LINEAR ALGEBRA AND ITS APPLICATIONS

ELSEVIER Linear Algebra and its Applications 297 (1999) 157–175
www.elsevier.com/locate/laa

Majorization polytopes

Geir Dahl*

Department of Mathematics and Department of Informatics, University of Oslo, P.O.Box 1080, Blindern 0316 Oslo, Norway

Received 3 June 1999; accepted 27 June 1999

Submitted by R.A. Brualdi

Abstract

We study polytopes related to the concept of *matrix majorization*: for two real matrices $\mathbf{A}$ and $\mathbf{B}$ having m rows we say that $\mathbf{A}$ majorizes $\mathbf{B}$ if there is a row-stochastic matrix $\mathbf{X}$ with $\mathbf{AX} = \mathbf{B}$. In that case we write $\mathbf{A} \succ \mathbf{B}$ and the associated majorization polytope $\mathscr{M}(\mathbf{A} \succ \mathbf{B})$ is the set of row stochastic matrices $\mathbf{X}$ such that $\mathbf{AX} = \mathbf{B}$. We investigate some properties of $\mathscr{M}(\mathbf{A} \succ \mathbf{B})$ and obtain e.g., generalizations of some results known for vector majorization. Relations to transportation polytopes and network flow theory are discussed. A complete description of the vertices of majorization polytopes is found for some special cases. © 1999 Elsevier Science Inc. All rights reserved.

Keywords: Majorization; Row-stochastic matrices; Polytopes

1. Introduction

Let $\mathscr{M}_{n,p}$ denote the set of $n \times p$ row-stochastic matrices, i.e., $\mathbf{X} = [x_{i,j}] \in \mathscr{M}_{n,p}$ means that $x_{i,j} \geqslant 0$ for all i, j and $\sum_{j=1}^{p} x_{i,j} = 1$ for all i. We recall the notion of *matrix majorization* as defined in [5]. Let $\mathbf{A}$ and $\mathbf{B}$ be two real matrices with m rows, say $\mathbf{A} \in \mathbb{R}^{m,n}$ and $\mathbf{B} \in \mathbb{R}^{m,p}$. We say that $\mathbf{A}$ *majorizes* $\mathbf{B}$, and write $\mathbf{A} \succ \mathbf{B}$ (or $\mathbf{B} \prec \mathbf{A}$), provided that there exists a row-stochastic matrix $\mathbf{X}$ such that

$$\mathbf{AX} = \mathbf{B}.$$

*Tel.: +47-22-85-2425; fax: +47-22-85-2401.

E-mail address: geird@ifi.uio.no (G. Dahl)

PII: S 0 0 2 4 - 3 7 9 5 (9 9) 0 0 1 4 8 - 2

Note that the number of columns in the two matrices $\mathbf{A}$ and $\mathbf{B}$ may be different. Matrix majorization is a preorder on the set of real matrices with m rows. For different properties of $\succ$ and characterizations of $\mathbf{A} \succ \mathbf{B}$, see [5].

Corresponding to $\mathbf{A} \in \mathbb{R}^{m,n}$ and $\mathbf{B} \in \mathbb{R}^{m,p}$ we define the *majorization polytope*

$$\mathscr{M}(\mathbf{A} \succ \mathbf{B}) = \{\mathbf{X} \in \mathscr{M}_{n,p} \colon \mathbf{AX} = \mathbf{B}\}.$$

This set is nonempty iff $\mathbf{A} \succ \mathbf{B}$, and in that case, it is a bounded polyhedron, i.e., a polytope in the vector space $\mathbb{R}^{n,p}$. We let $\mathbf{e}$ denote a (column) vector of all ones (properly dimensioned). Note that if $\mathbf{A} \succ \mathbf{B}$ then $\mathbf{Ae} = \mathbf{Be}$ (as $\mathbf{Be} = \mathbf{AXe} = \mathbf{Ae}$).

For other (related) majorization notions for matrices see e.g. [1,7,9] (further references are in [5]). Majorization in connection with measure families was studied in e.g. [10].

Matrix majorization generalizes the classical concept of majorization between vectors. Recall that if $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ one says that $\mathbf{a}$ *majorizes* $\mathbf{b}$, denoted by $\mathbf{a} \succ \mathbf{b}$, provided that $\sum_{j=1}^{k} a_{[j]} \geqslant \sum_{j=1}^{k} b_{[j]}$ for $k = 1, \ldots, n-1$ and $\sum_{j=1}^{n} a_j = \sum_{j=1}^{n} b_j$. (Here $a_{[j]}$ denotes the jth largest number among the components of $\mathbf{a}$.) It now follows from our definition above that the matrix majorization

$$\begin{bmatrix} 1 & \ldots & 1 \\ a_1 & \ldots & a_n \end{bmatrix} \succ \begin{bmatrix} 1 & \ldots & 1 \\ b_1 & \ldots & b_n \end{bmatrix}$$

holds if and only if there is a doubly stochastic matrix $\mathbf{X} \in \mathbb{R}^{n,n}$ such that $\mathbf{a}^{\mathrm{T}}\mathbf{X} = \mathbf{b}^{\mathrm{T}}$. But this is equivalent to $\mathbf{a} \succ \mathbf{b}$ according to a well-known theorem of Hardy-Littlewood and Pólya (see [9]).

The goal of this paper is to investigate the majorization polytope $\mathscr{M}(\mathbf{A} \succ \mathbf{B})$ under certain assumptions on the matrices involved.

For vector majorization, say $\mathbf{a} \succ \mathbf{b}$, the majorization polytope $\Omega(\mathbf{a} \succ \mathbf{b})$ consisting of all doubly stochastic matrices $\mathbf{S}$ satisfying $\mathbf{Sa} = \mathbf{b}$ was studied in [2] and different combinatorial properties of this polytope were found. In particular the support matrix of the majorization was determined. A related study for majorization polytopes in connection with *multivariate majorization* is found in [3]. In that paper (see also [9]) one says that an $m \times n$ matrix $\mathbf{A}$ multivariate majorizes another $m \times n$ matrix $\mathbf{B}$, and write $\mathbf{A} \succ_{\mathrm{d}} \mathbf{B}$, provided that there is a *doubly* stochastic matrix $\mathbf{X}$ such that $\mathbf{AX} = \mathbf{B}$. Matrix majorization generalizes this notion as we have that

$$\begin{bmatrix} \mathbf{e}^{\mathrm{T}} \\ \mathbf{A} \end{bmatrix} \succ \begin{bmatrix} \mathbf{e}^{\mathrm{T}} \\ \mathbf{B} \end{bmatrix} \quad \Leftrightarrow \quad \mathbf{A} \succ_{\mathrm{d}} \mathbf{B}.$$

When $\mathbf{A}$ and $\mathbf{B}$ are $(0, 1)$-matrices with exactly one 1 in each row (or exactly two ones in each row) a complete description of all vertices of $\mathscr{M}(\mathbf{A} \succ \mathbf{B})$ was given in [5].

Some of our notation is explained next. $\mathbb{R}^{m,n}$ is the vector space of real $m \times n$ matrices. Let $\mathbf{A} \in \mathbb{R}^{m,n}$. Then the jth column vector of $\mathbf{A}$ is denoted by $\mathbf{a}^j$ and the ith row vector is denoted by $\mathbf{a}_i$. A matrix or vector with all components being zero is denoted by $\mathbf{0}$. If $S \subseteq \mathbb{R}^n$ the convex hull of S is denoted by $\mathrm{conv}(S)$.

2. Relation to transportation polytopes

Let $\mathbf{a}^{\mathrm{T}} = [a_1, \ldots, a_n]$ and $\mathbf{b}^{\mathrm{T}} = [b_1, \ldots, b_p]$ be vectors with nonnegative components and satisfying $\sum_{j=1}^{n} a_j = \sum_{j=1}^{p} b_j$. Define

$$\mathscr{T}(\mathbf{a}, \mathbf{b}) = \{\mathbf{Y} \in \mathbb{R}^{n,p}\colon \mathbf{Ye} = \mathbf{a},\ \mathbf{e}^{\mathrm{T}}\mathbf{Y} = \mathbf{b}^{\mathrm{T}},\ \mathbf{Y} \geqslant \mathbf{0}\}.$$

The polytope $\mathscr{T}(\mathbf{a}, \mathbf{b})$ is the well-known *transportation polytope* which arises in the transportation problem in linear programming. It is widely studied, see e.g. [11], and the 1-skeleton of this polytope is known, and its vertices correspond to spanning trees in the complete bipartite graph $K_{n,p}$. $\mathscr{T}(\mathbf{a}, \mathbf{b})$ is nonempty as $\sum_j a_j = \sum_j b_j$.

When $\mathbf{z} \in \mathbb{R}^n$ we let $\mathbf{D}(\mathbf{z}) \in \mathbb{R}^{n,n}$ denote the diagonal matrix with $d_{j,j} = z_j$ for $j = 1, \ldots, n$. Majorization polytopes are related to transportation polytopes as given in the following proposition. A *positive* matrix is a matrix with only positive entries.

Proposition 2.1. *Let* $\mathbf{A} \in \mathbb{R}^{m,n}$ *be positive and* $\mathbf{B} \in \mathbb{R}^{m,p}$ *nonnegative. Then*

$$\mathscr{M}(\mathbf{A} \succ \mathbf{B}) = \bigcap_{i=1}^{m} \mathbf{D}(\mathbf{a}_i)^{-1} \cdot \mathscr{T}(\mathbf{a}_i, \mathbf{b}_i).$$

$(\mathbf{D}(\mathbf{a}_i)^{-1} \cdot \mathscr{T}(\mathbf{a}_i, \mathbf{b}_i)$ *consists of the matrices* $\mathbf{D}(\mathbf{a}_i)^{-1}\mathbf{Y}$ *where* $\mathbf{Y} \in \mathscr{T}(\mathbf{a}_i, \mathbf{b}_i))$.

Proof. Assume that $\mathbf{X} \in \mathscr{M}(\mathbf{A} \succ \mathbf{B})$ so $\mathbf{X} \geqslant \mathbf{0}$, $\mathbf{Xe} = \mathbf{e}$ and $\mathbf{a}_i^{\mathrm{T}}\mathbf{X} = \mathbf{b}_i^{\mathrm{T}}$ for $i \leqslant m$. Consider a fixed $i \leqslant m$ and define $\mathbf{Y} = D(\mathbf{a}_i)\mathbf{X}$. Note that the diagonal matrix $D(\mathbf{a}_i)$ is positive and nonsingular, by assumption. Therefore $\mathbf{Y} \geqslant \mathbf{0}$, $\mathbf{Ye} = D(\mathbf{a}_i)\mathbf{X}\mathbf{e} = D(\mathbf{a}_i)\mathbf{e} = \mathbf{a}_i$ and $\mathbf{b}_i^{\mathrm{T}} = \mathbf{a}_i^{\mathrm{T}}\mathbf{X} = \mathbf{a}_i^{\mathrm{T}} D(\mathbf{a}_i)^{-1}\mathbf{Y} = \mathbf{e}^{\mathrm{T}}\mathbf{Y}$. This means that $\mathbf{Y} \in \mathscr{T}(\mathbf{a}, \mathbf{b})$ and therefore $\mathbf{X} \in D(\mathbf{a}_i)^{-1} \cdot \mathscr{T}(\mathbf{a}, \mathbf{b})$. Since this holds for every $i \leqslant m$ we conclude that $\mathscr{M}(\mathbf{A} \succ \mathbf{B}) \subseteq \bigcap_{i=1}^{m} \mathbf{D}(\mathbf{a}_i)^{-1} \cdot \mathscr{T}(\mathbf{a}_i, \mathbf{b}_i)$. The converse inclusion is shown similarly. □

Thus the majorization polytope $\mathscr{M}(\mathbf{A} \succ \mathbf{B})$ is the intersection of m "scaled" transportation polytopes. A similar correspondence exists in the general case where $\mathbf{A}$ is nonnegative and may contain zeros, see [5].

In general $\mathscr{M}(\mathbf{A} \succ \mathbf{B})$ may be very complex, but when $\mathbf{A}$ has a certain property the relation to transportation polytopes becomes very useful. Let $\mathrm{supp}(\mathbf{x})$ denote the support of a vector $\mathbf{x} \in \mathbb{R}^n$, i.e., $\mathrm{supp}(\mathbf{x}) = \{j \leqslant n\colon x_j \neq 0\}$. A matrix $\mathbf{A} \in \mathscr{M}_{m,n}$ is a *disjoint-row-support matrix* if the supports of its rows are pairwise disjoint. By suitable line permutations such a matrix $\mathbf{A}$ can be brought to the form

$$\begin{bmatrix} \bar{\mathbf{a}}_1^{\mathrm{T}} & \mathbf{0} & & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \bar{\mathbf{a}}_2^{\mathrm{T}} & & & \\ & & \ddots & & \\ \mathbf{0} & & & \bar{\mathbf{a}}_m^{\mathrm{T}} & \mathbf{0} \end{bmatrix}, \tag{1}$$

where for $i = 1, \ldots, m$, $\bar{\mathbf{a}}_i$ is a vector in $\mathbb{R}^{n_i}$ with only positive components and $n_0 = n - \sum_{i=1}^{m} n_i$.

Proposition 2.2. *Let* $\mathbf{A} \in \mathscr{M}_{m,n}$ *be a disjoint-row-support matrix as given in* (1), *and let* $\mathbf{B} \in \mathscr{M}_{m,p}$. *Then* $\mathbf{A} \succ \mathbf{B}$ *and* $\mathscr{M}(\mathbf{A} \succ \mathbf{B})$ *consists of the matrices* $\mathbf{X}$ *given by*

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_1 \\ \vdots \\ \mathbf{X}_m \\ \mathbf{X}_0 \end{bmatrix}, \tag{2}$$

where $\mathbf{X}_i = \mathbf{D}(\bar{\mathbf{a}}_i)^{-1}\mathbf{Y}_i$ *and* $\mathbf{Y}_i \in \mathscr{T}(\bar{\mathbf{a}}_i, \mathbf{b}_i)$ *for* $i = 1, \ldots, m$ *and* $\mathbf{X}_0 \in \mathscr{M}_{n_0,p}$.

Proof. Comparing the *i*th row in the matrix equation $\mathbf{AX} = \mathbf{B}$ we get $\bar{\mathbf{a}}_i^{\mathrm{T}}\mathbf{X}_i = \mathbf{b}_i^{\mathrm{T}}$ for $i = 1, \ldots, m$. Thus, arguments as in the proof of Proposition 2.1 give the desired result. □

In the situation given in Proposition 2.2 we also see that the vertices of $\mathscr{M}(\mathbf{A} \succ \mathbf{B})$ are the matrices $\mathbf{X}$ in (2) where $\mathbf{Y}_i$ is a vertex of $\mathscr{T}(\bar{\mathbf{a}}_i, \mathbf{b}_i)$ and $\mathbf{X}_0 \in \mathscr{M}_{n_0,p}$ is integral. As remarked above, such vertices $\mathbf{Y}_i$ correspond to spanning trees in the complete bipartite graph $K_{n_i,p}$.

The result of Proposition 2.2 also applies to an arbitrary disjoint-row-support matrix $\mathbf{A} \in \mathscr{M}_{m,n}$. This is due to the fact that when $\mathbf{P}$ and $\mathbf{Q}$ are permutations matrices then $\mathbf{AX} = \mathbf{B}$ is equivalent to $(\mathbf{PAQ})(\mathbf{Q}^{\mathrm{T}}\mathbf{X}) = \mathbf{PB}$. Here one may choose $\mathbf{P}$ and $\mathbf{Q}$ so that $\mathbf{PAQ}$ has the form (1).

3. Relation to network flow theory

The purpose of this section is demonstrate a relation between majorization polytopes and the theory of network flows.

First we discuss some questions concerning integral matrices in majorization polytopes. An integral row-stochastic matrix is a $(0, 1)$-matrix with exactly one nonzero, a one, in each row. Following [5] we define the *Markotope* $\mathscr{M}(\mathbf{A}; k)$ associated with $\mathbf{A} \in \mathbb{R}^{m,n}$ and a positive integer k by

$$\mathscr{M}(\mathbf{A}; k) = \{\mathbf{AX}: \mathbf{X} \in \mathscr{M}_{n,k}\}.$$

Thus, letting $\mathbf{B} \in \mathbb{R}^{m,p}$ and $k = p$, we see that $\mathbf{A} \succ \mathbf{B}$ if and only if $\mathbf{B} \in \mathscr{M}(\mathbf{A}; p)$. The Markotope $\mathscr{M}(\mathbf{A}; k)$ is a polytope in $\mathbb{R}^{m,k}$ and (see [5]) each vertex of $\mathscr{M}(\mathbf{A}; k)$ may be written

$$\left[\sum_{j \in J_1} \mathbf{a}^j, \ldots, \sum_{j \in J_k} \mathbf{a}^j\right],$$

where $J_1, \ldots, J_k$ is a partition of $\{1, \ldots, n\}$ (some of the sets may be empty in which case the vector sum should be understood as the zero vector). Thus, the vertices are those matrices of the form **AX** where **X** is an integral row-stochastic matrix. From this we directly obtain the following result concerning integral matrices in the majorization polytope.

Proposition 3.1. *Let* $\mathbf{A} \in \mathbb{R}^{m,n}$ *and* $\mathbf{B} \in \mathbb{R}^{m,p}$ *satisfy* $\mathbf{A} \succ \mathbf{B}$. *Then* $\mathscr{M}(\mathbf{A} \succ \mathbf{B})$ *contains an integral matrix if and only if* $\mathbf{B}$ *is a vertex of* $\mathscr{M}(\mathbf{A};\ p)$, *or, equivalently, there is a partition* $J_1, \ldots, J_p$ *of* $\{1, \ldots, n\}$ *such that* $\mathbf{B} = \left[\sum_{j \in J_1} \mathbf{a}^j, \ldots, \sum_{j \in J_p} \mathbf{a}^j\right]$.

We now turn to network flows. Let $G = (V, E)$ be a directed graph with node set V and arc set E. Let m and n denote the number of nodes and arcs, respectively. The set of arcs with terminal end node (head) v is denoted by $\delta^-(v)$ and the set of arcs with initial endnode (tail) v is denoted by $\delta^+(v)$. A vector $\mathbf{b} \in \mathbb{R}^V$ (or $\mathbb{R}^m$) with $\sum_{v \in V} b_v = 0$ is a called a *demand* (*vector*) and a vector $\mathbf{x} \in \mathbb{R}^E$ satisfying

$$\begin{aligned} &\text{(i)} \quad \sum_{e \in \delta^-(v)} x_e - \sum_{e \in \delta^+(v)} x_e = b_v \quad \text{for all } v \in V; \\ &\text{(ii)} \quad x_e \geqslant 0 \quad \text{for all } e \in E \end{aligned} \tag{3}$$

is called a **b**-*flow*. One can interpret x_e where $e = (v, w)$ as a flow from node v to node w along the arc e and then Eq. (3)(i) says that the net flow into node v equals b_v for each $v \in V$. Let $\mathbf{u} \in \mathbb{R}^E$be a nonnegative vector, called a *capacity*, and let $\mathbf{b}^1, \ldots, \mathbf{b}^s$ be different demands. If $\mathbf{x}^j$ is a $\mathbf{b}^j$-flow for $j = 1, \ldots, s$ and

$$\sum_{j=1}^{s} \mathbf{x}^j \leqslant \mathbf{u}$$

we call $(\mathbf{x}^1, \ldots, \mathbf{x}^p)$ a *multicommodity flow* w.r.t. $(\mathbf{b}^1, \ldots, \mathbf{b}^s; \mathbf{u})$. These constraints say that the total flow (summed over all commodities) in each arc e does not exceed the capacity u_e.

Let $\mathbf{A} \in \mathbb{R}^{m,n}$ be the node-arc incidence matrix of the digraph G. Thus, $\mathbf{A}$ is the $(-1, 0, 1)$-matrix with a row for each $v \in V$ and a column for each arc $e \in E$ and $a_{v,e} = 1$ if $e \in \delta^-(v)$, $a_{v,e} = -1$ if $e \in \delta^+(v)$ and $a_{v,e} = 0$ otherwise. Moreover, let $\mathbf{b}^1, \ldots, \mathbf{b}^{p-1}$ be demand vectors in $\mathbb{R}^V$ and define $\mathbf{b}^p = \sum_{j=1}^{n} \mathbf{a}^j - \sum_{j=1}^{p-1} \mathbf{b}^j$. Thus, $\mathbf{Ae} = \mathbf{Be}$ where $\mathbf{B} = [\mathbf{b}^1, \ldots, \mathbf{b}^p]$ (a necessary majorization condition).

The following immediate result connects majorization and flows.

Proposition 3.2. *When* $\mathbf{A}$ *and* $\mathbf{B}$ *are as above, then* $\mathbf{A} \succ \mathbf{B}$ *if and only if there exists a multicommodity flow* $(\mathbf{x}^1, \ldots, \mathbf{x}^{p-1})$ *w.r.t.* $(\mathbf{b}^1, \ldots, \mathbf{b}^{p-1}; \mathbf{e})$. *Moreover,* $\mathbf{X} \in \mathscr{M}(\mathbf{A} \succ \mathbf{B})$ *if and only if* $\mathbf{X} = [\mathbf{x}^1, \ldots, \mathbf{x}^p]$ *where* $(\mathbf{x}^1, \ldots, \mathbf{x}^{p-1})$ *is a multicommodity flow w.r.t.* $(\mathbf{b}^1, \ldots, \mathbf{b}^{p-1}; \mathbf{e})$ *and* $\mathbf{x}^p = \mathbf{e} - \sum_{j=1}^{p-1} \mathbf{x}^j$.

Proof. $\mathbf{AX} = \mathbf{B}$ means that $\mathbf{Ax}^j = \mathbf{b}^j$ for $j = 1, \ldots, p$. Here the equation $\mathbf{Ax}^p = \mathbf{b}^p$ may be replaced by $\sum_{j=1}^{p-1} \mathbf{x}^j \leqslant \mathbf{e}$ as $\mathbf{Ae} = \mathbf{Be}$. □

Thus, the question of whether the majorization $\mathbf{A} \succ \mathbf{B}$ holds corresponds to the existence question for multicommodity flows, and the majorization polytope is essentially the set of multicommodity flows. Consider the special case of $p = 2$. Then $\mathbf{A} \succ \mathbf{B}$ holds if and only if there is a $\mathbf{b}^1$-flow $\mathbf{x}$ satisfying $\mathbf{x} \leqslant \mathbf{e}$. From network flow theory such a flow exists if and only if $\sum_{v \in V} b_v^1 = 0$ and

$$|\delta^-(S)| \geqslant \sum_{v \in S} b_v^1 \quad \text{for all } S \subset V,$$

where $\delta^-(S)$ is the set of all arcs with head in S and tail outside S. More generally, in the multicommodity case (arbitrary p) conditions assuring the existence of a multicommodity flow w.r.t. $(\mathbf{b}^1, \ldots, \mathbf{b}^{p-1}; \mathbf{u})$ are well known in the network flow literature (these conditions are derived from Farkas' lemma). The computational problem of checking if there exists a multicommodity flow may be solved efficiently by linear programming. This is not so, however, if we ask for an *integral* multicommodity flow. Consider the special case of the situation in Proposition 3.2 where each $\mathbf{b}^j$ for $j = 1, \ldots, p-1$ contains a -1 and a 1 while all other components are zero, say $b_v^j = -1$ if $v = r^j$, $b_v^j = 1$ if $v = s^j$ and $b_v^j = 0$ otherwise. Then, an integral multicommodity flow simply corresponds to arc-disjoint directed paths $Q^1, \ldots, Q^{p-1}$ where, for $j = 1, \ldots, p-1$, Q^j goes from r^j to s^j. The computational problem of checking the existence of such paths (in a given directed graph) is known to be *NP*-complete, even if $p = 3$ (see [6]). This means, confer Propositions 3.1 and 3.2, that even if $\mathbf{A} \succ \mathbf{B}$ and $p = 3$, it is *NP*-hard to decide if $\mathscr{M}(\mathbf{A} \succ \mathbf{B})$ contains an integral matrix.

4. The full row-rank case

Throughout this section we consider a given majorization $\mathbf{A} \succ \mathbf{B}$ where $\mathbf{A} \in \mathbb{R}^{m,n}$, $\mathbf{B} \in \mathbb{R}^{m,p}$ and $\mathbf{A}$ has full row-rank (so $n \geqslant m$).

Note that if $\mathbf{P}$ is an $n \times n$ permutation matrix and $\mathbf{AX} = \mathbf{B}$, then $\mathbf{APP}^{\mathrm{T}}\mathbf{X} = \mathbf{B}$. So, permuting columns of the matrix $\mathbf{A}$ simply corresponds to permuting rows of the matrices in the majorization polytope. Thus, we may assume that $\mathbf{A}$ is partitioned as

$$\mathbf{A} = [\mathbf{A}_1, \mathbf{A}_2], \tag{4}$$

where $\mathbf{A}_1 \in \mathbb{R}^{m,m}$ is nonsingular.

Proposition 4.1. *Assume that* $\mathbf{A}$ *is partitioned as in* (4) *with* $\mathbf{A}_1 \in \mathbb{R}^{m,m}$ *nonsingular. Then*

$$\mathscr{M}(\mathbf{A} \succ \mathbf{B}) = \left\{ \begin{bmatrix} \mathbf{X}_1 \\ \mathbf{Y} \end{bmatrix} : \ \mathbf{X}_1 = \mathbf{A}_1^{-1}(\mathbf{B} - \mathbf{A}_2\mathbf{Y}), \ \mathbf{Y} \in \mathscr{M}^*(\mathbf{A} \succ \mathbf{B}) \right\},$$

where

$$\mathscr{M}^*(\mathbf{A} \succ \mathbf{B}) = \left\{ \mathbf{Y} \in \mathscr{M}_{n-m,p} : \mathbf{A}_1^{-1}\mathbf{A}_2\mathbf{Y} \leqslant \mathbf{A}_1^{-1}\mathbf{B} \right\}.$$

Proof. We partition $\mathbf{X} \in \mathbb{R}^{m,n}$ as

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_1 \\ \mathbf{Y} \end{bmatrix}$$

(where $\mathbf{X}_1 \in \mathbb{R}^{m,m}$ and $\mathbf{Y} \in \mathbb{R}^{n-m,p}$) and see that $\mathbf{AX} = \mathbf{B}$ is equivalent to $\mathbf{A}_1\mathbf{X}_1 + \mathbf{A}_2\mathbf{Y} = \mathbf{B}$. But $\mathbf{A}_1$ is nonsingular so the system becomes $\mathbf{X}_1 = \mathbf{A}_1^{-1}(\mathbf{B} - \mathbf{A}_2\mathbf{Y})$ where $\mathbf{Y}$ is arbitrary. The additional constraints on $\mathbf{X}$, i.e., that $\mathbf{X}$ is row-stochastic, now translate into both $\mathbf{X}_1$ and $\mathbf{Y}$ being row-stochastic. Clearly, $\mathbf{X}_1 \geqslant \mathbf{0}$ is equivalent to $\mathbf{A}_1^{-1}\mathbf{A}_2\mathbf{Y} \leqslant \mathbf{A}_1^{-1}\mathbf{B}$. Moreover, when $\mathbf{Y}$ is row-stochastic we obtain $\mathbf{X}_1\mathbf{e} = \mathbf{A}_1^{-1}(\mathbf{B} - \mathbf{A}_2\mathbf{Y})\mathbf{e} = \mathbf{A}_1^{-1}(\mathbf{Be} - \mathbf{A}_2\mathbf{Ye}) = \mathbf{A}_1^{-1}(\mathbf{Ae} - \mathbf{A}_2\mathbf{e}) = \mathbf{A}_1^{-1}\mathbf{A}_1\mathbf{e} = \mathbf{e}$. The desired result now follows. □

Therefore, a study of the majorization polytope $\mathscr{M}(\mathbf{A} \succ \mathbf{B})$ reduces to a study of the *reduced majorization polytope* $\mathscr{M}^*(\mathbf{A} \succ \mathbf{B})$ which lies in a lower-dimensional space $\mathbb{R}^{n-m,m}$. The two polytopes are affinely isomorphic. When $n - m$ is small, this may make it possible to obtain much more information about these polytopes. When $n = m$ we trivially have that $\mathscr{M}(\mathbf{A} \succ \mathbf{B}) = \{\mathbf{A}^{-1}\mathbf{B}\}$. More interestingly, we now give a complete description of all the vertices of $\mathscr{M}(\mathbf{A} \succ \mathbf{B})$ in the case $n - m = 1$.

Let $n = m + 1$ so $\mathbf{A} = [\mathbf{A}_1, \mathbf{a}_n]$. Moreover, let $\mathbf{Y} = \mathbf{y}^{\mathrm{T}}$ where $\mathbf{y} \in \mathbb{R}^p$ and we want to find the (column) vectors $\mathbf{y}$ in the reduced majorization polytope (viewed as a polytope in $\mathbb{R}^p$ now). Define $\mathbf{C} = \mathbf{A}_1^{-1}\mathbf{B}$ and $\mathbf{d} = \mathbf{A}_1^{-1}\mathbf{a}_n$. We see that $\mathbf{y} \in \mathscr{M}^*(\mathbf{A} \succ \mathbf{B})$ if and only if $\mathbf{dy}^{\mathrm{T}} \leqslant \mathbf{C}$, or equivalently,

$$(*) \quad d_i \cdot y_j \leqslant c_{i,j} \quad \text{for all } i \leqslant m, \quad j \leqslant p.$$

This system just provides lower and upper bounds on each variable. Let I_+, I_0 and I_- denote the set of indices $i \leqslant p$ such that d_i is positive, zero or negative, respectively. Then $(*)$ is equivalent to

$$\begin{aligned} &\text{(i)} \;\; c_{i,j} \geqslant 0 && \text{for all } i \in I_0, \quad j \leqslant p, \\ &\text{(ii)} \;\; l_j \leqslant y_j \leqslant u_j && \text{for all } j \leqslant p, \end{aligned} \tag{5}$$

where $l_j := \max\{c_{i,j}/d_i : i \in I_-, \, j \leqslant p\}$ and $u_j := \min\{c_{i,j}/d_i : i \in I_+, \, j \leqslant p\}$. From this discussion we arrive at the following result (with the notation introduced above).

Proposition 4.2. *When $n = m + 1$ and $\mathbf{A}_1$ is nonsingular we have that $\mathbf{A} \succ \mathbf{B}$ if and only if* (5)(i) *holds, $l_j \leqslant u_j$ for all $j \leqslant p$ and $\sum_j l_j \leqslant 1 \leqslant \sum_j u_j$. Moreover, when these conditions hold, $\mathscr{M}^*(\mathbf{A} \succ \mathbf{B})$ is the solution set of* (5) *and the vertices are of the form $y_j \in \{l_j, u_j\}$ for all but possibly one j and $\sum_j y_j = 1$.*

From this, due to Proposition 4.1, one gets a complete description of both facets and vertices of majorization polytopes for the case when $n = m + 1$. We remark that all the vertices of the polytope $\Omega_3(\mathbf{x} \succ \mathbf{y})$ when $n = 3$ were determined in [4] ($\Omega_3(\mathbf{x} \succ \mathbf{y})$ consists of the doubly stochastic 3×3-matrices $\mathbf{S}$ satisfying $\mathbf{Sx} = \mathbf{y}$ for given $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$).

As a small example consider

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 \\ 6 & 3 & 1 \end{bmatrix}, \qquad \mathbf{B} = \begin{bmatrix} 1 & 1 & 1 \\ 4 & 4 & 2 \end{bmatrix}.$$

Then $\mathbf{A} \succ \mathbf{B}$ as $(6, 3, 1) \succ (4, 4, 2)$. Some calculation shows that the linear system $(*)$ defining the reduced majorization polytope is $0 \leqslant y_1 \leqslant 2/5,\ \ 0 \leqslant y_2 \leqslant 2/5,\ \ 1/2 \leqslant y_3 \leqslant 4/5$ and the vertices are $(0, 2/5, 3/5)$, $(2/5, 0, 3/5)$, $(2/5, 1/10, 1/2)$, $(1/10, 2/5, 1/2)$ and $(1/5, 0, 4/5)$.

5. The case of two rows

Throughout this section we consider the case when $m = 2$ so $\mathbf{A}$ and $\mathbf{B}$ are matrices

$$\mathbf{A} = \begin{bmatrix} a_{1,1} & \dots & a_{1,n} \\ a_{2,1} & \dots & a_{2,n} \end{bmatrix}, \qquad \mathbf{B} = \begin{bmatrix} b_{1,1} & \dots & b_{1,p} \\ b_{2,1} & \dots & b_{2,p} \end{bmatrix}.$$

We shall assume that (i) both matrices are nonnegative, (ii) $a_{1,j} > 0$ for $j \leqslant n$ and $b_{1,j} > 0$ for $j \leqslant p$, and (iii) $\mathbf{Ae} = \mathbf{Be}$. For instance, the matrices in Section 1 in connection with vector majorization fit into this framework. We define $w_i := \sum_{j=1}^{n} a_{i,j} = \sum_{j=1}^{p} b_{i,j}$ for $i = 1, 2$ and $\mathbf{w} = (w_1, w_2)$. Note that $w_1 > 0$.

We need some sets and functions associated with the matrix $\mathbf{A}$ (and similar concepts and notation are used in connection with $\mathbf{B}$). As usual the jth column of $\mathbf{A}$ is $\mathbf{a}^j$. The set

$$Z_A := \sum_{j=1}^{n} \mathrm{conv}(\{\mathbf{0}, \mathbf{a}^j\})$$

is a zonotope (a vector sum of line segments) in $\mathbb{R}^2$ which is symmetric around the point $(1/2)\mathbf{w}$. An example is shown in Fig. 1. Note that the zonotopes Z_A and Z_B have the same point of symmetry. The "upper boundary" of Z_A may be seen as the graph of a function $\beta_A\colon [0, w_1] \to \mathbb{R}$ given by

$$\begin{aligned} \beta_A(h) &= \max\{y\colon (h, y) \in Z_A\} \\ &= \max\left\{ \sum_{j=1}^{n} a_{2,j} v_j : \sum_{j=1}^{n} a_{1,j} v_j \leqslant h, \quad 0 \leqslant v_j \leqslant 1 \right. \\ &\qquad\qquad \left. \text{for } j = 1, \dots, n \right\} \end{aligned}$$

for $0 \leqslant h \leqslant w_1$. The function β_A is piecewise linear, concave, nondecreasing and continuous (and its graph has $\mathbf{0}$ and $\mathbf{w}$ as its endpoints). We also define

$$\Delta^A(j) = a_{2,j}/a_{1,j} \quad \text{for } j = 1, \dots, n$$

so $\Delta^A(j)$ is the slope of the line segment $\mathrm{conv}(\{\mathbf{0}, \mathbf{a}_j\})$.

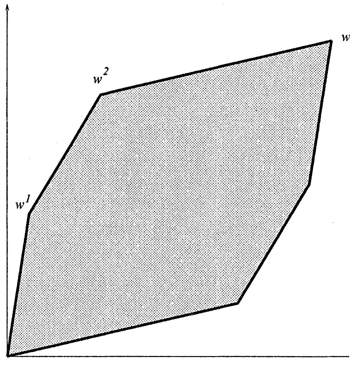

Fig. 1. The zonotope $Z_{\mathbf{A}}$.

We hereafter assume that the columns of **A** have been permuted so that $\Delta^A(j)$ is nonincreasing as a function of *j* and we then say that **A** is *monotone*. Similarly, we assume that **B** is monotone. This is with no loss of generality as permutations of columns of **A** and **B** correspond to line permutations of matrices in the majorization polytope. It follows from the monotonicity that $\beta_A(\sum_{j=1}^{k} a_{1,j}) = \sum_{j=1}^{k} a_{2,j}$ for $k = 0, 1, \ldots, n$ (and β_A is linear on each interval $[\sum_{j=1}^{k} a_{1,j}, \sum_{j=1}^{k+1} a_{1,j}]$).

The following result was shown in [5]. It gives a geometrical characterization of matrix majorization.

Theorem 5.1. *The following conditions are equivalent for nonnegative matrices* $\mathbf{A} \in \mathbb{R}^{2,n}$ *and* $\mathbf{B} \in \mathbb{R}^{2,p}$ *with* $\mathbf{Ae} = \mathbf{Be}$:

(i) $\mathbf{A} \succ \mathbf{B}$.

(ii) $Z_A \supseteq Z_B$.

(iii) $\beta_A \geqslant \beta_B$.

(iv) $\beta_A\left(\sum_{j=1}^{k} b_{1,j}\right) \geqslant \sum_{j=1}^{k} b_{2,j}$ *for* $k = 1, \ldots, p-1$.

Condition (iv) has special interest, it can be seen as a generalization of the (defining) partial sum ordering of vector majorization. Thus, when $a_{1,j} = b_{1,j} = 1$ for all *j* (iv) specializes into $\sum_{j=1}^{k} a_{[j]} \geqslant \sum_{j=1}^{k} b_{[j]}$ for $k = 1, \ldots, n-1$.

Since **A** is monotone, there are integers $0 = i_0 < i_1 < \cdots < i_r = n$ and numbers $\Delta_1^A > \Delta_2^A > \cdots > \Delta_r^A$ such that $\Delta^A(j) = \Delta_k^A$ for all $i_{k-1} < j \leqslant i_k$ and $1 \leqslant k \leqslant r$. Similarly, for **B**, we may construct the numbers $\Delta_1^B > \cdots > \Delta_s^B$.

Remark. Consider the matrix $\mathbf{A}'$ obtained from **A** by replacing the columns $\mathbf{a}^j$ for $i_{k-1} < j \leqslant i_k$ by the single column vector $\sum_{j=i_{k-1}+1}^{i_k} \mathbf{a}^j$. Onecan show that $\mathbf{A}' \succ$

$\mathbf{A}$ and $\mathbf{A} \succ \mathbf{A}'$, i.e., these matrices are equivalent with respect to the majorization preorder. Moreover, all the numbers $\Delta_j^{A'}$ are distinct. Similarly we may construct $\mathbf{B}'$ from $\mathbf{B}$. The benefit of this approach is that the analysis of $\mathscr{M}(\mathbf{A}' \succ \mathbf{B}')$ becomes less technical. However, this process does influence the majorization polytope, so we prefer to treat the more general case in the following.

We need a result on the representation of points on the upper boundary of Z_A. The vertices on the upper boundary are the points $\mathbf{w}^k = \sum_{j=1}^{i_k} \mathbf{a}_j$ for $k = 0, \ldots, r$ (where $\mathbf{w}^0 = \mathbf{0}$ and $\mathbf{w}^r = \mathbf{w}$). Note that we have $0 = w_1^0 < w_1^1 < \cdots < w_1^r = w_1$. Let now $0 \leqslant h \leqslant w_1$. Define $k_1(h) = \max\{k : w_1^k \leqslant h\}$ and $k_2(h) = \min\{k : w_1^k \geqslant h\}$. Thus, there are two possibilities: (i) $k_1(h) = k_2(h)$ (i.e., h is the first coordinate of one of the vertices $\mathbf{w}^{\mathrm{t}}$, $t \leqslant k$), and (ii) $k_1(h) = k_2(h) - 1$.

Lemma 5.2. *Let* $\mathbf{h} = (h_1, h_2)$ *where* $0 \leqslant h_1 \leqslant w_1$ *and* $h_2 = \beta_A(h_1)$, *and let* $k_i = k_i(h_1)$ *for* $i = 1, 2$. *Consider a point* $\mathbf{z} \in [0, 1]^n$ *with* $\mathbf{Az} = \mathbf{h}$.

Then $\mathbf{z}$ *satisfies* $z_i = 1$ *for* $i \leqslant i_{k_1}$ *and* $z_i = 0$ *for* $i > i_{k_2}$. *In particular, when* $k_1 = k_2$ *the point* $\mathbf{z}$ *is unique.*

Proof. The point $\mathbf{h}$ lies on the upper boundary of Z_A. Assume first that $\mathbf{h} = \mathbf{w}^k$ (a vertex of Z_A). Then there is a vector $\mathbf{c} \in \mathbb{R}^2$ such that $\mathbf{c}^{\mathrm{T}}\mathbf{a}^j$ is positive when $j \leqslant i_k$ and negative otherwise ($\mathbf{c}$ is an outward normal vector to Z_A at $\mathbf{h}$). Then $\mathbf{h}$ is the unique optimal solution to the linear program $\max\{\mathbf{c}^{\mathrm{T}}\mathbf{x} : \mathbf{x} \in Z_A\}$. But each $\mathbf{x} \in Z_A$ has the form $\mathbf{x} = \mathbf{Az}$ for some $\mathbf{z} \in [0, 1]^n$ and $\mathbf{c}^{\mathrm{T}}\mathbf{x} = \mathbf{c}^{\mathrm{T}}\mathbf{Az} = \sum_{j=1}^{n} (\mathbf{c}^{\mathrm{T}}\mathbf{a}_j) z_j$. Thus, this linear function is maximized precisely when we let $z_j = 1$ for $j \leqslant i_k$ and $z_j = 0$ otherwise. This proves (i). Statement (ii) is proved similarly, except that the vector $\mathbf{c}$ is now an outward normal vector to the edge of Z_A between $\mathbf{w}^k$and $\mathbf{w}^{k+1}$ (so $\mathbf{c}^{\mathrm{T}}\mathbf{a}^j$ is positive, zero or negative according to whether $j \leqslant i_k$, $i_k < j \leqslant i_{k+1}$ or $j > i_{k+1}$, respectively). □

We say that the majorization $\mathbf{A} \succ \mathbf{B}$ has a *coincidence at* h where $0 < h < w_1$ if $\beta_A(h) = \beta_B(h)$. Otherwise, we say that $\mathbf{A} \succ \mathbf{B}$ has no coincidence. If there is a coincidence at h, then there must also be a coincidence at one of the points $\sum_{j=1}^{r} b_{1,j}$ (this follows from the properties of β_A and β_B). If $\Delta_B(1) = \cdots = \Delta_B(p)$ we say that $\mathbf{B}$ is a Δ-*constant* matrix. Note that if $\mathbf{B}$ is Δ-constant, then the function β_B is linear and Z_B degenerates into the line segment $\mathrm{conv}(\{\mathbf{0}, \mathbf{w}\})$.

The following result of the structure of matrices in $\mathscr{M}(\mathbf{A} \succ \mathbf{B})$ may be seen as a generalization of a result in [8].

Theorem 5.3. *Assume that the majorization* $\mathbf{A} \succ \mathbf{B}$ *has a coincidence at* $h_1 = \sum_{j=1}^{r} b_{1,j}$ *and let* $k_i = k_i(h_1)$ *for* $i = 1, 2$. *Then each* $\mathbf{X} \in \mathscr{M}(\mathbf{A} \succ \mathbf{B})$ *satisfies* $x_{i,j} = 0$ *for* $i \leqslant i_{k_1}$, $j > r$ *and* $x_{i,j} = 0$ *for* $i > i_{k_2}$, $j \leqslant r$.

Proof. Let $\mathbf{h} = (h_1, h_2)$ where $h_2 = \beta_B(h_1) = \sum_{j=1}^{r} b_{2,j}$. Thus $\mathbf{h} = \sum_{j=1}^{r} \mathbf{b}_j$. Since $\mathbf{X} \in \mathscr{M}(\mathbf{A} \succ \mathbf{B})$, we have $\mathbf{A}\mathbf{x}^j = \mathbf{b}^j$ for $j \leqslant n$ and therefore

$$\mathbf{h} = \sum_{j=1}^{r} \mathbf{b}^j = \sum_{j=1}^{r} \mathbf{A}\mathbf{x}^j = \mathbf{A}\sum_{j=1}^{r} \mathbf{x}^j.$$

The vector $\mathbf{z} := \sum_{j=1}^{r} \mathbf{x}^j$ also satisfies $\mathbf{z} \in [0, 1]^n$ (as $\mathbf{X}$ is row-stochastic). From Lemma 5.2 we obtain that $\sum_{j=1}^{r} x_{i,j} = 1$ for $i \leqslant i_{k_1}$ and $\sum_{j=1}^{r} x_{i,j} = 0$ for $i > i_{k_2}$. This implies the desired conclusion. □

We shall need a continuity result saying that the set Z_A depends continuously on the matrix $\mathbf{A}$. To state this more precisely, we let $\rho(U, V)$ denote the Hausdorff distance between two sets U and v (in $\mathbb{R}^2$), i.e., $\rho(U, V) = \max\{\max_{\mathbf{u}\in U} \min_{\mathbf{v}\in V} \|\mathbf{u} - \mathbf{v}\|, \max_{\mathbf{v}\in V} \min_{\mathbf{u}\in U} \|\mathbf{u} - \mathbf{v}\|\}$.

Lemma 5.4. *Assume that* $\mathbf{A}, \hat{\mathbf{A}} \in \mathbb{R}^{2,n}$ *and* $\mathbf{A}\mathbf{e} = \hat{\mathbf{A}}\mathbf{e}$. *If* $\|\mathbf{a}_j - \hat{\mathbf{a}}_j\| \leqslant \epsilon$ *for all* $j \leqslant n$, *then* $\rho(Z_A, Z_{\hat{A}}) \leqslant n\epsilon$.

Proof. Assume that $\|\mathbf{a}_j - \hat{\mathbf{a}}_j\| \leqslant \epsilon$ for all j. Let $\mathbf{x} \in Z_A$ so there are numbers $z_1, \ldots, z_n$ in $[0, 1]$ such that $\mathbf{x} = \sum_{j=1}^{n} z_j \mathbf{a}^j$. We let $\hat{\mathbf{z}} = \sum_{j=1}^{n} z_j \hat{\mathbf{a}}_j$ and note that $\hat{\mathbf{z}} \in Z_{\hat{A}}$. Moreover, $\|\mathbf{z} - \hat{\mathbf{z}}\| = \|\sum_{j=1}^{n} z_j(\mathbf{a}^j - \hat{\mathbf{a}}^j)\| \leqslant \sum_{j=1}^{n} \|\mathbf{a}_j - \hat{\mathbf{a}}_j\| \leqslant n\epsilon$. This implies that $\max_{\mathbf{z}\in Z_A} \min_{\hat{\mathbf{z}}\in Z_{\hat{A}}} \|\mathbf{z} - \hat{\mathbf{z}}\| \leqslant n\epsilon$ and (by symmetry) $\rho(Z_A, Z_{\hat{A}}) \leqslant n\epsilon$. □

The following theorem generalizes a result of [2] about the existence of a positive matrix in majorization polytopes.

Theorem 5.5. *Let* $\mathbf{A} \succ \mathbf{B}$. *Then* $\mathscr{M}(\mathbf{A} \succ \mathbf{B})$ *contains a positive matrix if and only if* $\mathbf{A} \succ \mathbf{B}$ *has no coincidence or* $\mathbf{B}$ *is a* Δ*-constant matrix.*

Proof. Assume that $\mathbf{A} \succ \mathbf{B}$ has a coincidence and that $\mathbf{B}$ is not a Δ-constant matrix. As $\mathbf{A} \succ \mathbf{B}$ this implies that $\mathbf{A}$ is not a Δ-constant matrix. Moreover, as remarked above, we may assume that $\mathbf{A} \succ \mathbf{B}$ has a coincidence at $h_1 = \sum_{j=1}^{r} b_{1,j}$ for some $r \in \{1, \ldots, p-1\}$. But then it follows from Theorem 5.3 that each $\mathbf{X} \in \mathscr{M}(\mathbf{A} \succ \mathbf{B})$ has some zero entries, so there cannot be any positive matrix in the majorization polytope.

To prove the converse, assume first that $\mathbf{B}$ is a Δ-constant matrix so we have that $b_{2,j} = b_{1,j} \cdot \theta$ for some (positive) number θ. We then have $w_2 = \sum_j b_{2,j} = \theta w_1$. Define $\mathbf{X} \in \mathbb{R}^{n,p}$ by $x_{i,j} = b_{1,j}/w_1$ for all $i \leqslant n$ and $j \leqslant p$ (recall that $w_1 > 0$). Then $x_{i,j} > 0$ for all i, j and $\sum_j x_{i,j} = (1/w_1)\sum_j b_{1,j} = 1$, thus $\mathbf{X}$ is a positive row-stochastic matrix. Moreover $(\mathbf{AX})_{i,j} = \sum_t a_{i,t}x_{t,j} = (b_{1,j}/w_1)\sum_t a_{i,t} = b_{1,j} \cdot w_i/w_1 = b_{i,j}$ for all i, j (as $w_2 = \theta w_1$). Thus, $\mathbf{AX} = \mathbf{B}$ and we have found a positive matrix in $\mathscr{M}(\mathbf{A} \succ \mathbf{B})$.

Assume next that $\mathbf{A} \succ \mathbf{B}$ has no coincidence. This means that $\beta_A(h) > \beta_B(h)$ for all $0 < h < w_1$. Let ϵ be a small postive number and define the matrix $\mathbf{X}(\epsilon) = [x_{i,j}(\epsilon)] \in \mathbb{R}^{n,n}$ by $x_{i,j}(\epsilon) = 1 - (n-1)\epsilon$ when $i = j$ and $x_{i,j}(\epsilon) = \epsilon$ when $i \neq j$. For ϵ small enough $\mathbf{X}(\epsilon)$ is a positive row-stochastic (in fact doubly stochastic) matrix. Define $\mathbf{A}(\epsilon) \in \mathbb{R}^{2,n}$ by $\mathbf{A}(\epsilon) = \mathbf{A}\mathbf{X}(\epsilon)$ and let $\mathbf{a}^j(\epsilon)$ denote the jth column of $\mathbf{A}(\epsilon)$. Then we obtain (as $\mathbf{w} = \sum_j \mathbf{a}^j$)

$$\mathbf{a}^j(\epsilon) = (1-(n-1)\epsilon)\mathbf{a}^j + \epsilon \sum_{i \neq j} \mathbf{a}^i = (1 - n\epsilon)\mathbf{a}^j + \epsilon\mathbf{w},$$

which gives $\|\mathbf{a}^j - \mathbf{a}^j(\epsilon)\| \leqslant \epsilon(n\|\mathbf{a}^j\| + \|\mathbf{w}\|)$. It follows from Lemma 5.4 that we can get $Z_{A(\epsilon)}$ arbitrarily close to Z_A by choosing ϵ small enough. This means that we have

$$\beta_A(h) \geqslant \beta_{A(\epsilon)}(h) > \beta_B(h) \quad \text{for all } 0 < h < w_1.$$

Therefore, by Theorem 5.1, $\mathbf{A}(\epsilon) \succ \mathbf{B}$ and there is a row-stochastic matrix $\mathbf{Y}$ with $\mathbf{A}(\epsilon)\mathbf{Y} = \mathbf{B}$. Observe that $\mathbf{Y}$ has no zero column (as that would imply that $\mathbf{B}$ has a zero column which contradicts that $b_{1,j} > 0$ for all j). Collecting our results we now get

$$\mathbf{B} = \mathbf{A}(\epsilon)\mathbf{Y} = \mathbf{A}\mathbf{X}(\epsilon)\mathbf{Y}.$$

But the matrix $\mathbf{X}(\epsilon)\mathbf{Y}$ is positive (as $\mathbf{X}(\epsilon)$ is positive and no column of $\mathbf{Y}$ is zero) and row-stochastic and the proof is complete. □

The following result holds for arbitrary m, although we shall only use it for $m = 2$ in the following lemma.

Lemma 5.6. *Let $\mathbf{A} \in \mathbb{R}^{m,n}$, $\mathbf{B} \in \mathbb{R}^{m,p}$ and assume that $\mathbf{A} \succ \mathbf{B}$. Then the dimension of the affine set $\{\mathbf{X} \in \mathbb{R}^{n,p} : \mathbf{A}\mathbf{X} = \mathbf{B},\ \mathbf{X}\mathbf{e} = \mathbf{e}\}$ is equal to $np - (p-1)\text{rank}(\mathbf{A})$.*

Proof. The matrix equations $\mathbf{A}\mathbf{X} = \mathbf{B}$, $\mathbf{X}\mathbf{e} = \mathbf{e}$ may be written as the following linear system with variables being the columns of $\mathbf{X}$

$$(\star) \quad \mathbf{A}\mathbf{x}^j = \mathbf{b}^j \ \text{ for } j \leqslant p, \quad \mathbf{x}^1 + \cdots + \mathbf{x}^p = \mathbf{e}.$$

We need to determine the dimension of the affine set consisting of the solutions $\mathbf{x}^1, \ldots, \mathbf{x}^p$ of $(\star)$. Note that the system is consistent as $\mathbf{A} \succ \mathbf{B}$. We may eliminate $\mathbf{x}^p$ from the last equation in $(\star)$, so $\mathbf{x}^p = \mathbf{e} - \sum_{j=1}^{p-1} \mathbf{x}^j$ and then $\mathbf{A}\mathbf{x}^p = \mathbf{b}^p$ becomes $\sum_{j=1}^{p-1} \mathbf{A}\mathbf{x}^j = \mathbf{e} - \mathbf{b}^p$. Thus, $\dim(\mathscr{M}(\mathbf{A} \succ \mathbf{B})) = np - \text{rank}(\hat{\mathbf{A}})$ where $\hat{\mathbf{A}}$ is the $pm \times (p-1)n$-dimensional block matrix

$$\hat{\mathbf{A}} = \begin{bmatrix} \mathbf{A} & & \\ & \ddots & \\ & & \mathbf{A} \\ \mathbf{A} & \cdots & \mathbf{A} \end{bmatrix}.$$

Now, we note that the last (block) row of $\hat{\mathbf{A}}$ is the sum of the other rows, so we may delete the last block row in order to compute rank($\hat{\mathbf{A}}$). The resulting matrix is the direct product of $p-1$ matrices each being equal to $\mathbf{A}$, so rank($\hat{\mathbf{A}}$) $=$ $(p-1)$rank($\mathbf{A}$) and the result follows. $\square$

By applying this lemma to our case of $m = 2$ we may calculate the dimension of the majorization polytope in a certain situation.

Proposition 5.7. *Consider again the case of* $m = 2$. *Assume that the majorization* $\mathbf{A} \succ \mathbf{B}$ *holds and that it has no coincidence. Then* $\dim(\mathscr{M}(\mathbf{A} \succ \mathbf{B})) = (n-2)p + 2$.

Proof. When $\mathbf{A} \succ \mathbf{B}$ has no coincidence we know from Theorem 5.5 that $\mathscr{M}(\mathbf{A} \succ \mathbf{B})$ contains a positive matrix. Therefore none of the nonnegatity constraints $x_{i,j} \geqslant 0$ can be an implicit equality for the majorization polytope. Thus, the implicit equalities are simply $\mathbf{AX} = \mathbf{B}$, $\mathbf{Xe} = \mathbf{e}$ and Lemma 5.7 gives $\dim(\mathscr{M}(\mathbf{A} \succ \mathbf{B})) = np - (p-1)\text{rank}(\mathbf{A})$. We have that rank($\mathbf{A}$) $\leqslant 2$. The first row of $\mathbf{A}$ is positive so the rank is nonzero. If rank($\mathbf{A}$) $= 1$, Z_A would be the line segment conv($\{\mathbf{0}, \mathbf{w}\}$) so $Z_A = Z_B$ and there would be a coincidence. Thus, rank($\mathbf{A}$) $= 2$ and the desired result follows. $\square$

We may now derive a result on the structure of $\mathscr{M}(\mathbf{A} \succ \mathbf{B})$. Let $\mathbf{A} \succ \mathbf{B}$ so we have $\beta_A \geqslant \beta_B$. Consider the set $S = \{h \in [0, w_1]\colon \beta_A(h) = \beta_B(h)\}$. Assume, for simplicity, that S is finite and that each $h \in S$ is the first coordinate of a vertex of Z_A. This implies that there are integers $r_1, \ldots, r_s$ such that $S = \{0, \sum_{j=1}^{r_1} b_{2,j}, \ldots, \sum_{j=1}^{r_s} b_{2,j}, 1\}$. We here have that $k_1(\sum_{j=1}^{r_\nu} b_{2,j}) = k_2(\sum_{j=1}^{r_\nu} b_{2,j}) := k(\nu)$ for $\nu = 1, \ldots, s$. Moreover, each $\mathbf{X} \in \mathscr{M}(\mathbf{A} \succ \mathbf{B})$ has the form

$$\mathbf{X} = \mathbf{X}_1 \oplus \cdots \oplus \mathbf{X}_s,$$

where $\mathbf{X}_\nu$ is a $(i_{k(\nu+1)} - i_{k(\nu)}) \times (r_{\nu+1} - r_\nu)$-dimensional row-stochastic matrix, for $\nu = 0, \ldots, s-1$ (and $r_0 = 0$, $i_{k(0)} = 0$). Define now $\mathbf{A}^{(\nu)}$ as the submatrix of $\mathbf{A}$ consisting of the columns $\mathbf{a}^j$ where $i_{k(\nu)} < j \leqslant i_{k(\nu+1)}$. Let $\mathbf{B}^{(\nu)}$ be the submatrix of $\mathbf{B}$ consisting of the columns $\mathbf{b}^j$ where $r_\nu < j \leqslant r_{\nu+1}$. We then have that

$$\mathscr{M}(\mathbf{A} \succ \mathbf{B}) = \mathscr{M}(\mathbf{A}^{(1)} \succ \mathbf{B}^{(1)}) \oplus \cdots \oplus \mathscr{M}(\mathbf{A}^{(s)} \succ \mathbf{B}^{(s)})$$

and therefore (from Proposition 5.7)

$$\begin{aligned}\dim(\mathscr{M}(\mathbf{A} \succ \mathbf{B}) &= \sum_{\nu=1}^{s} \dim(\mathscr{M}(\mathbf{A}^{(\nu)} \succ \mathbf{B}^{(\nu)}) \\ &= \sum_{\nu=1}^{s} [(i_{k(\nu+1)} - i_{k(\nu)} - 2)(r_{\nu+1} - r_\nu) + 2].\end{aligned}$$

In general (without our simplifying assumption on the set S) the majorization polytope has a “stair-case” nonzero pattern. This was discussed in [2] for the special case of vector majorization. We omit the technicalities of such a description here.

We now proceed to show that the majorization polytope contains a matrix which is the product of certain simple row-stochastic matrices.

Let $k \in \{0, \ldots, p\}$, $0 \leqslant \alpha \leqslant 1$ and $0 \leqslant \gamma \leqslant 1$ and consider the $n \times n$ row stochastic matrix

$$\mathbf{S}(\alpha, \gamma; k) = \begin{bmatrix} \mathbf{S}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix},$$

where $\mathbf{I}$ is the identity matrix and the matrix $\mathbf{S}_1$ has order $k + 1$ and is given by

$$\mathbf{S}_1 = \begin{bmatrix} \gamma & 1-\gamma & 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ \gamma & 1-\gamma & 0 & 0 & \ldots & 0 \\ \gamma\alpha & (1-\gamma)\alpha & 1-\alpha & 0 & \ldots & 0 \end{bmatrix}.$$

If $\mathbf{A} \in \mathbb{R}^{m,n}$ and $\mathbf{C} = \mathbf{AS}(\alpha, \gamma; k)$ then the columns of $\mathbf{C}$ are $\mathbf{c}^1 = \gamma(\sum_{i=1}^{k} \mathbf{a}^i + \alpha\mathbf{a}^{k+1})$, $\mathbf{c}^2 = (1-\gamma)(\sum_{i=1}^{k} \mathbf{a}^i + \alpha\mathbf{a}^{k+1})$, $\mathbf{c}^3 = (1-\alpha)\mathbf{a}^{k+1}$, $\mathbf{c}^4 = \cdots = \mathbf{c}^{k+1} = \mathbf{0}$ and $\mathbf{c}^j = \mathbf{a}^j$ for $j = k+2, \ldots, n$. Each matrix obtained from $\mathbf{S}(\alpha, \gamma; k)$ by permuting its lines (rows and columns) and possibly deleting columns with all zeros is called an *S-matrix*. Every S-matrix is row-stochastic so if $\mathbf{AX} = \mathbf{B}$, where $\mathbf{X}$ is a product of S-matrices, then $\mathbf{A} \succ \mathbf{B}$ and $\mathbf{X} \in \mathscr{M}(\mathbf{A} \succ \mathbf{B})$. More interestingly, the converse also holds as stated in the following theorem.

Theorem 5.8. *Let* $\mathbf{A} \in \mathbb{R}^{2,n}$ *and* $\mathbf{B} \in \mathbb{R}^{2,p}$ *and assume that* $\mathbf{A} \succ \mathbf{B}$. *Then* $\mathscr{M}(\mathbf{A} \succ \mathbf{B})$ *contains a matrix which is the product of at most p S-matrices.*

Proof. Let $\mathbf{A} \succ \mathbf{B}$ and we may assume that both $\mathbf{A}$ and $\mathbf{B}$ are monotone. The proof is by induction on p.

If $p = 1$, then $\mathbf{B} = [\mathbf{b}^1]$ and since $\mathbf{A} \succ \mathbf{B}$ we get $\mathbf{Ae} = \mathbf{Be} = \mathbf{B}$. Here the $n \times 1$ matrix $\mathbf{X} = [\mathbf{e}]$ is an S-matrix (obtained from $\mathbf{S}(1, 1; n-1)$ be deleting all columns except the first which is $\mathbf{e}$) so the desired result holds for $p = 1$.

Assume that the theorem holds when $\mathbf{B}$ has at most $p - 1$ columns. We may assume that both $\mathbf{A}$ and $\mathbf{B}$ are monotone. Since $\mathbf{A} \succ \mathbf{B}$ we have that $\beta_A \geqslant \beta_B$. Moreover, as $\beta_A(0) = \beta_B(0) = 0$ it follows that $\Delta^A(1) \geqslant \Delta^B(1)$. Therefore, there is a $\lambda \geqslant 1$ such that the point $\mathbf{h} = \lambda\mathbf{b}^1$ lies on the graph of β_A. Then we can find (confer $\mathbf{z}$ in Lemma 5.2) a $k \in \{0, \ldots, p\}$ and $0 \leqslant \alpha \leqslant 1$ such that $\mathbf{h} = \sum_{i=1}^{k} \mathbf{a}^i + \alpha\mathbf{a}^{k+1}$. Letting $\gamma = 1/\lambda$ (so $0 < \gamma \leqslant 1$) we now get $\mathbf{b}^1 = \gamma(\sum_{i=1}^{k} \mathbf{a}^i + \alpha\mathbf{a}^{k+1})$. This implies that the first column of the matrix $\mathbf{A}' := \mathbf{AS}(\alpha, \gamma; k)$ equals $\mathbf{b}^1$. Moreover, we have $\beta_A \geqslant \beta_{A'} \geqslant \beta_B$ as the graph of $\beta_{A'}$ is linear between $\mathbf{0}$ and $\mathbf{h}$ (and $\mathbf{b}^1$ lies on this line segment) and thereafter it coincides with the graph of β_A. Thus $\mathbf{A}' \succ \mathbf{B}$ and therefore $\mathbf{A}'_1 \succ \mathbf{B}_1$ where these two matrices are obtained from $\mathbf{A}'$and $\mathbf{B}$ respectively by deleting the first column (which is $\mathbf{b}_1$ in both matrices). But $\mathbf{B}_1$ has $p - 1$ columns so by induction $\mathbf{A}_1'\mathbf{X}' = \mathbf{B}_1$ for some matrix $\mathbf{X}'$ which is the product of at most $p-1$ S-matrices $\mathbf{S}_i'$. From this we see that $\mathbf{AX} = \mathbf{B}$ where $\mathbf{X}$ is the product of $\mathbf{S}(\alpha, \gamma; k)$ and the matrices

$$\begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{S}_i' \end{bmatrix}$$

each being an S-matrix. Thus $\mathbf{X}$ is the product of at most p S-matrices which completes the induction proof. □

This theorem is along the same lines as a basic fact for vector majorization: for row vectors $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ we have that $\mathbf{a} \succ \mathbf{b}$ if and only if $\mathbf{aX} = \mathbf{b}$ for a doubly stochastic matrix $\mathbf{X}$ which is the product of at most n matrices corresponding to T-transforms. (Each such matrix is a convex combination of the identity matrix and a permutation matrix corresponding to a transposition.) We remark that the geometrical idea in the proof is to gradually "move" the curve β_A towards β_B. This may be done in several ways which proves the existence of related matrices in the majorization polytope.

6. Support-majorization

In this final section we consider some combinatorial properties of matrix majorization and majorization polytopes.

Let $\mathbf{A} \in \mathscr{M}_{m,n}$ and $\mathbf{B} \in \mathscr{M}_{m,p}$. Define $\text{supp}(\mathbf{A}) = \{(i, j): a_{i,j} > 0\}$. Let the *support-class* of $\mathbf{A}$, denoted by $\mathscr{S}(\mathbf{A})$, consist of those $m \times n$ row-stochastic matrices $\mathbf{A}'$ satisfying $\text{supp}(\mathbf{A}') = \text{supp}(\mathbf{A})$. $\mathscr{S}(\mathbf{B})$ is defined similarly. We say that $\mathbf{A}$ *support-majorizes* $\mathbf{B}$, and write $\mathbf{A} \succ^{\mathrm{s}} \mathbf{B}$, provided that for every $\mathbf{A}' \in \mathscr{S}(\mathbf{A})$ and $\mathbf{B}' \in \mathscr{S}(\mathbf{B})$ it holds that $\mathbf{A}' \succ \mathbf{B}'$. This means that $\mathbf{A} \succ \mathbf{B}$ and that the majorization is preserved under every change of the entries of the matrices as long as one stays in the respective support-classes. An even stronger notion than $\succ^{\mathrm{s}}$ is the following. We say that $\mathbf{A}$ *strongly support-majorizes* $\mathbf{B}$ if $\mathbf{A} \succ^{\mathrm{s}} \mathbf{B}$ and all the matrices in the majorization polytopes $\mathscr{M}(\mathbf{A}' \succ \mathbf{B}')$ for $\mathbf{A}' \in \mathscr{S}(\mathbf{A})$ and $\mathbf{B}' \in \mathscr{S}(\mathbf{B})$ belong to the same support-class (so this class only depends on $\mathbf{A}$and $\mathbf{B}$, not $\mathbf{A}'$ or $\mathbf{B}'$).

Consider two distinct row indices i and i' (where $i, i' \leqslant m$). Recall that $\mathbf{a}_i$ ($\mathbf{b}_i$) is the ith row of $\mathbf{A}$ ($\mathbf{B}$). If $\text{supp}(\mathbf{a}_i) \cap \text{supp}(\mathbf{a}_{i'}) \neq \emptyset$ implies that there is a $j \leqslant p$ such that $\mathbf{b}_i = \mathbf{b}_{i'} = \mathbf{e}_j$ (the jth coordinate vector), we say that rows i and i' are *nonconflicting*; otherwise they are in *conflict*.

Theorem 6.1. *Let* $\mathbf{A} \in \mathscr{M}_{m,n}$ *and* $\mathbf{B} \in \mathscr{M}_{m,p}$. *Then* $\mathbf{A} \succ^{\mathrm{s}} \mathbf{B}$ *if and only if no pair of rows is in conflict.*

Proof. Assume that rows i and i' are in conflict. Then there is a $k \in \text{supp}(\mathbf{a}_i) \cap \text{supp}(\mathbf{a}_{i'})$ and two distinct indices $j, j' \leqslant p$ with $j \in \text{supp}(\mathbf{b}_i)$ and $j' \in \text{supp}(\mathbf{b}_{i'})$.

Therefore we can find a matrix $\mathbf{B}' \in \mathscr{S}(\mathbf{B})$ with $b'_{i,j}, b'_{i',j'} > 1 - \epsilon$ where ϵ is a suitably small positive number (see later). Moreover, we can find a matrix $\mathbf{A}' \in \mathscr{S}(\mathbf{A})$ with $a'_{i,k} = a'_{i',k} > 1 - \epsilon$. We claim that

$$(*) \quad \begin{bmatrix} \mathbf{a}'_i \\ \mathbf{a}'_{i'} \end{bmatrix} \not\succ \begin{bmatrix} \mathbf{b}'_i \\ \mathbf{b}'_{i'} \end{bmatrix}.$$

This follows from Theorem 5.1 for by choosing $\epsilon > 0$ small enough we obtain $\beta_{A'_1} \not\succcurlyeq \beta_{B'_1}$. But from $(*)$ we conclude that $\mathbf{A}' \not\succ \mathbf{B}'$ (for if there were a row-stochastic matrix $\mathbf{X}$ with $\mathbf{A}'\mathbf{X} = \mathbf{B}'$ then

$$\begin{bmatrix} \mathbf{a}'_i \\ \mathbf{a}'_{i'} \end{bmatrix} \mathbf{X} = \begin{bmatrix} \mathbf{b}'_i \\ \mathbf{b}'_{i'} \end{bmatrix},$$

which contradicts $(*)$). We have therefore shown that a necessary condition for $\mathbf{A} \succ^{\mathrm{s}} \mathbf{B}$ is that no pair of rows is in conflict.

Assume that each pair of rows is nonconflicting. Construct the graph G with node set $I = \{1, \ldots, m\}$ and with an edge $[i, i']$ whenever $\mathrm{supp}(\mathbf{a}_i) \cap \mathrm{supp}(\mathbf{a}_{i'})$ is nonempty. Let the connected components of G be (the node sets) $I_1, \ldots, I_{r'}$ where the trivial components (a single node) are $I_{r+1}, \ldots, I_{r'}$ (where $1 \leqslant r \leqslant r'$). It follows from the nonconflicting assumption that there are column indices $j_1, \ldots, j_r$ such that $\mathbf{b}_i = \mathbf{e}_{j_k}$ for all $i \in I_k$, $k = 1, \ldots, r$. We may find a permutation matrix $\mathbf{P}$ of order m such that $\mathbf{PA}$ has the rows in I_k before all rows in $I_{k'}$ when $k < k'$. Next we may find a permutation matrix $\mathbf{Q}$ such that

$$\bar{\mathbf{A}} = \mathbf{PAQ} = \begin{bmatrix} \bar{\mathbf{A}}_1 & & \\ & \ddots & \\ & & \bar{\mathbf{A}}_{r+1} \end{bmatrix},$$

where $\bar{\mathbf{A}}_1, \ldots, \bar{\mathbf{A}}_r$, correspond to the rows $I_1, \ldots, I_r$, respectively (the nontrivial components) and $\bar{\mathbf{A}}_{r+1}$ is a disjoint-row-support matrix of the form (1). Further, the matrix $\bar{\mathbf{B}} = \mathbf{PB}$ may be written

$$\bar{\mathbf{B}} = \begin{bmatrix} \bar{\mathbf{B}}_1 \\ \vdots \\ \bar{\mathbf{B}}_{r+1} \end{bmatrix},$$

where for $i = 1, \ldots, r$ the matrix $\bar{\mathbf{B}}_i$ (with rows corresponding to I_i) has a column of all ones and the remaining columns are zero. The matrix $\bar{\mathbf{B}}_{r+1}$ may be an arbitrary row-stochastic matrix. In order to show that $\mathbf{A} \succ^{\mathrm{s}} \mathbf{B}$ it suffices to show that $\bar{\mathbf{A}} \succ^{\mathrm{s}} \bar{\mathbf{B}}$. So let $\mathbf{A}' \in \mathscr{S}(\bar{\mathbf{A}})$ and $\mathbf{B}' \in \mathscr{S}(\bar{\mathbf{B}})$. The matrices $\mathbf{A}'$ and $\mathbf{B}'$ may be written in the same form as $\bar{\mathbf{A}}$ and $\bar{\mathbf{B}}$ in terms of the submatrices $\mathbf{A}_1, \ldots, \mathbf{A}_{r+1}$ and $\mathbf{B}_1, \ldots, \mathbf{B}_{r+1}$, respectively. Let $k \leqslant r$. Then the j_kth column of $\mathbf{B}_k$ (as in $\bar{\mathbf{B}}_k$) is all ones and the other columns are zero. Then there is a row-stochastic matrix $\mathbf{X}_k$ such that $\mathbf{A}_k\mathbf{X}_k = \mathbf{B}_k$; just let the j_kth column of $\mathbf{X}_k$ be all ones while all other columns are zero. (The matrix equation holds as $\mathbf{A}_k$ is row-stochastic.) Furthermore, there is a row-stochastic matrix $\mathbf{X}_{k+1}$ such that $\mathbf{A}_{r+1}\mathbf{X}_{r+1} = \mathbf{B}_{r+1}$. This follows from Proposition 2.2 because $\mathbf{A}_{r+1}$ is a disjoint-row-support matrix. We let

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_1 \\ \vdots \\ \mathbf{X}_{r+1} \end{bmatrix}$$

and note that $\mathbf{X}$ is row-stochastic. Moreover, $\bar{\mathbf{A}}\mathbf{X} = \bar{\mathbf{B}}$ so $\bar{\mathbf{A}} \succ \bar{\mathbf{B}}$. This proves that $\mathbf{A} \succ^{\mathrm{s}} \mathbf{B}$ and the proof is complete. □

We see from the proof of Theorem 6.1 that when $\mathbf{A} \succ^{\mathrm{s}} \mathbf{B}$ both these matrices may be constructed in a certain manner (see the decomposition of $\bar{\mathbf{A}}$ and $\bar{\mathbf{B}}$). Moreover, the proof also indicates that the matrices in the majorization polytope have a certain structure. This structure is exploited next to obtain a characterization of the notion of strong majorization. First, we need a result on the dimension of the transportation polytope (the proof is easy and omitted).

Lemma 6.2. *Let* $\mathbf{a} \in \mathbb{R}^n$ *and* $\mathbf{b} \in \mathbb{R}^p$ *be nonnegative vectors with* $\sum_{j=1}^{n} a_j = \sum_{j=1}^{p} b_j$. *Let* n^+ *and* p^+ *be the number of positive elements in* $\mathbf{a}$ *and* $\mathbf{b}$, *respectively. Then the transportation polytope* $\mathscr{T}(\mathbf{a}, \mathbf{b})$ *has dimension* $(n^+ - 1)(p^+ - 1)$.

Let I_0 consist of those row indices $i \leqslant m$ such that $\mathrm{supp}(\mathbf{a}_i) \cap \mathrm{supp}(\mathbf{a}_k) = \emptyset$ for all $k \leqslant m, k \neq i$, these are the indices of rows of $\mathbf{A}$ that have disjoint support from every other row.

Corollary 6.3. *Let* $\mathbf{A} \in \mathscr{M}_{m,n}$ *and* $\mathbf{B} \in \mathscr{M}_{m,p}$. *Then* $\mathbf{A}$ *strongly majorizes* $\mathbf{B}$ *if and only if the following conditions hold*:

(i) *no pair of rows is in conflict,*

(ii) *if* $\mathbf{A}$ *has a zero column then* $p = 0$ *and*

(iii) *for each* $i \in I_0$ *at least one of the rows* $\mathbf{a}_i$ *and* $\mathbf{b}_i$ *is integral* (*so it is a unit vector*).

Proof. Assume that $\mathbf{A}$ strongly majorizes $\mathbf{B}$, so by Theorem 6.1 (i) holds. Then there are permutation matrices $\mathbf{P}$ and $\mathbf{Q}$ such that $\bar{\mathbf{A}} = \mathbf{PAQ}$ and $\bar{\mathbf{B}} = \mathbf{PB}$ have the form explained in the proof of Theorem 6.1. Thus, if $\mathbf{X} \in \mathscr{M}(\mathbf{A} \succ \mathbf{B})$, then $\mathbf{PAQQ}^{\mathrm{T}}\mathbf{X} = \mathbf{PB}$ so $\bar{\mathbf{A}}\mathbf{Z} = \bar{\mathbf{B}}$ where $\mathbf{Z} = \mathbf{Q}^{\mathrm{T}}\mathbf{X}$. We partition $\mathbf{Z}$ by

$$\mathbf{Z} = \begin{bmatrix} \mathbf{Z}_1 \\ \vdots \\ \mathbf{Z}_{r+1} \end{bmatrix}$$

and then $\bar{\mathbf{A}}\mathbf{Z} = \bar{\mathbf{B}}$ becomes $\bar{\mathbf{A}}_k\mathbf{Z}_k = \bar{\mathbf{B}}_k$ for $k = 1, \ldots, r+1$. Let first $k \leqslant r$. We recall the structure of the matrices involved: $\bar{\mathbf{B}}_k$ has a column of all ones and the remaining columns are zero, and $\bar{\mathbf{A}}_k$ has no zero column. Moreover, since both $\bar{\mathbf{A}}_i$ and $\mathbf{Z}_k$ are nonnegative, we deduce that also $\mathbf{Z}_k$ has a column of all ones and the remaining columns are zero. Next, consider $k = r + 1$. $\bar{\mathbf{A}}_{r+1}$ is a disjoint-row-support matrix while $\bar{\mathbf{B}}_{r+1}$ is arbitrary (but row-stochastic). Then the solution $\mathbf{Z}_{k+1}$ of $\bar{\mathbf{A}}_{k+1}\mathbf{Z}_{k+1} = \bar{\mathbf{B}}_{k+1}$ has the form (2) given in Proposition 2.2, say

$$\mathbf{Z}_{k+1} = \begin{bmatrix} \mathbf{Z}_{k+1,1} \\ \vdots \\ \mathbf{Z}_{k+1,p} \\ \mathbf{Z}_{k+1,0} \end{bmatrix},$$

where $\mathbf{Z}_{k+1,i} = \mathbf{D}(\bar{\mathbf{a}}_i)^{-1}\mathbf{Y}_i$ and $\mathbf{Y}_i \in \mathscr{T}(\bar{\mathbf{a}}_i, \mathbf{b}_i)$ for each i and $\mathbf{Z}_{k+1,0} \in \mathscr{M}_{n_0,p}$. Note that the last matrix corresponds to n_0 columns of $\mathbf{A}$ that are zero. Assume that $n_0 > 0$ and $p > 0$. Then $\mathscr{M}_{n_0,p}$ contains two matrices in different support classes and so does $\mathscr{M}(\mathbf{A} \succ \mathbf{B})$ (recall that $\mathbf{Z} = \mathbf{Q}^{\mathrm{T}}\mathbf{X}$ so $\mathbf{X}$ is obtained from $\mathbf{Z}$ by some permutation of its rows). But this contradicts that $\mathbf{A}$ strongly majorizes $\mathbf{B}$, so (ii) holds. Finally, if property (iii) were violated, the dimension of the transportation polytope $\mathscr{T}(\bar{\mathbf{a}}_i, \mathbf{b}_i)$ would be at least one (see Lemma 6.2), and then this polytope would have vertices with distinct support. This would again give solutions in the majorization polytope with different supports, a contradiction. Thus, property (iii) must hold. This, proves that (i)–(iii) all hold. The converse implication is shown using arguments as in the proof of Theorem 6.1. In fact, conditions (i)–(iii) imply that there is a unique $\mathbf{X}$ in $\mathscr{M}(\mathbf{A} \succ \mathbf{B})$ so then $\mathbf{A}$ strongly majorizes $\mathbf{B}$. We omit the details here. □

Thus, strong majorization is indeed a very strong requirement as the majorization polytope contains a unique element in that case. Note that this element is an integral row-stochastic matrix, so Proposition 3.1 gives a further description of the relation between the columns of $\mathbf{A}$ and $\mathbf{B}$ in this situation.

References

[1] T. Ando, Majorization and inequalities in matrix theory, Linear Algebra Appl. 199 (1994) 17–67.
[2] R.A. Brualdi, The doubly stochastic matrices of a vector majorization, Linear Algebra Appl. 61 (1984) 141–154.
[3] G.-S Cheon, Y.-H. Lee, The doubly stochastic matrices of a multivariate majorization, J. Korean Math. Soc. 32 (4) (1995) 857–867.
[4] G.-S. Cheon, S.-Z. Song, On the extreme points of the majorization polytope $\Omega_3(yx)$, Linear Algebra Appl. 269 (1) (1998) 47–52.
[5] G. Dahl, Matrix majorization, Linear Algebra Appl. 288 (1999) 53–73.
[6] B. Korte et al. (Eds.), Paths Flows and VLSI-Layout, vol. 9, Algorithms and Combinatorics, Springer, Berlin, 1990.
[7] S. Karlin, Y. Rinott, Comparison of measures, multivariate majorization, and applications to statistics, in: S. Karlin et al. (Ed.), Studies in Econometrics Time Series and Multivariate Statistics. (In commemoration of T.W. Anderson's 65th birthday), Academic Press, New York, 1983, pp. 465–489.
[8] R.B. Levow, A problem of Mirsky concerning nonsingular doubly stochastic matrices, Linear Algebra Appl. 5 (1972) 197–206.
[9] A.W. Marshall, I. Olkin, Inequalities: Theory of Majorization and its Applications, Academic Press, New York, 1979.

[10] E. Torgersen, Comparison of Statistical Experiments, vol. 36, Encyclopedia of Mathematics and its Applications, Cambridge University Press, Cambridge, 1992.
[11] V.A. Yemelichev, M.M. Kovalev, M.K. Kravtsov, Polytopes Graphs and Optimization, Cambridge University Press, Cambridge, 1984.