lk}\geq a_{ll}+c_{kl}-c_{ll}+x_l. \tag{11.4}$$

From (11.3) and (11.4), we have the system

$$\begin{cases}x_l\geq a_{kk}-a_{kl}\\ x_l\leq a_{lk}-a_{ll}+c_{ll}-c_{kl}.\end{cases} \tag{11.5}$$

The system (11.5) is solvable if and only if

$$a_{kk}-a_{kl}\leq a_{lk}-a_{ll}+c_{ll}-c_{kl},$$

which holds if and only if

$$a_{kk}+a_{ll}-a_{kl}-a_{lk}\leq c_{ll}-c_{kl}.$$

Equivalently, (11.5) is solvable if and only if

$$d\left(A\left[\{k,l\}:\{k,l\}\right]\right) \leq c_{ll} - c_{kl}.$$

□

11.4 Summary

By making use of the connections with Game Theory, we have been able to identify the unique candidate for the eigenvalue when $C = A - B$ is a symmetric matrix with a saddle point and we were able to give necessary conditions for the unique candidate to be an eigenvalue. In general, it is not clear whether or not the candidate is indeed an eigenvalue. It would be interesting to study the resulting special class of two-sided systems of equations in order to answer this question (see chapter 13).

12. Matrix roots

12.1 Introduction

We define positive integer roots of finite matrices in the 2×2 case. Where possible, we try to generalise results to finite $n \times n$ matrices. Note that it was shown in [57] that finding kth roots of Boolean matrices (equivalently kth roots of digraphs) for any fixed positive integer $k \geq 2$ is NP-complete.

12.2 Finite 2×2 matrices

Let $A = (a_{ij}) \in \mathbb{R}^{2\times 2}$. We define the *max discriminant of* A to be the quantity

$$d\,(A) = a_{11}a_{22}a_{12}^{-1}a_{21}^{-1},$$

(see (2.1), page 15). Similarly for $d\,(B)$. Also, define the *max trace of* A to be the quantity

$$tr\,(A) := a_{11} \oplus a_{22}.$$

Similarly for $tr\,(B)$. We use the shorthand $a = tr\,(A)$ and $b = tr\,(B)$. Note that $\sqrt{a}$ for $a \in \mathbb{R}$ should be taken to mean $\frac{a}{2}$ in the classical linear notation, while A for a matrix A should be taken in the usual sense of matrix roots.

Theorem 12.1. *Let $A \in \mathbb{R}^{2\times 2}$. Then there exists $B \in \mathbb{R}^{2\times 2}$ such that $B^2 = A$ if and only* *if $d(A) \geq 0$. Moreover, if $d(A) \geq 0$, the matrix B^* defined by*

$$B^* = \begin{pmatrix} \sqrt{a_{11}} & a_{12}\sqrt{a}^{-1} \\ a_{21}\sqrt{a}^{-1} & \sqrt{a_{22}} \end{pmatrix}$$

is such that $(B^*)^2 = A$.

Proof. First, we suppose $d(A) < 0$ and suppose, for a contradiction, that there exists $B = (b_{ij}) \in \mathbb{R}^{2\times 2}$ such that $B^2 = A$. Then

$$\begin{pmatrix} b_{11}^2 \oplus b_{12}b_{21} & b_{12}b \\ b_{21}b & b_{22}^2 \oplus b_{12}b_{21} \end{pmatrix} = A. \tag{12.1}$$

Case 1 Suppose $a_{11} < a_{22}$, so that $a = a_{22}$. Then

$$b_{11}^2 \oplus b_{12}b_{21} < b_{22}^2 \oplus b_{12}b_{21} = b_{22}^2 \tag{12.2}$$

since the inequality yields

$$b_{12}b_{21} < b_{22}^2 \tag{12.3}$$

by cancellation. In particular, by (12.2), $b_{11}^2 < b_{22}^2$, which gives

$$b_{11} < b_{22}. \tag{12.4}$$

It follows from (12.1), (12.3) and (12.4) that

$$A = \begin{pmatrix} b_{11}^2 \oplus b_{12}b_{21} & b_{12}b_{22} \\ b_{21}b_{22} & b_{22}^2 \end{pmatrix}. \tag{12.5}$$

Clearly then,

$$b_{22} = \sqrt{a_{22}}. \tag{12.6}$$

By (12.5) and (12.6) we have then

$$\begin{cases} b_{12} = a_{12}\left(\sqrt{a_{22}}\right)^{-1} \\ b_{21} = a_{21}\left(\sqrt{a_{22}}\right)^{-1}. \end{cases} \tag{12.7}$$

By (12.5) and (12.7) it follows that

$$a_{11} = b_{11}^2 \oplus a_{12}a_{21}a_{22}^{-1}$$

implying that

$$a_{11} \geq a_{12}a_{21}a_{22}^{-1},$$

contradicting $d(A) < 0$.

Case 2 The case where $a_{11} > a_{22}$ is similar and omitted here.

Case 3 Suppose $a_{11} = a_{22} = a$. Then

$$b_{11}^2 \oplus b_{12}b_{21} = b_{22}^2 \oplus b_{12}b_{21}. \tag{12.8}$$

It follows by (12.8) that

$$b_{12}b_{21} \geq b_{11}^2 \Leftrightarrow b_{12}b_{21} \geq b_{22}^2. \tag{12.9}$$

Assume both inequalties in (12.9) hold, then by (12.1)

$$d(A) = (b_{12}b_{21})^2 b_{12}^{-1}b_{21}^{-1}\left[(b)^2\right]^{-1} \tag{12.10}$$

$$= b_{12}b_{21}\left[(b)^2\right]^{-1}. \tag{12.11}$$

It follows by (12.9) that $b_{12}b_{21} \geq (b)^2$, which gives

$$\left[(b)^2\right]^{-1} \geq b_{12}^{-1}b_{21}^{-1}.$$

Into (12.10), this yields

$$\begin{aligned} 0 &> b_{12}b_{21}\left[(b)^2\right]^{-1} \\ &\geq b_{12}b_{21}b_{12}^{-1}b_{21}^{-1} \\ &= 0, \end{aligned}$$

a contradiction. It follows that neither inequality in (12.9) holds. That is, $b_{11}^2 = b_{22}^2 = b^2 > b_{12}b_{21}$ and

$$B^2 = \begin{pmatrix} b_{11}^2 & b_{12}b \\ b_{21}b & b_{22}^2 \end{pmatrix}.$$

Then we have

$$\begin{cases} b_{11} = b_{22} = \sqrt{a} \\ b_{12} = a_{12}\left(\sqrt{a}\right)^{-1} \\ b_{21} = a_{21}\left(\sqrt{a}\right)^{-1}. \end{cases} \tag{12.12}$$

Define $C = B^2$. Then, by (12.12)

$$C = \begin{pmatrix} a \oplus a_{12}a_{21}a^{-1} & a_{12}\left(\sqrt{a}\right)^{-1}\sqrt{a} \\ a_{21}\left(\sqrt{a}\right)^{-1}\sqrt{a} & a \oplus a_{12}a_{21}a^{-1} \end{pmatrix} \tag{12.13}$$

$$= \begin{pmatrix} a_{11} \oplus a_{12}a_{21}\left(a\right)^{-1} & a_{12} \\ a_{21} & a_{22} \oplus a_{12}a_{21}\left(a\right)^{-1} \end{pmatrix}. \tag{12.14}$$

Clearly, $C = A$ if and only if

$$a_{12}a_{21}\left(a\right)^{-1} \leq a$$

$$\Leftrightarrow d\left(A\right) \geq 0,$$

a contradiction.

Having considered all cases, we conclude that there does not exist a matrix $B \in \mathbb{R}^{2\times 2}$ such that $B^2 = A$.

Conversely now, suppose $d\left(A\right) \geq 0$. It suffices to show that the matrix B^* given by

$$B^* = \begin{pmatrix} \sqrt{a_{11}} & a_{12}\left(\sqrt{a}\right)^{-1} \\ a_{21}\left(\sqrt{a}\right)^{-1} & \sqrt{a_{22}} \end{pmatrix} \tag{12.15}$$

is such that $\left(B^*\right)^2 = A$.

We can see that

$$(B^*)^2 = \begin{pmatrix} a_{11} \oplus a_{12}a_{21}a^{-1} & a_{12}\left(\sqrt{a}\right)^{-1}\left(\sqrt{a}\right) \\ a_{21}\left(\sqrt{a}\right)^{-1}\left(\sqrt{a}\right) & a_{22} \oplus a_{12}a_{21}a^{-1} \end{pmatrix} \tag{12.16}$$

$$= \begin{pmatrix} a_{11} \oplus a_{12}a_{21}\left(a\right)^{-1} & a_{12} \\ a_{21} & a_{22} \oplus a_{12}a_{21}\left(a\right)^{-1} \end{pmatrix}. \tag{12.17}$$

Now, since $d(A) \geq 0$, we have $a_{11} \geq a_{12}a_{21}a_{22}^{-1}$, which implies

$$a_{11} \geq a_{12}a_{21}a^{-1}. \tag{12.18}$$

Similarly,

$$a_{22} \geq a_{12}a_{21}\left(a\right)^{-1}. \tag{12.19}$$

By (12.17), (12.18) and (12.19) we see $(B^*)^2 = A$. □

We now discuss matters of uniqueness in an attempt to define the quantity $\sqrt{A}$. Suppose $d(A) \geq 0$ and $B = (b_{ij}) \in \mathbb{R}^{2\times 2}$ such that $B^2 = A$. Define $C := B^2$. Then

$$C = \begin{pmatrix} b_{11}^2 \oplus b_{12}b_{21} & b_{12}b \\ b_{21}b & b_{22}^2 \oplus b_{12}b_{21} \end{pmatrix}.$$

We proceed by considering cases on the diagonal elements of C.

- It is easy to see that if

$$\begin{cases} c_{11} = b_{11}^2 \\ c_{22} = b_{22}^2, \end{cases}$$

  then $B = B^*$. Indeed, in this case $b_{11} = \sqrt{a_{11}}$ and $b_{22} = \sqrt{a_{22}}$. Then $b = \sqrt{a}$. It follows that $b_{12} = a_{12}\sqrt{a}^{-1}$ and, similarly, $b_{21} = a_{21}\sqrt{a}^{-1}$.

- Suppose

$$\begin{cases} c_{11} = b_{12}b_{21} \geq b_{11}^2 \\ c_{22} = b_{22}^2 \geq b_{12}b_{21}. \end{cases} \tag{12.20}$$

It follows that $b_{22} > b_{11}$, so

$$C = \begin{pmatrix} b_{12}b_{21} & b_{12}b_{22} \\ b_{21}b_{22} & b_{22}^2 \end{pmatrix}.$$

We then have that

$$\begin{cases} b_{11} \leq \sqrt{a_{12}a_{21}a_{22}^{-1}} \\ b_{22} = \sqrt{a_{22}} \\ b_{12} = a_{12}\left(\sqrt{a_{22}}\right)^{-1} \\ b_{21} = a_{21}\left(\sqrt{a_{22}}\right)^{-1}. \end{cases}$$

It follows that $a_{11} = b_{12}b_{21} = a_{12}a_{21}a_{22}^{-1}$, which implies $d(A) = 0$. So, $b_{11} \leq \sqrt{a_{12}a_{21}a_{22}^{-1}}$ if and only if $b_{11} \leq \sqrt{a_{11}}$.

We see in this case that B^* is the component-wise maximum over all matrices B satisfying $B^2 = A$.

- Similarly, if

$$\begin{cases} c_{11} = b_{11}^2 \geq b_{12}b_{21} \\ c_{22} = b_{12}b_{21} > b_{22}^2, \end{cases}$$

then $d(A) = 0$ and B^* is the component-wise maximum over all matrices B satisfying $B^2 = A$.

- Finally then, if

$$c_{11} = b_{12}b_{21} = c_{22} > b_{11}^2 \oplus b_{22}^2,$$

then

$$C = \begin{pmatrix} b_{12}b_{21} & b_{12}b \\ b_{21}b & b_{12}b_{21} \end{pmatrix}.$$

One can show in this case that

$$B = \begin{pmatrix} \sqrt{a_{12}a_{21}a^{-1}} & \sqrt{a_{12}a_{21}^{-1}a} \\ \sqrt{a_{21}a_{12}^{-1}a} & \sqrt{a_{12}a_{21}a^{-1}} \end{pmatrix}.$$

It follows that

$$B^2 = \begin{pmatrix} a_{12}a_{21}a^{-1} \oplus a & a_{12} \\ a_{21} & a_{12}a_{21}a^{-1} \oplus a \end{pmatrix}$$

and so $B^2 = A$ if and only if $a_{12}a_{21}a^{-1} \leq a$ if and only if $d(A) \geq 0$, which holds.

Interestingly then in this final case, there are exactly two matrices B satisfying $B^2 = A$. Namely

$$B_1 := \begin{pmatrix} \sqrt{a} & a_{12}\left(\sqrt{a}\right)^{-1} \\ a_{21}\left(\sqrt{a}\right)^{-1} & \sqrt{a_{22}} \end{pmatrix}, B_2 := \begin{pmatrix} \sqrt{a_{12}a_{21}a^{-1}} & \sqrt{a_{12}a_{21}^{-1}a} \\ \sqrt{a_{21}a_{12}^{-1}a} & \sqrt{a_{12}a_{21}a^{-1}} \end{pmatrix}.$$

Note that

$$d(B_1) = d(A) > 0, d(B_2) = d^{-1}(A) < 0.$$

It follows that there does not exist a matrix D such that $D^2 = B_2$, but we know there exists a matrix E such that $E^4 = A$. This inconsistency with the matrix B_2 leads to us to suggest that $B_1 = B^*$ is the natural choice for $\sqrt{A}$.

Having considered all cases, we conclude that B^* is the "natural choice" for $\sqrt{A}$, in that $(B^*)^2 = A$ and B^* is consistent with the properties that we would expect of the square root of a matrix.

Definition 12.2. *Let $A \in \mathbb{R}^{2\times 2}$. If $d(A) \geq 0$, then*

$$\sqrt{A} := \begin{pmatrix} \sqrt{a_{11}} & a_{12}\left(\sqrt{a}\right)^{-1} \\ a_{21}\left(\sqrt{a}\right)^{-1} & \sqrt{a_{22}} \end{pmatrix}.$$

In order to better understand matrix roots, we look here to better understand matrix powers.

Let $B = (b_{ij}) \in \mathbb{R}^{2\times 2}$. Denote $\underline{b} := b_{11} \oplus' b_{22}$ and recall $b = b_{11} \oplus b_{22}$. Note that

$$\underline{b}^2 b_{12}^{-1} b_{21}^{-1} \leq d(B) \leq b^2 b_{12}^{-1} b_{21}^{-1}. \tag{12.21}$$

Our aim is to give explicit formulas for natural powers of the 2×2 matrix B. We consider three cases:

1 $b^2 b_{12}^{-1} b_{21}^{-1} < 0$

2 $\underline{b}^2 b_{12}^{-1} b_{21}^{-1} \leq 0 \leq b^2 b_{12}^{-1} b_{21}^{-1}$

3 $0 < \underline{b}^2 b_{12}^{-1} b_{21}^{-1}$

Case 1

$$B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}.$$

$$\begin{aligned} B^2 &= \begin{pmatrix} b_{11}^2 \oplus b_{12} b_{21} & b_{12} b \\ b_{21} b & b_{22}^2 \oplus b_{12} b_{21} \end{pmatrix} \\ &= \begin{pmatrix} b_{12} b_{21} & b_{12} b \\ b_{21} b & b_{12} b_{21} \end{pmatrix}. \end{aligned}$$

$$B^3 = \begin{pmatrix} b_{12}b_{21} & b_{12}b \\ b_{21}b & b_{12}b_{21} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$$
$$= \begin{pmatrix} b_{11}b_{12}b_{21} \oplus b_{12}b_{21}b & b_{12}\left(b_{12}b_{21} \oplus b_{22}b\right) \\ b_{21}\left(b_{11}b \oplus b_{12}b_{21}\right) & b_{12}b_{21}b \oplus b_{12}b_{21}b_{22} \end{pmatrix}$$
$$= \begin{pmatrix} b_{12}b_{21}b & b_{12}^2b_{21} \\ b_{21}^2b_{12} & b_{12}b_{21}b \end{pmatrix}.$$

Now, let $P(l)$ be the statement:

$$\text{“}B^{2l} = \begin{pmatrix} b_{12}^l b_{21}^l & b_{12}^l b_{21}^{l-1} b \\ b_{21}^l b_{12}^{l-1} b & b_{12}^l b_{21}^l \end{pmatrix}\text{”}.$$

Clearly, $P(1)$ holds. Now suppose $P(l)$ holds for some $l \geq 1$. Then

$$B^{2(l+1)} = B^{2l}B^2$$
$$= \begin{pmatrix} b_{12}^l b_{21}^l & b_{12}^l b_{21}^{l-1} b \\ b_{21}^l b_{12}^{l-1} b & b_{12}^l b_{21}^l \end{pmatrix} \begin{pmatrix} b_{12}b_{21} & b_{12}b \\ b_{21}b & b_{12}b_{21} \end{pmatrix}$$
$$= \begin{pmatrix} b_{12}^l b_{21}^l \left(b_{12}b_{21} \oplus b^2\right) & b_{12}^{l+1} b_{21}^l b \oplus b_{12}^{l+1} b_{21}^l b \\ b_{21}^{l+1} b_{12}^l b \oplus b_{21}^{l+1} b_{12}^l b & b_{12}^l b_{21}^l \left(b^2 \oplus b_{12}b_{21}\right) \end{pmatrix}$$
$$= \begin{pmatrix} b_{12}^{l+1} b_{21}^{l+1} & b_{12}^{l+1} b_{21}^l b \\ b_{21}^{l+1} b_{12}^l b & b_{12}^{l+1} b_{21}^{l+1} \end{pmatrix},$$

and so $P(l+1)$ holds. By induction, $P(l)$ holds for all $l \in \mathbb{N}$.

Corollary 12.3. *Let $B = (b_{ij}) \in \mathbb{R}^{2\times 2}$ such that $b^2 b_{12}^{-1} b_{21}^{-1} < 0$. Let $l \in \mathbb{N}$. Then*

$$d\left(B^{2l}\right) = b_{12} b_{21} b^{-2} > 0.$$

Let $Q(l)$ be the statement:

$$\text{“}B^{2l+1} = \begin{pmatrix} b_{12}^l b_{21}^l b & b_{12}^{l+1} b_{21}^l \\ b_{21}^{l+1} b_{12}^l & b_{12}^l b_{21}^l b \end{pmatrix}\text{”}.$$

Clearly, $Q(1)$ holds. Suppose now that $Q(l)$ holds for some $l \geq 1$. Then

$$\begin{aligned} B^{2(l+1)+1} &= B^{2l+1} B^2 \\ &= \begin{pmatrix} b_{12}^l b_{21}^l b & b_{12}^{l+1} b_{21}^l \\ b_{21}^{l+1} b_{12}^l & b_{12}^l b_{21}^l b \end{pmatrix} \begin{pmatrix} b_{12} b_{21} & b_{12} b \\ b_{21} b & b_{12} b_{21} \end{pmatrix} \\ &= \begin{pmatrix} b_{12}^{l+1} b_{21}^{l+1} b \oplus b_{12}^{l+1} b_{21}^{l+1} b & b_{12}^{l+1} b_{21}^l \left(b_{12} b_{21} \oplus b^2\right) \\ b_{21}^{l+1} b_{12}^l \left(b_{12} b_{21} \oplus b^2\right) & b_{12}^{l+1} b_{21}^{l+1} b \end{pmatrix} = \begin{pmatrix} b_{12}^{l+1} b_{21}^{l+1} b & b_{12}^{l+2} b_{21}^{l+1} \\ b_{21}^{l+2} b_{12}^{l+1} & b_{12}^{l+1} b_{21}^{l+1} b \end{pmatrix}, \end{aligned}$$

and so $Q(l+1)$ holds. It follows by induction that $Q(l)$ holds for all $l \in \mathbb{N}$.

Corollary 12.4. *Let $B = (b_{ij}) \in \mathbb{R}^{2\times 2}$ such that $b^2 b_{12}^{-1} b_{21}^{-1} < 0$. Let $l \in \mathbb{N}$. Then*

$$d\left(B^{2l+1}\right) = b^2 b_{12}^{-1} b_{21}^{-1} < 0.$$

Case 2 Recall

$$\underline{b}^2 b_{12}^{-1} b_{21}^{-1} \leq 0 \leq b^2 b_{12}^{-1} b_{21}^{-1}.$$

Without loss of generality, assume $\underline{b} = b_{11} \leq b_{22} = b$. (The case $b_{11} \geq b_{22}$ is similar). Now,

$$B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b \end{pmatrix}.$$

$$\begin{aligned} B^2 &= \begin{pmatrix} b_{11}^2 \oplus b_{12}b_{21} & b_{12}b \\ b_{21}b & b^2 \oplus b_{12}b_{21} \end{pmatrix} \\ &= \begin{pmatrix} b_{12}b_{21} & b_{12}b \\ b_{21}b & b^2 \end{pmatrix}. \end{aligned}$$

$$\begin{aligned} B^3 &= \begin{pmatrix} b_{12}b_{21} & b_{12}b \\ b_{21}b & b^2 \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b \end{pmatrix} \\ &= \begin{pmatrix} b_{12}b_{21}\left(b_{11} \oplus b\right) & b_{12}\left(b_{12}b_{21} \oplus b^2\right) \\ b_{21}\left(b_{11}b \oplus b^2\right) & b_{12}b_{21}b \oplus b^3 \end{pmatrix} \\ &= \begin{pmatrix} b_{12}b_{21}b & b_{12}b^2 \\ b_{21}b^2 & b^3 \end{pmatrix}. \end{aligned}$$

Let $P(l)$ be the statement:

$$\text{“}B^l = \begin{pmatrix} b_{12}b_{21}b^{l-2} & b_{12}b^{l-1} \\ b_{21}b^{l-1} & b^l \end{pmatrix}\text{”}.$$

Clearly $P(2)$ and $P(3)$ hold. Suppose now that $P(l)$ holds for some $l \geq 2$. Then

$$
\begin{aligned}
B^{l+1} &= \begin{pmatrix} b_{12}b_{21}b^{l-2} & b_{12}b^{l-1} \\ b_{21}b^{l-1} & b^{l} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b \end{pmatrix} \\
&= \begin{pmatrix} b_{12}b_{21}\left(b^{l-2}b_{11} \oplus b^{l-1}\right) & b_{12}b^{l-2}\left(b_{12}b_{21} \oplus b^{2}\right) \\ b_{21}b^{l-1}\left(b_{11} \oplus b\right) & b^{l-1}\left(b_{12}b_{21} \oplus b^{2}\right) \end{pmatrix} \\
&= \begin{pmatrix} b_{12}b_{21}b^{l-1} & b_{12}b^{l} \\ b_{21}b^{l} & b^{l+1} \end{pmatrix},
\end{aligned}
$$

and so $P(l)$ holds. It follows that $P(l)$ holds for all $l \in \mathbb{N}$.

Similarly, if $b_{11} \geq b_{22}$, then

$$
B^{l} = \begin{pmatrix} b^{l} & b_{12}b^{l-1} \\ b_{21}b^{l-1} & b_{12}b_{21}b^{l-2} \end{pmatrix}.
$$

Corollary 12.5. *Let $B = (b_{ij}) \in \mathbb{R}^{2\times 2}$ such that $\underline{b}^{2}b_{12}^{-1}b_{21}^{-1} \leq 0 \leq b^{2}b_{12}^{-1}b_{21}^{-1}$. Let $l \in \mathbb{N}$. Then*

$$
d\left(B^{l}\right) = 0.
$$

Case 3 Recall

$$
0 < \underline{b}^{2}b_{12}^{-1}b_{21}^{-1}.
$$

Assume without loss of generality $\underline{b} = b_{11} \leq b_{22} = b$. (The case $b_{11} \geq b_{22}$ is similar).

$$B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b \end{pmatrix}.$$

$$\begin{aligned} B^2 &= \begin{pmatrix} b_{11}^2 \oplus b_{12}b_{21} & b_{12}b \\ b_{21}b & b^2 \oplus b_{12}b_{21} \end{pmatrix} \\ &= \begin{pmatrix} b_{11}^2 & b_{12}b \\ b_{21}b & b^2 \end{pmatrix}. \end{aligned}$$

$$\begin{aligned} B^3 &= \begin{pmatrix} b_{11}^2 & b_{12}b \\ b_{21}b & b^2 \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b \end{pmatrix} \\ &= \begin{pmatrix} b_{11}^3 \oplus b_{12}b_{21}b & b_{12}\left(b_{11}^2 \oplus b^2\right) \\ b_{21}\left(b_{11}b \oplus b^2\right) & b_{12}b_{21}b \oplus b^3 \end{pmatrix} \qquad = \begin{pmatrix} b_{11}^3 \oplus b_{12}b_{21}b & b_{12}b^2 \\ b_{21}b^2 & b^3 \end{pmatrix}. \end{aligned}$$

Let $P(l)$ be the statement

$$\text{“}B^l = \begin{pmatrix} b_{11}^l \oplus b_{12}b_{21}b^{l-2} & b_{12}b^{l-1} \\ b_{21}b^{l-1} & b^l \end{pmatrix}\text{”}.$$

Clearly $P(2)$ and $P(3)$ hold. Suppose $P(l)$ holds for some $l \geq 2$. Then

$$
\begin{aligned}
B^{l+1} &= \begin{pmatrix} b_{11}^{l} \oplus b_{12}b_{21}b^{l-2} & b_{12}b^{l-1} \\ b_{21}b^{l-1} & b^{l} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b \end{pmatrix} \\
&= \begin{pmatrix} b_{11}^{l+1} \oplus b_{12}b_{21}\left(b_{11}b^{l-1} \oplus b^{l-1}\right) & b_{12}\left(b_{11}^{l} \oplus b_{12}b_{21}b^{l-2}\right) \oplus b_{12}b^{l} \\ b_{21}b_{11}b^{l-1} \oplus b_{21}b^{l} & b_{12}b_{21}b^{l-1} \oplus b^{l+1} \end{pmatrix} \\
&= \begin{pmatrix} b_{11}^{l+1} \oplus b_{12}b_{21}b^{l-1} & b_{12}b^{l} \\ b_{21}b^{l} & b^{l} \end{pmatrix},
\end{aligned}
$$

and so $P(l+1)$ holds. It follows by induction that $P(l)$ holds for all $l \in \mathbb{N}$.

Similarly, if $b_{11} \geq b_{22}$, then

$$
B^{l} = \begin{pmatrix} b^{l} & b_{12}b^{l-1} \\ b_{21}b^{l-1} & b_{22}^{l} \oplus b_{12}b_{21}b^{l-2} \end{pmatrix}.
$$

Lemma 12.6. *Let $B = (b_{ij}) \in \mathbb{R}^{2\times 2}$ such that $\underline{b}^{2}b_{12}^{-1}b_{21}^{-1} > 0$. Let $l \in \mathbb{N}$. Then*

$$
d\left(B^{l}\right) \geq 0.
$$

Proof. Without loss of generality assume $b_{11} \leq b_{22}$ (the case $b_{11} \geq b_{22}$ is similar). Now,

$$
\begin{aligned}
d\left(B^{l}\right) &= b_{11}^{l}b^{l}b_{12}^{-1}b_{21}^{-1}b^{2-2l} \oplus b_{12}b_{21}b^{l-2}b^{l}b_{12}^{-1}b_{21}^{-1}b^{2-2l} \\
&= b_{11}^{l}b^{2-l}b_{12}^{-1}b_{21}^{-1} \oplus 0 \\
&\geq 0.
\end{aligned}
$$

□

12.2.1 Odd roots

Let $A = (a_{ij}) \in \mathbb{R}^{2\times 2}$. Let $k \in \mathbb{N}$. We wish to find $B = (b_{ij}) \in \mathbb{R}^{2\times 2}$ such that $B^{2k+1} = A$.

- $d(A) < 0$

  It follows from Corollaries 12.4, 12.5 and Lemma 12.6 that if there exists B such that $B^{2k+1} = A$, then $b^2 b_{12}^{-1} b_{21}^{-1} < 0$.

  It follows that

$$A = \begin{pmatrix} b_{12}^k b_{21}^k b & b_{12}^{k+1} b_{21}^k \\ b_{21}^{k+1} b_{12}^k & b_{12}^k b_{21}^k b \end{pmatrix}.$$

  So,

$$\begin{cases} a_{11} = a_{22} = a = b_{12}^k b_{21}^k b \\ a_{12} = b_{12}^{k+1} b_{21}^k \\ a_{21} = b_{21}^{k+1} b_{12}^k \end{cases}$$

  This is a system of three linear equations in three unknowns and can be solved easily (by Gaussian elimination, say). The unique solution is

$$\begin{cases} b = a\left(a_{12}a_{21}\right)^{\frac{-k}{2k+1}} \\ b_{12} = a_{12}\left(a_{12}a_{21}\right)^{\frac{-k}{2k+1}} \\ b_{21} = a_{21}\left(a_{12}a_{21}\right)^{\frac{-k}{2k+1}}. \end{cases}$$

  It follows that

$$B = (a_{12}a_{21})^{\frac{-k}{2k+1}} \begin{pmatrix} \lambda & a_{12} \\ a_{21} & \mu \end{pmatrix},$$

where $\lambda \oplus \mu = a$.

In particular, the component-wise maximum over all such B is given by

$$B^* = (a_{12}a_{21})^{\frac{-k}{2k+1}} \begin{pmatrix} a & a_{12} \\ a_{21} & a \end{pmatrix}. \tag{12.22}$$

- $d(A) > 0$

It follows from Corollaries 12.4, 12.5 and Lemma 12.6 that $\underline{b}^2 b_{12}^{-1} b_{21}^{-1} > 0$. Let us assume first that $b_{11} \leq b_{22}$ (the case $b_{11} \geq b_{22}$ is similar). Now, if there exists B such that $B^{2k+1} = A$, then

$$A = \begin{pmatrix} b_{11}^{2k+1} \oplus b_{12}b_{21}b^{2k-1} & b_{12}b^{2k} \\ b_{21}b^{2k} & b_{22}^{2k+1} \end{pmatrix}.$$

We have then

$$\begin{cases} b_{22} = a_{22}^{\frac{1}{2k+1}} = a_{22}a_{22}^{\frac{-2k}{2k+1}} \\ b_{12} = a_{12}a_{22}^{\frac{-2k}{2k+1}} \\ b_{21} = a_{21}a_{22}^{\frac{-2k}{2k+1}}. \end{cases}$$

Note that

$$
\begin{aligned}
b_{12}b_{21}b^{2k-1} &= a_{12}a_{22}^{\frac{-2k}{2k+1}}a_{21}a_{22}^{\frac{-2k}{2k+1}}a_{22}^{\frac{2k-1}{2k+1}} \\
&= a_{12}a_{21}a_{22}^{\frac{-(2k+1)}{2k+1}} \\
&= a_{12}a_{21}a_{22}^{-1} \\
&< a_{11}.
\end{aligned}
$$

We then have

$$
b_{11}^{2k+1} = a_{11} \Leftrightarrow b_{11} = a_{11}^{\frac{1}{2k+1}}.
$$

To conclude,

$$
\begin{aligned}
B &= (b_{ij}) \\
&= \begin{pmatrix} a_{11}^{\frac{1}{2k+1}} & a_{12}a_{22}^{\frac{-2k}{2k+1}} \\ a_{21}a_{22}^{\frac{-2k}{2k+1}} & a_{22}^{\frac{1}{2k+1}} \end{pmatrix} \\
&= \begin{pmatrix} a_{11}a_{11}^{\frac{-2k}{2k+1}} & a_{12}a_{22}^{\frac{-2k}{2k+1}} \\ a_{21}a_{22}^{\frac{-2k}{2k+1}} & a_{22}a_{22}^{\frac{-2k}{2k+1}} \end{pmatrix} \\
&= \left(a_{ij}\left(a_{ii} \oplus a_{jj}\right)^{\frac{-2k}{2k+1}}\right).
\end{aligned}
$$

Similarly, if $b_{11} \geq b_{22}$, then

$$
B = (b_{ij}) = \left(a_{ij}\left(a_{ii} \oplus a_{jj}\right)^{\frac{-2k}{2k+1}}\right).
$$

- $d\left(A\right)=0$

  It follows from Corollaries 12.4, 12.5 and Lemma 12.6 that either

  - $\underline{b}^2 b_{12}^{-1} b_{21}^{-1} \leq 0 \leq b^2 b_{12}^{-1} b_{21}^{-1}$ or
  - $\underline{b}^2 b_{12}^{-1} b_{21}^{-1} > 0.$

  Firstly, we consider $\underline{b}^2 b_{12}^{-1} b_{21}^{-1} \leq 0 \leq b^2 b_{12}^{-1} b_{21}^{-1}$. Let us assume that $b_{11} \leq b_{22}$ (the case $b_{11} \geq b_{22}$ is similar). We have then that

$$B^{2k+1} = \begin{pmatrix} b_{12} b_{21} b^{2k-1} & b_{12} b_{22}^{2k} \\ b_{21} b^{2k} & b^{2k+1} \end{pmatrix}.$$

  We have then

$$\begin{cases} b = b_{22} = a_{22}^{\frac{1}{2k+1}} = a_{22} a_{22}^{\frac{-2k}{2k+1}} \\ b_{12} = a_{12} a_{22}^{\frac{-2k}{2k+1}} \\ b_{21} = a_{21} a_{22}^{\frac{-2k}{2k+1}}. \end{cases}$$

  Note that

$$\begin{cases} b_{11}^2 \leq b_{12} b_{21} \\ b_{11}^{2k-1} \leq b_{22}^{2k-1}. \end{cases}$$

  Together, these imply

$$
\begin{aligned}
b_{11}^{2k+1} &\leq b_{12}b_{21}b_{22}^{2k-1} \\
&= a_{12}a_{22}^{\frac{-2k}{2k+1}}a_{21}a_{22}^{\frac{-2k}{2k+1}}a_{22}^{\frac{2k-1}{2k+1}} \\
&= a_{12}a_{21}a_{22}^{\frac{-(2k+1)}{2k+1}} \\
&= a_{12}a_{21}a_{22}^{-1} \\
&= a_{11}.
\end{aligned}
$$

We then have

$$
b_{11} \leq a_{11}^{\frac{1}{2k+1}} = a_{11}a_{11}^{\frac{-2k}{2k+1}}.
$$

Conversely, if $b_{11} = a_{11}^{\frac{1}{2k+1}}$, then it follows that

$$
\begin{cases}
b_{11}^2 = b_{12}b_{21} \\
b_{11}^{2k-1} = b_{22}^{2k-1}.
\end{cases}
$$

It follows that $b_{11} = b_{22}$ and

$$
\underline{b}^2 b_{12}^{-1}b_{21}^{-1} = d\left(B\right) = b^2 b_{12}^{-1}b_{21}^{-1} = 0.
$$

We can show in this case that B^{2k+1} is indeed equal to A, showing that the inequality $b_{11} \leq a_{11}^{\frac{1}{2k}}$ is the tightest possible.

We conclude that

$$
B = \left(\begin{array}{cc}
\lambda a_{11}^{\frac{-2k}{2k+1}} & a_{12}a_{22}^{\frac{-2k}{2k+1}} \\
a_{21}a_{22}^{\frac{-2k}{2k+1}} & a_{22}a_{22}^{\frac{-2k}{2k+1}}
\end{array} \right),
$$

where $\lambda \le a_{11}$. It follows that the component-wise maximum over all such B is given by

$$\begin{aligned} B^* &= (b_{ij}) \\ &= \begin{pmatrix} a_{11}a_{11}^{\frac{-2k}{2k+1}} & a_{12}a_{22}^{\frac{-2k}{2k+1}} \\ a_{21}a_{22}^{\frac{-2k}{2k+1}} & a_{22}a_{22}^{\frac{-2k}{2k+1}} \end{pmatrix} \\ &= \left(a_{ij}\left(a_{ii} \oplus a_{jj}\right)^{\frac{-2k}{2k+1}}\right). \end{aligned}$$

Similarly, when $b_{11} \ge b_{22}$ we find

$$B^* = (b_{ij}) = \left(a_{ij}\left(a_{ii} \oplus a_{jj}\right)^{\frac{-2k}{2k+1}}\right).$$

Next, consider $0 < \underline{b}^2 b_{12}^{-1} b_{21}^{-1}$. Again, let us suppose $b_{11} \le b_{22}$ (the same results are obtained in the case $b_{11} \ge b_{22}$). In this case we have

$$B = \begin{pmatrix} b_{11}^{2k+1} \oplus b_{12}b_{21}b^{2k-1} & b_{12}b^{2k} \\ b_{21}b^{2k} & b^{2k+1} \end{pmatrix}.$$

This yields

$$\begin{cases} b = a_{22}^{\frac{1}{2k+1}} = a_{22}a_{22}^{\frac{-2k}{2k+1}} \\ b_{12} = a_{12}a_{22}^{\frac{-2k}{2k+1}} \\ b_{21} = a_{21}a_{22}^{\frac{-2k}{2k+1}}. \end{cases}$$

Note that

$$\begin{aligned}
b_{12}b_{21}b^{2k-1} &= a_{12}a_{22}^{\frac{-2k}{2k+1}}a_{21}a_{22}^{\frac{-2k}{2k+1}}a_{22}^{\frac{2k-1}{2k+1}} \\
&= a_{12}a_{21}a_{22}^{\frac{-2k-1}{2k+1}} \\
&= a_{12}a_{21}a_{22}^{-1} \\
&= a_{11}.
\end{aligned}$$

It follows that

$$B = \begin{pmatrix} \lambda a_{11}^{\frac{-2k}{2k+1}} & a_{12}a_{22}^{\frac{-2k}{2k+1}} \\ a_{21}a_{22}^{\frac{-2k}{2k+1}} & a_{22}a_{22}^{\frac{-2k}{2k+1}} \end{pmatrix},$$

where $\lambda \leq a_{11}$.

The component-wise maximum over all such B is given by

$$\begin{aligned}
B^* &= (b_{ij}) \\
&= \begin{pmatrix} a_{11}a_{11}^{\frac{-2k}{2k+1}} & a_{12}a_{22}^{\frac{-2k}{2k+1}} \\ a_{21}a_{22}^{\frac{-2k}{2k+1}} & a_{22}a_{22}^{\frac{-2k}{2k+1}} \end{pmatrix} \\
&= \left(a_{ij}\left(a_{ii} \oplus a_{jj}\right)^{\frac{-2k}{2k+1}}\right).
\end{aligned}$$

Similarly, if $b_{11} \geq b_{22}$, then we find

$$B^* = (b_{ij}) = \left(a_{ij}\left(a_{ii} \oplus a_{jj}\right)^{\frac{-2k}{2k+1}}\right).$$

12.2.2 Even roots

Let $A = (a_{ij}) \in \mathbb{R}^{2\times 2}$. Let $k \in \mathbb{N}$. We wish to find $B = (b_{ij}) \in \mathbb{R}^{2\times 2}$ such that $B^{2k} = A$.

- $d(A) < 0$

  By Corollaries 12.3, 12.5 and Lemma 12.6 we see that there does not exist B such that $B^{2k} = A$.

- $d(A) > 0$

  Suppose that $b_{11} \leq b_{22}$, the results for $b_{11} \geq b_{22}$ are similar. By Corollaries 12.3, 12.5 and Lemma 12.6 we see that

$$B^{2k} = \begin{pmatrix} b_{11}^{2k} \oplus b_{12}b_{21}b_{22}^{2k-2} & b_{12}b_{22}^{2k-1} \\ b_{21}b_{22}^{2k-1} & b_{22}^{2k} \end{pmatrix}.$$

  From this we have

$$\begin{cases} b_{22} = a_{22}^{\frac{1}{2k}} = a_{22}a_{22}^{\frac{-2k+1}{2k}} \\ b_{12} = a_{12}a_{22}^{\frac{-2k+1}{2k}} \\ b_{21} = a_{21}a_{22}^{\frac{-2k+1}{2k}}. \end{cases}$$

  Note that

$$
\begin{aligned}
b_{12}b_{21}b_{22}^{2k-2} &= a_{12}a_{22}^{\frac{-2k+1}{2k}}\,a_{21}a_{22}^{\frac{-2k+1}{2k}}\,a_{22}^{\frac{2k-2}{2k}} \\
&= a_{12}a_{21}a_{22}^{\frac{-2k}{2k}} \\
&= a_{12}a_{21}a_{22}^{-1} \\
&< a_{11}.
\end{aligned}
$$

It follows that

$$
\begin{aligned}
& b_{11}^{2k} = a_{11} \\
\Leftrightarrow & b_{11} = a_{11}^{\frac{1}{2k}} = a_{11}a_{11}^{\frac{-2k+1}{2k}}.
\end{aligned}
$$

To conclude,

$$
\begin{aligned}
B &= (b_{ij}) \\
&= \begin{pmatrix} a_{11}a_{11}^{\frac{-2k+1}{2k}} & a_{12}a_{22}^{\frac{-2k+1}{2k}} \\ a_{21}a_{22}^{\frac{-2k+1}{2k}} & a_{22}a_{22}^{\frac{-2k+1}{2k}} \end{pmatrix} \\
&= \left(a_{ij}\left(a_{ii} \oplus a_{jj}\right)^{\frac{-2k+1}{2k}}\right).
\end{aligned}
$$

Similarly, if we have $b_{11} \geq b_{22}$, then we obtain

$$
B = \left(a_{ij}\left(a_{ii} \oplus a_{jj}\right)^{\frac{-2k+1}{2k}}\right).
$$

- $d(A) = 0$

  By Corollaries 12.3, 12.5 and Lemma 12.6 we have two possible cases, namely:

  - $\underline{b}^2 b_{12}^{-1} b_{21}^{-1} \leq 0 \leq b^2 b_{12}^{-1} b_{21}^{-1}$ or
  - $0 < \underline{b}^2 b_{12}^{-1} b_{21}^{-1}$.

  First, let us consider $\underline{b}^2 b_{12}^{-1} b_{21}^{-1} \leq 0 \leq b^2 b_{12}^{-1} b_{21}^{-1}$. Suppose $b_{11} \leq b_{22}$, the results are similar for $b_{11} \geq b_{22}$. We have

$$B^{2k} = \begin{pmatrix} b_{12} b_{21} b_{22}^{2k-2} & b_{12} b_{22}^{2k-1} \\ b_{21} b_{22}^{2k-1} & b_{22}^{2k} \end{pmatrix}.$$

  It follows that

$$\begin{cases} b_{22} = a_{22}^{\frac{1}{2k}} = a_{22} a_{22}^{\frac{-2k+1}{2k}} \\ b_{12} = a_{12} a_{22}^{\frac{-2k+1}{2k}} \\ b_{21} = a_{21} a_{22}^{\frac{-2k+1}{2k}}. \end{cases}$$

  Note

$$\begin{aligned} b_{12} b_{21} b_{22}^{2k-2} &= a_{12} a_{22}^{\frac{-2k+1}{2k}} a_{21} a_{22}^{\frac{-2k+1}{2k}} a_{22}^{\frac{2k-2}{2k}} \\ &= a_{12} a_{21} a_{22}^{\frac{-2k}{2k}} \\ &= a_{12} a_{21} a_{22}^{-1} \\ &= a_{11}. \end{aligned}$$

  We have then that

$$b_{11} \leq a_{11}^{\frac{1}{2k}} = a_{11} a_{11}^{\frac{-2k+1}{2k}}.$$

In conclusion,

$$B = \begin{pmatrix} \lambda & a_{12} a_{22}^{\frac{-2k+1}{2k}} \\ a_{21} a_{22}^{\frac{-2k+1}{2k}} & a_{22} a_{22}^{\frac{-2k+1}{2k}} \end{pmatrix},$$

where

$$\lambda \leq a_{11} a_{11}^{\frac{-2k+1}{2k}}.$$

The component-wise maximum over all such B is given by

$$\begin{aligned} B^* &= (b_{ij}) \\ &= \begin{pmatrix} a_{11} a_{11}^{\frac{-2k+1}{2k}} & a_{12} a_{22}^{\frac{-2k+1}{2k}} \\ a_{21} a_{22}^{\frac{-2k+1}{2k}} & a_{22} a_{22}^{\frac{-2k+1}{2k}} \end{pmatrix} \\ &= \left(a_{ij} \left(a_{ii} \oplus a_{jj} \right)^{\frac{-2k+1}{2k}} \right). \end{aligned}$$

Similarly, if $b_{11} \geq b_{22}$, then

$$B^* = \left(a_{ij} \left(a_{ii} \oplus a_{jj} \right)^{\frac{-2k+1}{2k}} \right).$$

Next, consider $0 < \underline{b}^2 b_{12}^{-1} b_{21}^{-1}$. Suppose $b_{11} \leq b_{22}$, the results are similar for $b_{11} \geq b_{22}$. We have

$$B^{2k} = \begin{pmatrix} b_{11}^{2k} \oplus b_{12} b_{21} b_{22}^{2k-2} & b_{12} b_{22}^{2k-1} \\ b_{21} b_{22}^{2k-1} & b_{22}^{2k} \end{pmatrix}.$$

Note that

$$\begin{aligned} 0 &= d\left(A\right) \\ &= d\left(B^{2k}\right) \\ &= b_{11}^{2k} b_{22}^{2-2k} b_{12}^{-1} b_{21}^{-1} \oplus 0. \end{aligned}$$

It follows that

$$b_{11}^{2k} = b_{12} b_{21} b_{22}^{2k-2},$$

and so

$$B^{2k} = \begin{pmatrix} b_{11}^{2k} & b_{12} b_{22}^{2k-1} \\ b_{21} b_{22}^{2k-1} & b_{22}^{2k} \end{pmatrix}.$$

We have

$$\begin{cases} b_{11} = a_{11}^{\frac{1}{2k}} = a_{11} a_{11}^{\frac{-2k+1}{2k}} \\ b_{22} = a_{22}^{\frac{1}{2k}} = a_{22} a_{22}^{\frac{-2k+1}{2k}} \\ b_{12} = a_{12} a_{22}^{\frac{-2k+1}{2k}} \\ b_{21} = a_{22}^{\frac{-2k+1}{2k}}. \end{cases}$$

To conclude,

$$
\begin{aligned}
B &= (b_{ij}) \\
&\begin{pmatrix} a_{11}a_{11}^{\frac{-2k+1}{2k}} & a_{12}a_{22}^{\frac{-2k+1}{2k}} \\ a_{21}a_{22}^{\frac{-2k+1}{2k}} & a_{22}a_{22}^{\frac{-2k+1}{2k}} \end{pmatrix} \\
&= \left(a_{ij}\left(a_{ii} \oplus a_{jj}\right)^{\frac{-2k+1}{2k}}\right).
\end{aligned}
$$

Similarly, when $b_{11} \geq b_{22}$, we obtain

$$
B = \left(a_{ij}\left(a_{ii} \oplus a_{jj}\right)^{\frac{-2k+1}{2k}}\right).
$$

We summarise our results with some Theorems.

Theorem 12.7. *Let $A = (a_{ij}) \in \mathbb{R}^{2\times 2}$ such that $d(A) \geq 0$. Let $k \in \mathbb{N}$. Define*

$$
B^* := \bigoplus \left\{B : B^k = A\right\}.
$$

Then

$$
B^* = (b_{ij}) = \left(a_{ij}\left(a_{ii} \oplus a_{jj}\right)^{\frac{1-k}{k}}\right).
$$

Corollary 12.8. *Let $A = (a_{ij}) \in \mathbb{R}^{2\times 2}$ such that $d(A) \geq 0$. In addition, suppose that $a_{11} = a_{22} =: a$. Let $k \in \mathbb{N}$. Then*

$$
\begin{aligned}
B^* &= (b_{ij}) \\
&= \left(a_{ij} a^{\frac{1-k}{k}}\right) \\
&= a^{\frac{1-k}{k}} A \\
&= \left(a^2 \oplus a_{12} a_{21}\right)^{\frac{1-k}{2k}} A.
\end{aligned}
$$

Theorem 12.9. *Let $A = (a_{ij}) \in \mathbb{R}^{2\times 2}$ such that $d(A) < 0$.*

- *If k is even, then there does not exist B such that $B^k = A$.*
- *If k is odd, then there exists B such that $B^k = A$ if and only if $a_{11} = a_{22}$. Moreover, if $a_{11} = a_{22}$, then*

  $$B = \alpha A,$$

  where

  $$\alpha = (a_{12} a_{21})^{\frac{1-k}{2k}}.$$

Note that

$$\lim_{d(A)\to 0} (a_{12} a_{21})^{\frac{1-k}{2k}} = a^{\frac{1-k}{k}}.$$

Corollary 12.10. *Let $A = (a_{ij}) \in \mathbb{R}^{2\times 2}$ such that $d(A) < 0$. In addition, suppose that $a_{11} = a_{22} =: a$. Let $k \in \mathbb{N}$ such that k is odd. Then $B^k = A$, where*

$$
\begin{aligned}
B &= (a_{12}a_{21})^{\frac{1-k}{2k}} A \\
&= \left(a^2 \oplus a_{12}a_{21}\right)^{\frac{1-k}{2k}} A.
\end{aligned}
$$

To conclude:

Theorem 12.11. *Let $A = (a_{ij}) \in \mathbb{R}^{2\times 2}$ such that $a_{11} = a_{22} = a$. Let $k \in \mathbb{N}$. Then there exists B such that $B^k = A$ if and only if $(B^*)^k = A$, where*

$$
\begin{aligned}
B^* &= \left(a^2 \oplus a_{12}a_{21}\right)^{\frac{1-k}{2k}} A \\
&= perm\,(A)^{\frac{1-k}{2k}} A
\end{aligned}
$$

and this is true when k is odd.

Remark 12.12. *Note that B^* in Theorem 12.11 is always a kth root of a 2×2 matrix when k is odd. The result does not generalise to larger matrices, as in the following example.*

Example 12.13. *Let*

$$
A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{pmatrix}, \quad k = 3.
$$

Then

$$
perm\,(A)^{\frac{1-k}{2k}} = 3^{\frac{-2}{6}} = -1
$$

and so

$$B^* = \begin{pmatrix} 0 & -1 & 0 \\ -1 & 0 & -1 \\ -1 & 1 & 0 \end{pmatrix}.$$

We can then show that

$$(B^*)^3 = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \neq A.$$

12.3 A generalisation to special types of $n \times n$ matrices

In this Section, $A \in \mathbb{R}^{n\times n}$ is a matrix such that

$$(\forall i)\,(\forall j)\,(\forall t)\, t \neq i,j; \quad a_{ij}a_{tt} \geq a_{it}a_{tj}. \tag{12.23}$$

Let $k \geq 1$ be fixed and for the remainder of this section define

$$B = (b_{ij}) = \left(a_{ij}\,(a_{ii} \oplus a_{jj})^{\frac{1-k}{k}}\right), \tag{12.24}$$

so

$$(\forall i)\, b_{ii} = a_{ii}^{\frac{1}{k}}. \tag{12.25}$$

Theorem 12.14.

$$(\forall l \geq 1)\, B^l = \left(b_{ij}^{(l)}\right) = \left(a_{ij}\,(a_{ii} \oplus a_{jj})^{\frac{l-k}{k}}\right).$$

The following lemma may be useful in the proof of Theorem 12.14.

Lemma 12.15. *Let $a, b \in \mathbb{R}$ and suppose $r, s \geq 0$. Then*

$$(a \oplus b)^{r+s} \geq a^r b^s.$$

Proof. [Proof of Lemma 12.15] We complete the proof with the use of classical notation, as follows.

$$\begin{aligned}(a \oplus b)^{r+s} &= (r+s)\max(a,b) \\ &= r\max(a,b) + s\max(a,b) \\ &\geq ra + sb \\ &= a^r b^s.\end{aligned}$$

□

Proof. [Proof of Theorem 12.14]

Let $P(l)$ be the statement

$$\text{“}B^l = \left(b_{ij}^{(l)}\right) = \left(a_{ij}\left(a_{ii} \oplus a_{jj}\right)^{\frac{l-k}{k}}\right)\text{”}$$

for $l \leq k$.

Clearly, $P(1)$ holds.

Let $1 \leq l \leq k-1$ and suppose $P(l)$ holds, that is

$$(\forall i)(\forall j)\, b_{ij}^{(l)} = a_{ij}\left(a_{ii} \oplus a_{jj}\right)^{\frac{l-k}{k}}.$$

So

$$(\forall i)\, b_{ii}^{(l)} = a_{ii}^{\frac{l}{k}}.$$

We show that $P(l+1)$ holds also.

(1) Let $i \in \mathbb{N}$. Then

$$\begin{aligned}
b_{ii}^{(l+1)} &= \bigoplus_{t\in N} b_{it}^{(l)} b_{ti} \\
&= b_{ii}^{(l)} b_{ii} \oplus \bigoplus_{t\neq i} b_{it}^{(l)} b_{ti} \\
&= a_{ii} \left(a_{ii} \oplus a_{ii}\right)^{\frac{l-k}{k}} a_{ii} \left(a_{ii} \oplus a_{ii}\right)^{\frac{1-k}{k}} \\
&\quad \oplus \bigoplus_{t\neq i} a_{it} \left(a_{ii} \oplus a_{tt}\right)^{\frac{l-k}{k}} a_{ti} \left(a_{tt} \oplus a_{ii}\right)^{\frac{1-k}{k}} \\
&= a_{ii}^{\frac{l+1}{k}} \oplus \bigoplus_{t\neq i} a_{it} a_{ti} \left(a_{ii} \oplus a_{tt}\right)^{\frac{l+1-2k}{k}} .
\end{aligned}$$

Let $t \neq i$ be fixed. We claim that

$$a_{ii}^{\frac{l+1}{k}} \geq a_{it} a_{ti} \left(a_{ii} \oplus a_{tt}\right)^{\frac{l+1-2k}{k}} .$$

To see this, first observe

$$\left(a_{ii} \oplus a_{tt}\right)^{\frac{2k-l-1}{k}} \geq a_{ii}^{\frac{k-l-1}{k}} a_{tt}^{\frac{k}{k}} \tag{12.26}$$

by Lemma 12.15 (note $k - l - 1 \geq 0$). We then have

$$\begin{aligned}
& a_{ii}^{\frac{l+1}{k}} \geq a_{it} a_{ti} \left(a_{ii} \oplus a_{tt}\right)^{\frac{l+1-2k}{k}} \\
\Leftrightarrow & a_{ii}^{\frac{l+1}{k}} \overbrace{\left(a_{ii} \oplus a_{tt}\right)^{\frac{2k-l-1}{k}}}^{\geq a_{ii}^{\frac{k-l-1}{k}} a_{tt}^{\frac{k}{k}} \text{ by (12.26)}} \geq a_{it} a_{ti} \\
\Leftarrow & a_{ii}^{\frac{l+1}{k}} a_{ii}^{\frac{k-1-l}{k}} a_{tt}^{\frac{k}{k}} \geq a_{it} a_{ti} \\
\Leftrightarrow & a_{ii} a_{tt} \geq a_{it} a_{ti},
\end{aligned}$$

which holds.

It follows

$$b_{ii}^{(l+1)} = a_{ii}^{\frac{l+1}{k}} = a_{ii} \left(a_{ii} \oplus a_{ii}\right)^{\frac{l+1-k}{k}},$$

as required.

(2) Let $i, j \in N, i \neq j$. Fist observe

$$a_{ii}^{\frac{l}{k}} \oplus a_{jj}^{\frac{1}{k}} \left(a_{ii} \oplus a_{jj}\right)^{\frac{l-1}{k}} = \left(a_{ii} \oplus a_{jj}\right)^{\frac{l}{k}}. \tag{12.27}$$

We then have

$$\begin{aligned}
b_{ij}^{(l+1)} =& \bigoplus_{t \in N} b_{it}^{(l)} b_{tj} \\
=& b_{ii}^{(l)} b_{ij} \oplus b_{ij}^{(l)} b_{jj} \oplus \bigoplus_{t \neq i,j} b_{it}^{(l)} b_{tj} \\
=& a_{ii}^{\frac{l}{k}} a_{ij} \left(a_{ii} \oplus a_{jj}\right)^{\frac{1-k}{k}} \oplus a_{ij} \left(a_{ii} \oplus a_{jj}\right)^{\frac{l-k}{k}} a_{jj}^{\frac{1}{k}} \oplus \bigoplus_{t \neq i,j} b_{it}^{(l)} b_{tj} \\
=& a_{ij} \left(a_{ii} \oplus a_{jj}\right)^{\frac{1-k}{k}} \overbrace{\left[a_{ii}^{\frac{l}{k}} \oplus a_{jj}^{\frac{1}{k}} \left(a_{ii} \oplus a_{jj}\right)^{\frac{l-1}{k}}\right]}^{=\left(a_{ii} \oplus a_{jj}\right)^{\frac{l}{k}} \text{ by (12.27)}} \oplus \bigoplus_{t \neq i,j} b_{it}^{(l)} b_{tj} \\
=& a_{ij} \left(a_{ii} \oplus a_{jj}\right)^{\frac{l+1-k}{k}} \oplus \bigoplus_{t \neq i,j} a_{it} \left(a_{ii} \oplus a_{tt}\right)^{\frac{l-k}{k}} a_{tj} \left(a_{tt} \oplus a_{jj}\right)^{\frac{1-k}{k}}.
\end{aligned}$$

Let $t \neq i, j$ be fixed. We claim that

$$a_{ij} \left(a_{ii} \oplus a_{jj}\right)^{\frac{l+1-k}{k}} \geq a_{it} \left(a_{ii} \oplus a_{tt}\right)^{\frac{l-k}{k}} a_{tj} \left(a_{tt} \oplus a_{jj}\right)^{\frac{1-k}{k}}. \tag{12.28}$$

To see this, first observe

$$\left(a_{ii} \oplus a_{jj}\right)^{\frac{l+1-k}{k}} \left(a_{ii} \oplus a_{tt}\right)^{\frac{k-l}{k}} \left(a_{tt} \oplus a_{jj}\right)^{\frac{k-1}{k}} \geq a_{tt}. \tag{12.29}$$

This follows since

$$\begin{aligned}
\left(a_{ii} \oplus a_{jj}\right)^{\frac{l+1-k}{k}} \left(a_{ii} \oplus a_{tt}\right)^{\frac{k-l}{k}} \left(a_{tt} \oplus a_{jj}\right)^{\frac{k-1}{k}} &\geq \left(a_{ii} \oplus a_{jj}\right)^{\frac{l+1-k}{k}} a_{ii}^{\frac{k-l-1}{k}} a_{tt}^{\frac{1}{k}} a_{tt}^{\frac{k-1}{k}} \\
&= \left(a_{ii} \oplus a_{jj}\right)^{\frac{l+1-k}{k}} a_{ii}^{\frac{k-l-1}{k}} a_{tt} \\
&= a_{ii}^{\frac{l+1-k}{k}} a_{ii}^{\frac{k-l-1}{k}} a_{tt} \oplus a_{jj}^{\frac{l+1-k}{k}} a_{ii}^{\frac{k-l-1}{k}} a_{tt} \\
&= a_{tt} \oplus a_{jj}^{\frac{l+1-k}{k}} a_{ii}^{\frac{k-l-1}{k}} a_{tt} \\
&\geq a_{tt},
\end{aligned}$$

where the first inequality holds by two uses of Lemma 12.15 (note that $k-l-1 \geq 0$).

We can now show (12.28) holds, as follows.

$$\begin{aligned}
& a_{ij} \left(a_{ii} \oplus a_{jj}\right)^{\frac{l+1-k}{k}} \geq a_{it} \left(a_{ii} \oplus a_{tt}\right)^{\frac{l-k}{k}} a_{tj} \left(a_{tt} \oplus a_{jj}\right)^{\frac{1-k}{k}} \\
\Leftrightarrow & a_{ij} \overbrace{\left(a_{ii} \oplus a_{jj}\right)^{\frac{l+1-k}{k}} \left(a_{ii} \oplus a_{tt}\right)^{\frac{k-l}{k}} \left(a_{tt} \oplus a_{jj}\right)^{\frac{k-1}{k}}}^{\geq a_{tt} \text{ by } (12.29)} \geq a_{it} a_{tj} \\
\Leftarrow & a_{ij} a_{tt} \geq a_{it} a_{tj},
\end{aligned}$$

which holds.

Therefore

$$b_{ij}^{(l+1)} = a_{ij} \left(a_{ii} \oplus a_{jj}\right)^{\frac{l+1-k}{k}} .$$

We conclude $P\left(l+1\right)$ holds and the result follows by induction. □

Corollary 12.16.

$$B^k = A.$$

Proof.

$$B^k = \left(b_{ij}^{(k)}\right) = \left(a_{ij}\left(a_{ii} \oplus a_{jj}\right)^{\frac{k-k}{k}}\right) = (a_{ij}) = A.$$

□

Example 12.17. $A = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 2 & 2 \\ 0 & 0 & 1 \end{pmatrix}$ *satisfies the condition of Theorem 12.14. For* $k = 3$ *we have*

$$B = (b_{ij}) = \left(a_{ij}\left(a_{ii} \oplus a_{jj}\right)^{-\frac{2}{3}}\right) = \begin{pmatrix} \frac{1}{3} & -\frac{4}{3} & \frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ -\frac{2}{3} & -\frac{4}{3} & \frac{1}{3} \end{pmatrix}.$$

We then see that

$$B^2 = \begin{pmatrix} \frac{1}{3} & -\frac{4}{3} & \frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ -\frac{2}{3} & -\frac{4}{3} & \frac{1}{3} \end{pmatrix}\begin{pmatrix} \frac{1}{3} & -\frac{4}{3} & \frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ -\frac{2}{3} & -\frac{4}{3} & \frac{1}{3} \end{pmatrix}$$
$$= \begin{pmatrix} \frac{2}{3} & -\frac{2}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{4}{3} & \frac{4}{3} \\ -\frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \end{pmatrix}$$

and

$$B^3 = \begin{pmatrix} \frac{2}{3} & -\frac{2}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{4}{3} & \frac{4}{3} \\ -\frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \end{pmatrix} \begin{pmatrix} \frac{1}{3} & -\frac{4}{3} & \frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ -\frac{2}{3} & -\frac{4}{3} & \frac{1}{3} \end{pmatrix}$$
$$= \begin{pmatrix} 1 & 0 & 1 \\ 1 & 2 & 2 \\ 0 & 0 & 1 \end{pmatrix}$$
$$= A,$$

as expected.

Example 12.18. $A = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 2 & 2 \\ 1 & 1 & 1 \end{pmatrix}$ *does not satisfy the conditions of Theorem 12.14 since* $a_{21}a_{33} < a_{23}a_{31}$.

If we define

$$B = (b_{ij})$$
$$= \left(a_{ij}\left(a_{ii} \oplus a_{jj}\right)^{-\frac{2}{3}}\right)$$
$$= \begin{pmatrix} \frac{1}{3} & -\frac{4}{3} & \frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{1}{3} & -\frac{1}{3} & \frac{1}{3} \end{pmatrix},$$

then

$$B^2 = \begin{pmatrix} \frac{1}{3} & -\frac{4}{3} & \frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{1}{3} & -\frac{1}{3} & \frac{1}{3} \end{pmatrix} \begin{pmatrix} \frac{1}{3} & -\frac{4}{3} & \frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{1}{3} & -\frac{1}{3} & \frac{1}{3} \end{pmatrix}$$

$$= \begin{pmatrix} \frac{2}{3} & 0 & \frac{2}{3} \\ 1 & \frac{4}{3} & \frac{4}{3} \\ \frac{2}{3} & \frac{1}{3} & \frac{2}{3} \end{pmatrix}$$

and

$$B^3 = \begin{pmatrix} \frac{2}{3} & 0 & \frac{2}{3} \\ 1 & \frac{4}{3} & \frac{4}{3} \\ \frac{2}{3} & \frac{1}{3} & \frac{2}{3} \end{pmatrix} \begin{pmatrix} \frac{1}{3} & -\frac{4}{3} & \frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{1}{3} & -\frac{1}{3} & \frac{1}{3} \end{pmatrix}$$

$$= \begin{pmatrix} 1 & \frac{2}{3} & 1 \\ \frac{5}{3} & 2 & 2 \\ 1 & 1 & 1 \end{pmatrix}$$

$$\neq A.$$

Interestingly, the matrix B^3 is "off" in exactly two positions. We can also check that the matrix A violates exactly two of the conditions of Theorem 12.14. In particular, $a_{21}a_{33} < a_{23}a_{31}$ and $a_{12}a_{33} < a_{13}a_{32}$. All other conditions are satisfied. We have the following remark.

Remark 12.19. *In the proof of Theorem 12.14, we fix i, j and then prove $b_{ij}^{(l)}$ is as expected. Not all conditions of the theorem are used however. To prove $b_{ij}^{(l)} = a_{ij} \left(a_{ii} \oplus a_{jj}\right)^{\frac{l-k}{k}}$*

we require

$$(\forall t \neq i,j)\, a_{ij}a_{tt} \geq a_{it}a_{tj}. \tag{12.30}$$

We call (12.30) the *root constraints* for (i,j). We summarise with the following corollary.

Corollary 12.20. *Let $A \in \mathbb{R}^{n\times n}$ and $k \geq 2$ an integer. Define the matrix*

$$B = (b_{ij}) = \left(a_{ij}\left(a_{ii} \oplus a_{jj}\right)^{\frac{1-k}{k}}\right).$$

If

$$(\forall t \neq i,j)\, a_{ij}a_{tt} \geq a_{it}a_{tj},$$

then

$$(\forall l \geq 1)\, b_{ij}^{(l)} = a_{ij}\left(a_{ii} \oplus a_{jj}\right)^{\frac{l-k}{k}}.$$

It follows that if a matrix $A \in \mathbb{R}^{n\times n}$ satisfies the root constraints for (i,j) for most pairs i,j, then B may serve as a good approximation (in some sense) to a kth root of A.

12.4 Summary

We defined kth roots (for integer k) for 2×2, finite matrices (when such roots exist). We also explicitly described when such roots do and do not exist. We were able to generalise the formula for the kth root of a matrix to the general $n \times n$ case, provided the matrix satisfied some conditions on some of its 2×2 sub-matrices. Corollary 12.20 motivated a question about approximations of matrix roots - which may be useful considering finding exact roots for $n\times n$ matrices is NP-complete in the Boolean case.

13. Thesis conclusions and further research

We give here a brief summary of the most important results of the thesis and some questions arising.

In chapter 3 we studied the finite, $m \times n$ parametrised system $Ax = b(\alpha)$. By describing the full solution set we developed a tool for finding non-trivial solutions to other systems. In particular, in chapter 7 we showed that the two-sided system $Ax = By$, for $A \in \mathbb{R}^{m\times n}$ and $B \in \mathbb{R}^{m\times 2}$, can be viewed as a sequence of one-sided parametrised systems. In chapter 10 we showed that GEP can be viewed as a one-sided parametrised system if B (say) is an outer-product.
Question: Is it possible to increase the number of parameters in the right-hand side and still describe the whole set of solutions?

In chapter 4 we presented a strongly polynomial method for solving the two-sided system of inequalities $Ax \leq Bx$ for $A \in \mathbb{R}^{m\times n}$ and $B \in \overline{\mathbb{R}}^{m\times n}$ where B has exactly one finite entry per row. This is in contrast to the problem in which we replace " $\leq$ " with " $=$ ". In making this small change we move from being able to describe the set of generators of all finite solutions in strongly polynomial time to having no obvious way of finding even one solution in strongly polynomial time. This problem is suggesting, then, that two-sided systems of inequalities are, probably, more tractable than two-sided systems of equations.

Question: Is it possible to solve such systems in strongly polynomial time if B has exactly two finite entries per row?

In chapter 5 we have

Question: Can we find a solution to the two-sided system $Ax = Bx$ in polynomial time when A and B have exactly two finite entries per row but not necessarily in the same position?

In chapter 6 we saw that the Cancellation Rule can be a powerful tool. Using it carefully allowed us to give explicit solutions for the two-dimensional two-sided system $Ax = Bx$ for $A, B \in \mathbb{R}^{m\times 2}$ and the two-dimensional generalised eigenproblem $Ax = \lambda Bx$ for $A, B \in \mathbb{R}^{m\times 2}$.

Question: Can we give similar explicit solutions for the three-dimensional counterparts?

In chapter 7 we have

Question: Can we increase the number of columns in B to three and still solve the two-sided system $Ax = By$?

In chapter 8 we showed a potentially useful connection between square, finite two-sided systems $Ax = Bx$ and the assignment problem. This connection can help to find "partial solutions", or even full solutions in some cases, in strongly polynomial time.

Question: Can we identify minimally active systems easily?

Question: Can we identify essential systems easily?

Question: Can we generalise the results to non-square systems?

Question: Can we say anything useful for non-finite, square matrices when there is no finite permutation in C?

In chapter 9

Question: Can we perform a similar analysis for other structured matrices? For example, Toeplitz or Hankel matrices.

In chapter 11 we found a new connection between the generalised eigenproblem $Ax =$

λBx and Game Theory when the matrix $C = A - B$ is finite, symmetric and has a saddle point. We showed that the value of this saddle point is the unique candidate for the generalised eigenvalue and gave some necessary conditions for checking if it is indeed an eigenvalue. In the 3×3 case we were able to give necessary and sufficient conditions.

Question: Can we find a necessary and sufficient condition for when the unique candidate for the generalised eigenvalue is indeed an eigenvalue?

Question: In the positive case, can we find a corresponding eigenvector?

Finally, in chapter 12, we defined kth roots (for integer k) for 2×2, finite matrices (when such roots exist). We also explicitly described when such roots do and do not exist. We were able to generalise the formula for the kth root of a matrix to the general $n \times n$ case provided the matrix satisfied some *Monge-type* conditions on some of its 2×2 sub-matrices.

Question: Is the following statement true: "The matrix $A \in \mathbb{R}^{n \times n}$ has a kth root (some fixed $k \geq 2$) if and only if $B^k = A$ (where B is as given in Theorem (12.14))"?

Question: Is B a good approximation to the kth root when A does not satisfy the conditions of Theorem 12.14?

13. Bibliography

[1] B. Addad, S. Amari, J. J. Lesage, *Networked conflicting timed event graphs representation in (Max,+) algebra*, Discrete Event Dynamic Systems, 22 (4), 2012 Automatic Control, AC-30: pp. 210-220, 1985.

[2] M. Akian, R. B. Bapat and S. Gaubert, *Max-plus algebras*, in Handbook of Linear Algebra (Discrete Mathematics and Its Applications), L. Hogben, ed.,Vol. 39, Chapman & Hall/CRC, pp. 25/1-25/17, 2006.

[3] M. Akian, S. Gaubert and A. Guterman, *Tropical polyhedra are equivalent to mean payoff games*, International of Algebra and Computation, 22 (1):125001 (43 pages), 2012.

[4] M. Akian, S. Gaubert, A. Marchesini, *Tropical bounds for eigenvalues of matrices*, Linear Algebra and its Applications, 446: pp. 281-303, 2014.

[5] X. Allamigeon, P. Benchimol, S. Gaubert and M. Joswig, *Combinatorial simplex algorithms can solve mean payoff games*, SIAM Journal on Optimization, 24 (4), pp. 2096-2117, 2014.

[6] X. Allamigeon, P. Benchimol, and S. Gaubert, *The tropical shadow-vertex algorithm solves mean payoff games in polynomial time on average*, Proceedings of the 41st International Colloquium on Automata, Languages, and Programming (ICALP), number 8572 in Lecture Notes in Computer Science, pp. 89100. Springer, 2014.

[7] X. Allamigeon, S. Gaubert and E. Goubault, *Computing the vertices of tropical polyhedra using directed hypergraphs*, Discrete Comput. Geom. 49 (2), pp. 247-279, 2013.

[8] F. L. Baccelli, G. Cohen, G. J. Olsder and J. P. Quadrat, *Synchronization and Linearity*, John Wiley, Chichester, New York, 1992.

[9] R. B. Bapat and T. E. S. Raghaven, *Nonnegative matrices and applications*, Cambridge: Cambridge University Press, 1997.

[10] R. B. Bapat, D. Stanford and P. van den Driessche, *The eigenproblem in max-algebra*, (DMS-631-IR). University of Victoria, British Colombia, 1993.

[11] M. Bezem, R. Nieuwenhuis and E. Rodríguez-Carbonell, *Hard problems in max-algebra, control theory, hypergraphs and other areas*. Information Processing Letters 110 (4): pp. 133-138, 2010.

[12] P. A. Binding and H. Volkmer, *A generalized eigenvalue problem in the max algebra*. Linear Algebra and Its Applications 422: pp. 360-371, 2007.

[13] R. Burkard, M. Dell'Amico and S. Martello, *Assignment Problems*, Society for Industrial and Applied Mathematics, Philadelphia, 2012.

[14] P. Butkovič, R. A. Cuninghame-Green and S. Gaubert, *Reducible spectral theory with applications to the robustness of matrices in max-algebra*, SIAM J. Matrix Anal. Appl. 31 (3), pp. 1412-1431, 2009.

[15] P. Butkovič and S. Gaubert, *Sign-non-singular matrices and matrices with unbalanced determinant in symmetrised semirings*, Linear Algebra and Appl., 301, pp. 195-201, 1999.

[16] P. Butkovič and G. Hegedüs, *An elimination method for finding all solutions of the system of linear equations over an extremal algebra*, Ekon. – Mat. Obzor 20, pp. 203-214, 1984.

[17] P. Butkovič and D. Jones, *On special cases of the generalized max-Plus eigenproblem*, SIAM J. Matrix Anal. Appl. Vol. 37, No. 3, pp. 1002-1021, 2016.

[18] P. Butkovič and H. Schneider, *Applications of max-algebra to diagonal scaling of matrices*, Electronic Journal of Linear Algebra, 13, pp. 262-273, 2005.

[19] P. Butkovič, *On certain properties of the systems of linear extremal equations*, Ekon. – Mat. Obzor 14, pp. 72-78, 1978.

[20] P. Butkovič, *Solution of systems of linear extremal equations*, Ekon. – Mat. Obzor 17, pp. 402-416, 1981.

[21] P. Butkovič, *On properties of solution sets of extremal linear programs*, Ann. Discrete Math. 19, pp. 41-54, 1984.

[22] P. Butkovič, *Necessary solvability conditions of systems of linear extremal equations*, Discrete Applied Mathematics 10, pp. 19-26, 1985.

[23] P. Butkovič, *Strong regularity of matrices - a survey of results*, Discrete Applied Mathematics, 48, pp. 45-68, 1994.

[24] P. Butkovič, *Max-algebra: the linear algebra of combinatorics?*, Linear Algebra and Its Applications, 367, pp. 313-335, 2003.

[25] P. Butkovič, *Max-linear Systems: Theory and Algorithms*, Springer Monographs in Mathematics, Springer-Verlag, London, 2010.

[26] B. A. Carré, *An algebra for network routing problems*, Journal of the Institute of Mathematics and Its Applications, 7 (3), pp. 273-294, 1971.

[27] W. Chen, X. Qi and S. Deng, *The eigenproblem and period analysis of the discrete event systems*, Systems Science and Mathematical Sciences, 3 (3), pp. 243-260, 1990.

[28] G. Cohen, D. Dubois, J. P. Quadrat, M. Viot, *A linear system-theoretic view of discrete event processes and its use for performance evaluation in manufacturing*, IEEE Transactions on Automatic Control, AC-30: pp. 210-220, 1985.

[29] G. Cohen, S. Gaubert, E. Mancinelli, J.P. Quadrat and E. Rofman, *On traffic light control of regular towns*, Mathematicae Notae, Boletin del Instituto de Matematica "Beppo Levi", Ano XLIII, pp. 51-62, Universidad Nacional de Rosario Editora, Rosario, Argentina, 2005.

[30] G. Cohen, S. Gaubert and J. P. Quadrat, *Max-plus algebra and system theory: where we are and where to go now*, Annual Reviews in Control, vol. 23, pp 207-219, 1999.

[31] R.A. Cuninghame-Green and P. Butkovič, *The equation Ax=By over (max, +)*, Theoretical Computer Science 293, pp. 3-12, 2003.

[32] R. A. Cuninghame-Green and P. Butkovič, *Generalised eigenproblem in max algebra*, IEEE Xplore, Discrete Event Systems (2008), 9th International Workshop WODES, pp. 236-241, 2008.

[33] R. A. Cuninghame-Green, *Process synchronisation in a steelworks - a problem of feasibility*, in Proc 2nd Int Conf on Operational Research, Banbury and Maitland, eds., English University Press, pp. 323-328, 1960.

[34] R. A. Cuninghame-Green, *Describing industrial processes with interference and approximating their steady-state behaviour*, Oper. Res. Quart. 13, pp. 95-100, 1962.

[35] R. A. Cuninghame-Green, *Minimax Algebra*, Lecture Notes in Economics and Mathematical Systems 166, Berlin, Springer, pp. 251-267, 1979.

[36] R. A. Cuninghame-Green, *Minimax algebra and applications*, Fuzzy Sets and Systems, 41, pp. 251-267, 1991.

[37] R. A. Cuninghame-Green, *Advances in imaging and electron physics*, Minimax algebra and applications (Vol. 90), New York: Academic Press, 1995.

[38] L. Elsner and P. van den Driessche, *On the power method in max-algebra*, Linear Algebra and Its Applications, 302/303, pp.17-32, 1999.

[39] G. M. Engel and H. Schneider, *Cyclic and diagonal products on a matrix*, Linear Algebra and Its Applications, 7, pp. 301-335, 1973.

[40] G. M. Engel and H. Schneider, *Diagonal similarity and equivalence for matrices over groups with 0*, Czechoslovak Mathematical Journal, 25, pp. 389-403, 1975.

[41] S. Gaubert, L. Grigori and M. Sharify, *Solution of the optimal assignment problem by diagonal scaling algorithms*, arXiv:1104.3830, 2011.

[42] S. Gaubert, R. D. Katz and S. Sergeev, *Tropical linear-fractional programming and parametric mean-payoff games*, Journal of Symbolic Computation 47, pp. 1447-1478, 2012.

[43] S. Gaubert and M. Plus, *Methods and applications of max-plus linear algebra*, Proceedings of the 14th Annual Symposium on Theoretical Aspects of Computer Science Lubeck, Germany February 27 - March 1, 1997 (STACS'97), pp. 261-282, LNCS 1200, Springer, 1997.

[44] S. Gaubert and S. Sergeev, *The level set method for the two-sided max-plus eigenproblem*, Discrete Event Dyn Syst, 23: pp. 105-134, 2013.

[45] S. Gaubert, *Théorie des systèmes linéaires dans les dioïdes*, Thèse, École des Mines de Paris, 1992 (in French).

[46] S. Gaubert, *Nonlinear Perron-Frobenius theory and discrete event systems*, Actes du colloque Modélisation de Systmes Ractifs (MSR'05), JESA vol. 39, pp. 175-190, 2005.

[47] B. Giffler, *Scheduling general production systems using schedule algebra*, Naval Research Logistics Quarterly, 10, pp. 237-255, 1963.

[48] B. Giffler, *Schedule algebra: a progress report*, Naval Research Logistics Quarterly, 15, pp. 255-280, 1968.

[49] M. Gondran and M. Minoux, *Valeurs propres et vecteur propres dans les dioïdes et leur interprétation en théorie des graphes*, Bulletin de la Direction Études et Recherches, Série C, Mathematiques et Informatiques 2, pp. 25-41, 1977 (in French).

[50] M. Gondran and M. Minoux, *L'Indépendance linéaire dans les dioïdes*, Bulletin de la Direction Études et Recherches. EDF, Série C, 1, , pp. 67-90, 1978 (in French).

[51] M. Gondran and M. Minoux, *Linear algebra of dioïdes, a survey of recent results*, Annals of Discrete Mathematics, 19, pp. 147-163, 1984.

[52] M. Gondran, *Path algebra and algorithms. In B. Roy (Ed.)*, NATO advanced study inst. ser., ser. C: Math. and phys. sci.: Vol. 19. Combinatorial programming: methods and applications (pp. 137-148), Reidel: Dordrecht. (Proc NATO Advanced Study Inst, Versailles, 1974).

[53] B. Heidergott, G. J. Olsder and J. van der Woude, *Max Plus at Work: Modeling and Analysis of Synchronized Systems*, A Course on Max-Plus Algebra, Princeton University Press, 2005.

[54] P. Klemperer, *The product-mix auction: a new auction design for differentiated goods*, Journal of the European Economic Association 8, pp. 526-536, 2010.

[55] D. König, *Gráfok és mátrixok*, Matematikai és Fizikai Lapok, 38, pp. 116-119, 1931 (in Hungarian).

[56] H. W. Kuhn, *The Hungarian method for the Assignment Problem*, Naval Reserach Logistics Quarterly, 3, pp. 253-258, 1956.

[57] M. Kutz, *The complexity of Boolean matrix root computation*, Theoretical Computer Science 325, pp. 373-390, 2004.

[58] V. Noferini, M. Sharify and F. Tisseur, *Tropical roots as approximations to eigenvalues of matrix polynomials*, SIAM. J. Matrix Anal. and Appl., 36(1), pp. 138-157, 2015.

[59] G. Owen, *Game Theory*, 3rd ed., Academic Press, 1995.

[60] H. Schneider and M. H. Schneider, *Max-balancing weighted directed graphs and matrix scaling*, Mathematics of Operations Research, 16(1), pp. 208-222, 1991.

[61] H. Schneider, *The influence of the marked reduced graph of a nonnegative matrix on the Jordan form and on related properties: a survey*, Linear Algebra and Its Applications, pp. 161-189, 1988.

[62] S. Sergeev, H. Schneider and P. Butkovič, *On visualisation scaling, subeigenvectors and Kleene stars in max algebra*, Linear Algebra and its Applications 431, pp. 2395-2406, 2009.

[63] S. Sergeev, *On the system $Ax = \lambda Bx$ in max algebra: every system of intervals is a spectrum*, Kybernetika, Volume 47, no. 5, pp. 715-721, 2011.

[64] N. N. Vorobyov, *Extremal algebra of positive matrices*, Elektronische Datenverarbeitung und Kybernetik, 3, pp. 39-71, 1967 (in Russian).

[65] N. N. Vorobyov, *Extremal algebra of non-negative matrices*, Elektronisch Informationsverarbeitung und Kybernetik, 6, pp. 302-312, 1970 (in Russian).

[66] E. A. Walkup and G. Boriello, *A general linear max-plus solution technique*, In: Gunawardena J. (ed) Idempotency, Cambridge, pp. 406-415, 1988.

[67] K. Zimmermann, *Extremální algebra*, Výzkumná publikace Ekonomicko - matematické laboratoře při Ekonomickém ústavě ČSAV, 46, Praha, 1976 (in Czech).

[68] U. Zimmermann, *Linear and combinatorial optimisation in ordered algebraic structures*, Annals of Discrete Mathematics: Vol. 10, Amsterdam, North Holland, 1981.