Linear Algebra and its Applications 585 (2020) 147–163

Contents lists available at ScienceDirect

Linear Algebra and its Applications

www.elsevier.com/locate/laa

Majorization for $(0,1)$-matrices ☆

Geir Dahl [a,*], Alexander Guterman [b,c,d], Pavel Shteyner [b,c,d]

[a] *Department of Mathematics, University of Oslo, Norway*
[b] *Department of Mathematics and Mechanics, Lomonosov Moscow State University, Moscow, Russia*
[c] *Moscow Center for Continuous Mathematical Education, Moscow, 119002, Russia*
[d] *Moscow Institute of Physics and Technology, Dolgoprudny, 141701, Russia*

A R T I C L E I N F O

Article history:
Received 16 June 2019
Accepted 28 September 2019
Available online 3 October 2019
Submitted by R. Brualdi

MSC:
06A06
15B51

Keywords:
Matrix majorization
Partial order
$(0,1)$-matrices

A B S T R A C T

This paper deals with the important notion of majorization. We study majorization for matrices, and focus on $(0,1)$-matrices. We prove several results concerning such matrix majorization orders on the set of (0,1)-matrices, including characterizations for certain orders, and separate sufficient and necessary conditions for the so-called matrix majorization order. Some of these results are of combinatorial nature.

1. Introduction

Majorization is an important order notion which arises in several areas of mathematics. An early treatment of majorization is the classical book by Hardy, Littlewood and Pólya

☆ The investigations of the second and the third authors were financially supported by RSF grant 16-11-10075.

* Corresponding author.
E-mail addresses: geird@math.uio.no (G. Dahl), guterman@list.ru (A. Guterman), pashteiner@ya.ru (P. Shteyner).

https://doi.org/10.1016/j.laa.2019.09.038

[6], and today a main reference is [10] which is a comprehensive treatment of both the theory and many applications in e.g., matrix theory, statistics, and combinatorics. The book [1] treats majorization in connection with combinatorial classes of matrices. Majorization is well-known as an order relation for vectors, but it has been extended, for example, to orders for matrices. There are different ways of doing such an extension. In our recent paper [4] we presented different matrix majorization orders, and introduced and studied its generalizations for classes of matrices. Matrix majorization has some of its roots in a functional analytic approach to basic notions in statistics, in connection with the theory of comparison of statistical experiments [14]. The theory deals with the information content in families of (probability) measures, as abstract representations of statistical experiments. These ideas were further developed in a matrix theoretic setting in [2] and [3], where related polytopes were studied.

The following notations will be used in the subsequent discussion. Let $M_{m,n}$ be the space of all real $m \times n$ matrices (where we write M_n if $m = n$). For a vector $x \in \mathbb{R}^n$ we let $x_{[j]}$ denote the jth largest number among the components of x. If $x, v \in \mathbb{R}^n$ we say that x is *majorized* by v, denoted by $x \preceq v$ (or $v \succeq x$), provided that $\sum_{j=1}^{k} x_{[j]} \leq \sum_{j=1}^{k} v_{[j]}$ for $k = 1, 2, \ldots, n$ where there is equality for $k = n$. A matrix is called *row stochastic* if it is (component-wise) nonnegative and each row sum is 1. The set of all such $n \times n$ matrices is denoted by Ω_n^{row}. If in addition each column sum is 1, the matrix is called *doubly stochastic*. The set Ω_n of all $n \times n$ doubly stochastic matrices is a polytope whose extreme points are the permutation matrices (the Birkhoff–von Neumann theorem), see [1] for an in-depth survey of properties of Ω_n, some related work may be found in [9], [13].

In this paper we study classical matrix majorizations and mainly focus on these notions for $(0, 1)$ matrices. This type of constraint on the matrix entries leads to some interesting questions and properties, some of combinatorial type. In particular, we find certain necessary and sufficient conditions for such matrices to be majorized and algorithms detecting the majorization.

There are two main motivations for the study of matrix majorization for $(0, 1)$-matrices. First, it is of interest to see if this restriction to the subclass of $(0, 1)$-matrices leads to simpler characterizations of the majorization order in question. This can give better insight into this order, and also possibly lead to new types of questions for larger matrix classes. Secondly, $(0, 1)$-matrices are essential to represent combinatorial objects, and therefore one may look at the meaning of such a matrix majorization order for two combinatorial objects (each associated with a $(0, 1)$-matrix). We leave this approach for future investigations, although we give a couple of such examples in this paper.

Our paper is organized as follows: Section 2 summarizes several types of matrix majorizations and discusses some connections between these notions. In Section 3 we characterize weak, directional and strong majorizations of $(0, 1)$-matrices. Section 4 is devoted to investigations of matrix majorization on $(0, 1)$ matrices.

For a matrix A its j'th column is denoted by $A^{(j)}$ and its i'th row is denoted by $A_{(i)}$. The set of all real $m \times n$ matrices with every element in $\{0, 1\}$ is denoted by $M_{m,n}(0, 1)$.

Let $R(A)$ denote the set of rows of a matrix A. The set of its columns is denoted by $C(A)$.

2. Different notions of matrix majorization

Vectors in $\mathbb{R}^n$ are considered as column vectors, and identified with corresponding n-tuples. The transpose of a matrix A is denoted by A^t. The j'th unit vector is denoted by e_j while e denotes the all 1s vector (of suitable length). The convex hull of a set $S \subseteq \mathbb{R}^n$ is denoted by $\text{conv}(S)$.

Let $\preceq$ denote a matrix majorization order. There are several such notions, and we mention some of them below. Let $A, B \in M_{m,n}$.

- Directional majorization: $A \preceq^d B$ when $Ax \preceq Bx$ for all $x \in \mathbb{R}^n$.
- Weak matrix majorization: $A \preceq^{wm} B$ when there is a row-stochastic matrix $X \in M_m$ such that $A = XB$.
- Strong majorization: $A \preceq^s B$ when there is an $X \in \Omega_m$ such that $A = XB$.
- Doubly stochastic majorization: $A \preceq^{ds} B$ when there is $X \in \Omega_n$ such that $A = BX$.
- Matrix majorization: $A \preceq^m B$ when there is a row-stochastic matrix $X \in M_n$ such that $A = BX$.

Classical vector majorization is a special case of strong majorization where matrices A and B have a single column. Also, doubly stochastic majorization is a special case of matrix majorization.

Matrix majorization was introduced and studied in [2], and has later been investigated in the linear algebra area. Also, recently in a *Nature Communications* paper [5], matrix majorization was the starting point for a generalization used to study basic problems in quantum mechanics and quantum thermodynamics.

Proposition 2.1. *Let $A, B \in M_{m,n}$.*

(i) *$A \preceq^s B$ if and only if $A^t \preceq^{ds} B^t$.*
(ii) *$A \preceq^d B$ implies that $A \preceq^{wm} B$.*
(iii) *$A \preceq^{wm} B$ does not imply $A \preceq^d B$ in general for $m \geq 2$.*
(iv) *$A \preceq^s B$ implies that $A \preceq^d B$.*
(v) *$A \preceq^d B$ also does not imply $A \preceq^s B$ for $m \geq 4$.*
(vi) *$A \preceq^{wm} B$ if and only if $R(A) \subseteq \text{conv}(R(B))$.*

Proof. (i) follows from the definitions. (ii) and (vi) follow by Proposition 3.3 in [11]. (iii) follows by [11, Example 1]. (iv) follows from the well-known fact for vector majorization, the Hardy-Littlewood-Pólya theorem, saying that $u \preceq v$ if and only if $u = Xv$ for some doubly stochastic matrix X. For (v) see [7, page 98(6)] and also [8]. □

Proposition 2.2. *There exist* $A, B \in M_{m,n}$ *such that*

(*i*) $A \preceq^s B$ *does not imply* $A \preceq^{ds} B$ *and conversely,*
(*ii*) *let* $\preceq$ *be either* $\preceq^s$, $\preceq^d$, *or* $\preceq^{wm}$, *then* $A \preceq B$ *does not imply* $C(A) \subseteq \text{conv}(C(B))$,
(*iii*) *let* $\preceq$ *be either* $\preceq^{ds}$ *or* $\preceq^m$, *then* $A \preceq B$ *does not imply* $R(A) \subseteq \text{conv}(R(B))$.

Proof. Let us consider $D = \begin{bmatrix} 3/4 & 0 & 1/4 \\ 0 & 1 & 0 \\ 1/4 & 0 & 3/4 \end{bmatrix} \in \Omega_n$, $B = \begin{bmatrix} 3 & -1 \\ 1 & 3 \\ 2 & 1 \end{bmatrix}$, $A = \begin{bmatrix} 11/4 & -1/2 \\ 1 & 3 \\ 9/4 & 1/2 \end{bmatrix}$.

Since $DB = A$, $A \preceq^s B$ and as a consequence $A \preceq^d B$ and $A \preceq^{wm} B$. It is easy to verify that $(11/4, 1, 9/4) \notin \text{conv}(C(B))$.

It follows that $R(A^t) \nsubseteq \text{conv}(R(B^t)) \Rightarrow A^t \not\preceq^{wm} B^t \Rightarrow A \not\preceq^{ds} B$. Thus we have proved (ii) and direct part of (i).

To conclude the proof one should do the same calculations with A^t, B^t and D^t. In this case $A^t = B^t D^t$, so $A^t \preceq^{ds} B^t$ and $A^t \preceq^m B^t$, so, (iii) and the inverse part of (i) are proved. □

Let $A, B \in M_{m,n}$, and let $\preceq$ be either $\preceq^{ds}$, $\preceq^m$, $\preceq^{wm}$ or $\preceq^s$. Then one can check if $A \preceq B$ holds in polynomial time using linear optimization. In fact, for each of these orders we look for a row-stochastic or doubly stochastic matrix X that satisfies a (finite) system of linear equations, namely $AX = B$ or $XA = B$ (depending of the order $\preceq$). Thus, one needs to decide if a certain system of linear inequalities has a solution, and this can be done efficiently by linear optimization. Moreover, the above holds for $\preceq^d$ too, since it can be reduced to weak majorization of finite number of matrices (see [11, Corollary 3.13]). Finally, we mention that a characterization of matrix majorization in terms of sublinear functions was given in [2].

3. Majorizations of types $\preceq^{wm}$, $\preceq^d$, $\preceq^s$ for $(0,1)$-matrices

Lemma 3.1. *Let* $v, w_1, w_2, \ldots, w_m$ *be* $(0,1)$*-vectors of the same size such that* $v \in \text{conv}(\{w_1, w_2, \ldots, w_m\})$. *Then* $v \in \{w_1, w_2, \ldots, w_m\}$.

Proof. Let $v = \sum_{j=1}^m \lambda_j w_j$ where $\lambda_j \geq 0$ for each j and $\sum_{j=1}^m \lambda_j = 1$. Then $v = \sum_{j \in J} \lambda_j w_j$ where $J = \{j : \lambda_j > 0\}$ is nonempty. We claim that $w_j = v$ for each $j \in J$. Indeed, if $v_i = 0$ for some i, then $(w_j)_i = 0$ for each $j \in J$ (otherwise $0 = v_i = \sum_{j \in J} \lambda_j (w_j)_i > 0$; a contradiction). If $v_i = 1$ for some i, then $(w_j)_i = 1$ for each $j \in J$ (otherwise $1 = v_i = \sum_{j \in J} \lambda_j (w_j)_i < 1$; a contradiction). The claim follows, and therefore the lemma. □

Corollary 3.2. *Let $v_1, v_2, \dots, v_n$ and $w_1, w_2, \dots, w_m$ be $(0,1)$-vectors of the same size such that $\{v_1, v_2, \dots, v_n\} \subseteq \operatorname{conv}(\{w_1, w_2, \dots, w_m\})$. Then $\{v_1, v_2, \dots, v_n\} \subseteq \{w_1, w_2, \dots, w_m\}$.*

Proof. We apply Lemma 3.1 to each of $v_1, v_2, \dots, v_n$. □

This corollary leads to a nice characterization of $\preceq^{wm}$ for $(0,1)$-matrices.

Proposition 3.3. *Let A, B be $(0,1)$-matrices of the same size. Then the following statements are equivalent:*

(*i*) $A \preceq^{wm} B$,
(*ii*) $R(A) \subseteq R(B)$,
(*iii*) $A = RB$ *for a $(0,1)$-matrix R with exactly one element equal to 1 in every row.*

Proof. If (i) holds, then by Proposition 2.1.(vi), $R(A) \subseteq \operatorname{conv}(R(B))$. Hence by Corollary 3.2 we obtain $R(A) \subseteq R(B)$, so (ii) holds. Moreover, (ii) implies (iii) by simple matrix computations, and (iii) implies (i), as such matrix R is row-stochastic. □

We now turn to strong and directional majorizations.

Lemma 3.4. *Let $A, B \in M_{m,n}(0,1)$ such that $A \preceq^d B$. Then the number of 1's in $A^{(j)}$ coincides with the number of 1's in $B^{(j)}$ for each $j \le n$.*

Proof. By the definition of the directional majorization, $Ax \preceq Bx$ for each x. By letting $v = e_j$ we obtain $Ae_j = A^{(j)} \preceq B^{(j)} = Be_j$. In particular, it follows that $\sum_{i=1}^m A_i^{(j)} = \sum_{i=1}^m B_i^{(j)}$ which gives the result since all entries are either 0 or 1. □

Let us characterize the majorization order $A \preceq^d B$ for $(0,1)$-matrices A and B by using purely combinatorial arguments.

Theorem 3.5. *Let $A, B \in M_{m,n}(0,1)$. Then*

(*i*) $A \preceq^d B$ *if and only if A is a row permutation of B, i.e., $A = PB$ for some permutation matrix P,*
(*ii*) $\preceq^d$ *is an equivalence relation on $M_{m,n}(0,1)$.*

Proof. For $x \in \{0,1\}^n$ let $x_A = |\{1 \le i \le m : A_{(i)} = x\}|$ and $x_B = |\{1 \le i \le m : B_{(i)} = x\}|$, the cardinalities of the sets of rows equal to x in A and B, correspondingly. By construction we get the identity

$$m = \sum_{x \in \{0,1\}^n} x_A = \sum_{x \in \{0,1\}^n} x_B. \tag{1}$$

Hence, in order to show that $x_A = x_B$ it is sufficient to show that $x_A \leq x_B$ for every $x \in \{0,1\}^n$.

For an arbitrary $x \in \{0,1\}^n$ define the (column) vector $v_x \in \mathbb{R}^n$ as follows: $(v_x)_i = 1$, if $x_i = 1$ and $(v_x)_i = -1$, if $x_i = 0$. Let $y \in \{0,1\}^n$. It is straightforward to see that $y^t v_x \leq x^t v_x$ and $y^t v_x = x^t v_x = x^t e$ if and only if $x = y$. It follows that the maximal possible elements of Av_x and Bv_x are $x^t e$ and the number of elements equal to $x^t e$ in Av_x (respectively, Bv_x) is precisely x_A (respectively, x_B). Since $Av_x \preceq Bv_x$ we may conclude that $x_A \leq x_B$. Indeed, otherwise $\sum_{i=1}^{x_A} (Av_x)_{[i]} > \sum_{i=1}^{x_A} (Bv_x)_{[i]}$, a contradiction.

By the equality (1) we have proved that $x_A = x_B$ for every $x \in \{0,1\}^n$. It follows that matrices A and B are equal up to permutation of rows.

Conversely, if A is a row permutation of B, then Az is a permutation of Bz, for any $z \in \mathbb{R}^n$, so $Az \preceq Bz$, and therefore $A \preceq^d B$.

Finally, since permutation matrices are invertible, it follows from the condition (i) that $A \preceq^d B$ if and only if $B \preceq^d A$, so this order is an equivalence relation. □

Corollary 3.6. *For* $(0,1)$*-matrices* $\preceq^d$ *and* $\preceq^s$ *are equivalent.*

Proof. $\preceq^s$ implies $\preceq^d$ for all matrices by Proposition 2.1.(iv). For (0,1)-matrices the converse implication follows from the characterization obtained in Theorem 3.5.(i). □

4. Matrix majorization for (0, 1)-matrices

We now consider the matrix majorization order, denoted by $\preceq^m$. This concept was introduced and studied in [2]. In particular, that paper contains characterizations of $A \preceq^m B$ when A and B are $(0,1)$-matrices with one, resp. two, 1's in each row. Matrix majorization is a more complex concept than weak matrix majorization. We show that this majorization differs from the others even in the case of $(0,1)$ matrices.

Observe that matrix majorization, unlike weak majorization, does not allow any analogues of Proposition 3.3, Lemma 3.4, and characterizations above. Namely, it can not be described in terms of row/column inclusion, as the following examples show: the first of them shows that row/column inclusion does not follow from $A \preceq^m B$, and the other one shows that converse implication does not hold as well.

Example 4.1. Let $A = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$, $R = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}$. Then $A = BR$, so $A \preceq^m B$. As we can see, $A^{(3)} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \notin C(B)$. Moreover, $B^{(1)} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \notin C(A)$ and $R(A) \cap R(B) = \emptyset$, so $A \not\preceq^{wm} B$. □

Example 4.2. Let $A = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}$. Then $Ae = Be$ and $C(A) \subseteq C(B)$. But it is easy to verify that $A \not\preceq^m B$. Indeed, suppose that there exists $R = \begin{bmatrix} a & b & \dots & \dots \\ d & e & \dots & \dots \\ g & h & \dots & \dots \\ j & k & \dots & \dots \end{bmatrix}$ such that $A = BR$. Since

$$BR = \begin{bmatrix} a+g & b+h & \dots & \dots \\ a+d & b+e & \dots & \dots \\ a+d+g+j & b+h+e+k & \dots & \dots \end{bmatrix},$$

it follows that $a + g = a + d = a + d + g + j = 1$. Then $d = g = j = 0$ and $a = 1$. For similar reasons, $b = 1$, and then R is not a row-stochastic matrix. □

The next example shows that, in the case of matrix majorization, one may have $A \preceq^m B$ for some $(0,1)$-matrices, but, still, the corresponding stochastic matrix R cannot be chosen from the set of $(0,1)$-matrices.

Example 4.3. There exist $A, B \in M_{m,n}(0,1)$ such that $A \preceq^m B$ and there is no $(0,1)$-matrix $R \in \Omega_n^{row}$ with $A = BR$.

Let $A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$.

Then one can check that $B^{-1} = \begin{bmatrix} 0.5 & 0.5 & -0.5 \\ 0.5 & -0.5 & 0.5 \\ -0.5 & 0.5 & 0.5 \end{bmatrix}$ and $B^{-1}A = \begin{bmatrix} 0.5 & 0.5 & 0 \\ 0.5 & 0.5 & 0 \\ 0.5 & 0.5 & 0 \end{bmatrix} =$ R. Obviously, this is the only R such that $A = BR$ and, since $R \in \Omega_n^{row}$, $A \preceq^m B$. □

Now we show that it is possible that $A \preceq^m B$, and $A = BR$ where B and R are $(0,1)$ matrices, but $A \notin M_{m,n}(0,1)$.

Example 4.4. There exist $(0,1)$-matrices B, R such that $A = BR$ is not a $(0,1)$-matrix.

Let $B = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$ and $R = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$. Then $A = BR = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}$ and $A \preceq^m B$. □

As mentioned, doubly stochastic majorization is a particular case of matrix majorization. We already know that $A \preceq^{ds} B$ if and only if $A = BP$ (i.e., A is a permutation of columns of B). So, below we mainly investigate those pairs $A, B \in M_{m,n}$ that $A \preceq^m B$, but $A \not\preceq^{ds} B$.

Lemma 4.5. *Let $A, B \in M_{m,n}(0,1)$ and $A \preceq^m B$. Then for every $i = 1, 2, \dots, m$ the number of 1's in the i'th row of A is equal to the number of 1's in the i'th row of B.*

Proof. By assumption, $A = BR$ for some $R \in \Omega_n^{row}$. Then $Ae = BRe = Be$ where $e = [1\ 1\ \cdots\ 1]^t$. Since both matrices are $(0,1)$, the i-th entry of Ae is the number of 1s in the i-th row of A. The same is for B. Hence, the result follows. □

Lemma 4.6. *Let $A, B \in M_{m,n}(0,1)$. Assume $A \preceq^m B$, but $A \not\preceq^{ds} B$. Then*

(i) there exists $R \in \Omega_n^{row}$ satisfying $A = BR$ such that R contains a zero column, and for each column sum c_i of R it holds that either $c_i = 0$ or $c_i \geq 1$,
(ii) if B does not contain a zero column, then for any $R \in \Omega_n^{row}$ satisfying $A = BR$ it holds that R contains a zero column, and for each column sum c_i of R it holds that either $c_i = 0$ or $c_i \geq 1$,
(iii) A contains a zero column.

Proof. Suppose that $A = BR$. It follows that the j'th column of A is a linear combination of columns of B with coefficients from j'th column of R. Since $R \in \Omega_n^{row}$, the sum of all elements in R is n. Since, by assumption, $A \not\preceq^{ds} B$, there is j such that the j'th column sum $c_j < 1$ and there is k such that the k'th column sum $c_k > 1$.

If $c_j < 1$ then $A^{(j)}$ is a zero column. Indeed, each element of $A^{(j)}$ is less than or equal to $c_j < 1$. Hence, it is 0.

(i) We are going to modify R in order to construct $R' \in \Omega_n^{row}$ such that $A = BR'$ and ${R'}^{(j)}$ is zero. Fix some k such that $c_k > 1$. Suppose that $r_{pj} \neq 0$. Since $0 = A^{(j)} = \sum_{i=1}^{n} r_{ij} B^{(i)} \geq r_{pj} B^{(p)}$, we obtain $B^{(p)} = 0$. It follows that change of $R_{(p)}$ does not affect A. Indeed, consider arbitrary a_{xy}. Since $A = BR$, $a_{xy} = \sum_{z=1}^{n} b_{xz} r_{zy} = \sum_{z \neq p} b_{xz} r_{zy} + b_{xp} r_{py} = \sum_{z \neq p} b_{xz} r_{zy}$.

We consider the matrix R', which is obtained from R by changing r_{pk} to $r_{pk} + r_{pj}$ and r_{pj} to 0, recall that k is some fixed index satisfying $c_k > 1$. We do the same for the rest of nonzero elements in $R^{(j)}$. Finally, we obtain R' such that $A = BR'$, ${R'}^{(j)}$ is a zero column, $c'_k > 1$ and $R^{(l)} = {R'}^{(l)}$ for $l \neq j, k$.

We repeat this procedure for every q with $0 < c_q < 1$. After several such substitutions we obtain R' such that for every q either $c'_q = 0$ or $c'_q \geq 1$ and R' contains a zero column, as required.

(ii) Suppose that B does not contain a zero column. Let j be such that $c_j < 1$. By above, $A^{(j)}$ is zero. Since $A^{(j)} = \sum_{i=1}^{n} r_{ij} B^{(i)}$ and all summands are non-negative, it follows that $r_{ij} = 0$ for all i.

Finally, if $c_j < 1$ then $c_j = 0$ and R contains a zero column.

(iii) By (i) we obtain that R' contains a zero column and, as a consequence, $A = BR'$ also contains a zero column. $\square$

Remark 4.7. Note that in general, if there is a zero column in B, then item (ii) above may not hold, i.e., there is an $R \in \Omega_n^{row}$ without zero columns satisfying $A = BR$. For instance, let

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Then

$$A = B \begin{bmatrix} 0.5 & 0 & 0.5 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{bmatrix}. \quad \square$$

Lemma 4.8. *Let A, B be $(0,1)$-matrices, and R be a nonnegative matrix such that $A = BR$ with $r_{ij} \neq 0$ for some i, j. Then $A^{(j)} \geq B^{(i)}$ (element-wise).*

Proof. Observe that $A^{(j)} = r_{ij}B^{(i)} + v$, where v is a certain nonnegative vector. Indeed, it is obtained by linear combination of nonnegative columns of B with nonnegative scalars r_{kj}. $\square$

Lemma 4.9. *Let A, B be such that $R \in \Omega_n^{row}$ with $A = BR$. Moreover, suppose that $A_{(1)} = B_{(1)} = \sum_{j=1}^{k} e_j^t$ for some $k \leq n$. Then $r_{ij} = 0$ whenever $i \leq k$ and $j > k$.*

Proof. The result follows from the fact that for $i = 1$ we have $a_{ij} = \sum_{q=1}^{n} b_{1q} r_{qj} = \sum_{q=1}^{k} b_{1q} r_{qj} = \sum_{q=1}^{k} r_{qj} = 0$ for $j > k$. $\square$

For an $m \times n$ matrix B let $\text{supp}(B_{(i)})$ denote the support of the i'th row of B, i.e., the set of column indices j such that (i, j)'th entry of B is non-zero. We consider the support as a subset of $\{1, 2, \ldots, n\}$. For a subset $S \subseteq \{1, 2, \ldots, n\}$ and a matrix $A \in M_{m,n}$ we let $A^{(S)}$ denote the submatrix induced by the columns with indices in S.

We are now ready to formulate and prove the following necessary condition concerning matrix majorization.

Lemma 4.10. *Let $A, B \in M_{m,n}(0,1)$ be such that $A \preceq^m B$. Then for every $i = 1, 2, \ldots, m$ there exists some permutation matrix P such that the following inequality holds element-wise: $A^{(\text{supp}(A_{(i)}))} \geq B^{(\text{supp}(B_{(i)}))}P$.*

Proof. Let $i \leq m$. Assume that there are k 1's in (the row) $A_{(i)}$, so, by Lemma 4.5, there are exactly k 1's in $B_{(i)}$. Without loss of generality assume that they are all in the first k columns in A and B and also suppose that $i = 1$.

Consider the submatrix D of R of size $k \times k$ consisting of the first k rows and columns of R. Now we may write B and R as block matrices: $B = \begin{bmatrix} X & Y \end{bmatrix}$ such that X is of size $m \times k$ with first row of 1s and Y is of size $m \times (n-k)$ with first row of 0s. By Lemma 4.9 $R = \begin{bmatrix} D & O \\ U & V \end{bmatrix}$ where D is a square matrix of order k, O is zero matrix of size $k \times (n-k)$ and U, V are arbitrary matrices of corresponding sizes. It is easy to verify that D is doubly stochastic. Indeed, by the proof above for any j the sum of elements of D in the j'th column equals $a_{1j} = 1$.

Finally, we conclude the proof using Lemma 4.8. It implies that if $r_{js} > 0$ then $A^{(s)} \geq B^{(j)}$. By Birkhoff–von Neumann theorem (see, for example, [10, Theorem 2.A.2]) $D = \lambda P + Q$ for some permutation matrix P and nonnegative Q. It follows that $A^{(\mathrm{supp}(A_{(i)}))} \geq B^{(\mathrm{supp}(B_{(i)}))} P$. □

Remark 4.11. The following example shows that the statement of the lemma above holds only for submatrices induced by supports of some fixed rows $A_{(k)}$ and $B_{(k)}$ respectively: Let A $= \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. It is easy to see that $A \preceq^m B$. In this case $\mathrm{supp}(A_{(i)}) = \{1\}$ and $\mathrm{supp}(B_{(i)}) = \{i\}$. Obviously $A^{(1)} \geq B^{(i)}$, but for any $M, N \subseteq \{1, 2, 3\}$ with $|M| = |N| > 1$ we have that $A^{(M)} \not\geq B^{(N)} P$ for any permutation matrix P. □

Lemma 4.12. *There is a polynomial algorithm to verify whether necessary condition in Lemma 4.10 holds.*

Proof. It is easy to see that this problem can be solved by an algorithm for finding a maximum matching in a bipartite graph. We consider columns of B with 1's in the chosen row (exactly k 1's) as vertices of the first part of the graph and columns of A with 1's in the same row as vertices in the second part. We write an edge between vertices $A^{(i)}$ and $B^{(j)}$ if $A^{(i)} \geq B^{(j)}$. After that we find maximum matching and if its size equals k, then the condition holds and does not hold otherwise. The complexity of this algorithm is not more than $O(mn^3)$. □

We observe that the necessary conditions given in Lemmas 4.5, 4.6 and 4.10, even combined, are still not sufficient for the matrix majorization as the following example shows.

Example 4.13. Indeed, it is easy to verify that

$$\begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \end{bmatrix} \not\preceq^m \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix}.$$

Suppose that there exists a row-stochastic matrix $R = \{r_{ij}\}$ such that $A = BR$. It follows directly from this equality that:

- $r_{11} + r_{21} = r_{11} + r_{41} = 1$ and $r_{21} + r_{31} = r_{31} + r_{41} = 0$. It follows that $r_{11} = 1$ and $r_{21} = r_{31} = r_{41} = r_{12} = r_{13} = r_{14} = 0$.
- $r_{12} + r_{22} = 1$ so $r_{22} = 1$. Also $r_{22} + r_{32} = 1$ so $r_{32} = 0$. Moreover, $r_{32} + r_{42} = 1$. It follows that $r_{42} = 1$ and $a_{42} = r_{12} + r_{42} = 1$, a contradiction. □

Let $A, B \in M_{m,n}$. We define a $(0, 1)$-matrix $Z = Z(A, B) = [z_{kj}]$ by the following rule. For a fixed index j we consider the set I of rows of A having 0 in the j'th column. If there exists $i \in I$ such that $b_{ik} = 1$, then $z_{kj} = 0$, otherwise it is 1. In other words, $z_{kj} = 0$ iff $k \in \bigcup_{i:a_{ij}=0} \operatorname{supp}(B_{(i)})$, $j = 1, 2, \ldots, n$.

Lemma 4.14. *Let $A, B \in M_{m,n}$ and assume that $A \preceq^m B$. Then every matrix $R = [r_{ij}] \in \Omega_n^{row}$ with $A = BR$ satisfies*

$$R \leq Z,$$

where $\leq$ denotes elementwise order here, namely, $U = (u_{ij}) \leq V = (v_{ij})$ iff $v_{ij} = 0$ for some i, j implies that $u_{ij} = 0$. In particular, Z has no zero row.

Proof. When $a_{ij} = 0$, and $A = BR$, it follows that the inner product of the i'th row in B and the j'th column in R is zero, and therefore

$$\sum_{k \in \operatorname{supp}(B_{(i)})} r_{kj} = 0.$$

Since R is nonnegative, it follows that $r_{kj} = 0$ for each $k \in \operatorname{supp}(B_{(i)})$. So, $r_{ij} = 0$ whenever $z_{ij} = 0$, and the first part follows. The last statement follows from the first part, since R cannot have a zero row being row-stochastic. □

Based on this lemma it is sometimes very easy to check that A is not majorized by B as the following two examples show.

Example 4.15. Let

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \; B = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}.$$

Then $Z = Z(A, B)$ is given by

$$Z = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

So, Z has a zero row and by Lemma 4.14 we obtain $A \not\preceq^m B$. □

Example 4.16. Let

$$A = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \end{bmatrix}, \; B = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix}.$$

Then $Z = Z(A, B)$ is given by

$$Z = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

For instance, $a_{13} = 0$, and $\mathrm{supp}(B_{(1)}) = \{1, 2\}$, so $z_{13} = z_{23} = 0$. So, a row-stochastic $R = [r_{ij}]$ with $A = BR$ must have the form

$$R = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & a & a' & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

for some $0 \le a \le 1$ and $a' = 1 - a$. But if $a > 0$, then $(BR)_{22} > 1$, and if $a' > 0$, then $(BR)_{33} > 1$; a contradiction. It follows that $A \not\preceq^m B$. □

Let n be a natural number. An *ordered n-partition* of the set $N_n := \{1, 2, \dots, n\}$ is an n-tuple $\pi = (S_1, S_2, \dots, S_n)$ of sets, where $S_1, S_2, \dots, S_n$ is a partition of the set N_n into the n disjoint subsets, and some sets S_i can be empty. Let $\mathcal{P}_n$ denote the set of all such ordered n-partitions. For instance, there are 27 ordered partitions in $\mathcal{P}_3$, one can consider $\pi_1 = (\{1, 3\}, \emptyset, \{2\})$ and $\pi_2 = (\emptyset, \emptyset, \{1, 2, 3\})$ for example.

Let $B = [B^{(1)} B^{(2)} \cdots B^{(n)}]$ be an $m \times n$ matrix, where $B^{(j)}$ is its j'th column, and let $\pi = (S_1, S_2, \dots, S_n) \in \mathcal{P}_n$. Define the $m \times n$ matrix

$$B_\pi = \left[\sum_{j \in S_1} B^{(j)} \quad \sum_{j \in S_2} B^{(j)} \quad \cdots \quad \sum_{j \in S_n} B^{(j)} \right]$$

where the sum is defined to be the zero vector if the summation set S_i is empty.

We now connect these notions to matrix majorization by introducing certain new majorization. Let A and B be real $m \times n$ matrices. We define the notion of $(0, 1)$-matrix majorization as follows: A is $(0, 1)$-*matrix majorized* by B, and we write $A \preceq^m_{0,1} B$, when there is an *integral* row-stochastic matrix $R \in M_n$ such that $A = BR$. Clearly, such R is a $(0, 1)$-matrix with exactly one 1 in every row. Let $\pi, \pi' \in \mathcal{P}_n$. We write $\pi \succeq \pi'$ if π' is a refinement of π, i.e., it is obtained by further partitioning of some of the subsets, and no permutations of subsets are allowed. This is a partial order on $\mathcal{P}_n$, its maximal elements are $(N_n, \emptyset, \dots, \emptyset)$ and its permutations, and its minimal elements are $(\{1\}, \{2\}, \dots, \{n\})$ and its permutations.

Theorem 4.17. *Let A and B be real $m \times n$ matrices. Then the following holds.*

- *(i) If $A \preceq^m_{0,1} B$, then $A \preceq^m B$.*
- *(ii) $A \preceq^m_{0,1} B$ if and only if $A = B_\pi$ for some $\pi \in \mathcal{P}_n$.*
- *(iii) $A \preceq^m B$ if and only if $A \in \operatorname{conv}(\{B_\pi : \pi \in \mathcal{P}_n\})$.*
- *(iv) If $\pi \succeq \pi'$, then $B_\pi \preceq^m B_{\pi'}$ for every $B \in M_{m,n}$.*

Proof. (i): This is clear as the integrality requirement on R makes $(0, 1)$-matrix majorization stronger than ordinary matrix majorization.

(ii): Assume $A \preceq^m_{0,1} B$, so $A = BR$ for an integral row-stochastic matrix $R = [r_{ij}]$. Let $S_j = \{i : r_{ij} = 1\}$ for each $1 \le j \le n$. Since R is row-stochastic, the sets S_j $(j \le n)$ are pairwise disjoint, and $\pi = (S_1, S_2, \dots, S_n)$ is an ordered partition. Moreover, the j'th column of BR is

$$\sum_{k \in S_j} B^{(k)}$$

and therefore $BR = B_\pi$, as desired. The converse implication also follows from this computation.

(iii): This was shown in [2] (Corollary 3.2), and essentially follows by convexity and the fact that the set of row-stochastic $m \times n$ matrices is the convex hull of integral row-stochastic $m \times n$ matrices, since this can be seen as the product (or direct sum) of the similar statement for each row separately, and this is just the standard simplex.

(iv): Assume $\pi, \pi' \in \mathcal{P}_n$ and $\pi \succeq \pi'$, and let $\pi = (S_1, S_2, \ldots, S_n)$, $\pi' = (S'_1, S'_2, \ldots, S'_n)$. Since $\pi \succeq \pi'$, for every nonempty S'_i there is unique S_j such that $S'_i \subseteq S_j$. Let $R \in M_n$. If $\emptyset \neq S'_i \subseteq S_j$, then set $r_{ij} = 1$, and otherwise set $r_{ij} = 0$. If $S'_k = \emptyset$, then set $R_{(k)} = e^t_1$. It is easy to see that $B_\pi = B_{\pi'} R$ and R is row stochastic. □

Note that Example 4.3 shows that the converse of the implication in (i) is not true even for $(0, 1)$-matrices A and B.

We also remark that property (iii) of the theorem can be seen as a generalization of Rado's theorem for vector majorization, see [10, Corollary 2.B.3]. Moreover, the set

$$P_B := \text{conv}(\{B_\pi : \pi \in \mathcal{P}_n\})$$

is a polytope, so it can be described in terms of its facets, i.e., as a solution of a finite system of linear inequalities. Namely, there exist $t \in \mathbb{N}$, matrices $C_i \in M_{m,n}$ and constants $b_i \in \mathbb{R}$ such that

$$P_B := \{A \in M_{m,n} : \langle C_i, A\rangle \leq b_i \text{ for } i = 1, \ldots, t\}.$$

The problem is to find C_i and b_i for each i, and this can only be done in very special cases. But the general fact here is that one has the characterization of matrix majorization: $A \preceq^m B$ if and only if $\langle C_i, A\rangle \leq b_i$ for $i = 1, \ldots, t$.

Corollary 4.18. *Let $B \in M_{m,n}(0, 1)$. Then the set of $m \times n$ $(0, 1)$-matrices A such that $A \preceq^m_{0,1} B$ consists of all matrices*

$$A = B_\pi$$

where $\pi = (S_1, S_2, \ldots, S_n) \in \mathcal{P}_n$ such that the vectors $B^{(j)}$ for $j \in S_i$ have pairwise disjoint supports $(i \leq n)$.

Proof. This follows from the definition of $(0, 1)$-matrix majorization and the assumptions that both A and B are $(0, 1)$-matrices: the condition of the theorem assures that the matrix B_π is a $(0, 1)$-matrix. □

The following example shows a combinatorial significance of Corollary 4.18.

Example 4.19. Let $G = (V, E)$ be a graph and B be its vertex-edge incidence matrix. Thus the columns of B correspond to the edges of G, and each such column has two nonzero entries, both equal to 1, and they are in the rows corresponding to the end-vertices of the edge.

Consider Corollary 4.18 and let A be a $(0,1)$-matrix satisfying $A \preceq^m_{0,1} B$. Then the nonzero columns of A correspond to a partition of the edge set E into matchings M_k $(k \le t)$ such that the column corresponding to M_k has 1's in rows corresponding to $V(M_k)$, i.e., the vertices covered by M_k. □

Item (iii) of Theorem 4.17 can be also reformulated in the following geometrical way.

Corollary 4.20. *Let $A, B \in M_{m,n}(0,1)$ and $\pi = (S_1, S_2, \ldots, S_n) \in \mathcal{P}_n$. Suppose that each subset S_i corresponds to $|S_i|$ columns of A: $A^{(i,1)}, A^{(i,2)}, \ldots, A^{(i,|S_i|)}$ in the following way:* ***Either*** *$A^{(i,1)} = \sum_{j \in S_i} B^{(j)}$ and $A^{(i,x)} = 0$ for $x = 2, \ldots, |S_i|$* ***or*** *$A^{(i,1)} = \cdots = A^{(i,k)} = \frac{1}{k} \sum_{j \in S_i} B^{(j)}$ and $A^{(i,x)} = 0$ for $x = k+1, \ldots, |S_i|$. Here $(i,x) \in \{1, \ldots, n\}$ and $(i,x) = (j,y)$ implies $i = j$ and $x = y$.*

Then $A \preceq^m B$.

The corollary above sometimes allows to check easily that A is majorized by B as the following example shows.

Example 4.21. Let

$$A = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} \text{ and } B = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

Moreover, let

$$R = \begin{bmatrix} 0.5 & 0.5 & 0 & 0 & 0 \\ 0.5 & 0.5 & 0 & 0 & 0 \\ 0.5 & 0.5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}.$$

Then it is easy to verify that $A = BR$. In this case $\{B^{(1)}, B^{(2)}, B^{(3)}\}$ corresponds to $\{A^{(1)}, A^{(2)}, A^{(3)}\}$, since $A^{(1)} = A^{(2)} = \frac{1}{2}(B^{(1)} + B^{(2)} + B^{(3)})$ and $A^{(3)} = 0$. $\{B^{(4)}, B^{(5)}\}$ corresponds to $\{A^{(4)}, A^{(5)}\}$ since $B^{(4)} + B^{(5)} = A^{(4)}$ and $A^{(5)} = 0$. □

The following characterization of matrix majorization is from [2]. Originally it was shown using Farkas' lemma (see, for example, [12, Corollary 22.3.1]). This was done by considering a nonnegative solution R to the system of linear equations $A = BR$, $Re = e$, here the matrix equations is written column-wise. For $m \times n$ matrices $U = [u_{ij}]$ and $V = [v_{ij}]$, we use the usual inner product $\langle U, V \rangle = \sum_{i,j} u_{ij} v_{ij}$, and also define

$$\rho(U) = \sum_j \max_i u_{ij}.$$

Theorem 4.22. [2, Theorem 3.6] *Let A and B be $m \times n$ matrices. Then $A \preceq^m B$ if and only if*

$$\langle A, Y \rangle \le \rho(Y^T B) \tag{2}$$

for all $Y \in M_{m,n}$.

Specializing this result to $(0,1)$-matrices we obtain the following characterization. Let $\mathrm{supp}(B^{(j)})$ denote the support of the j'th column of a matrix $B \in M_{m,n}$, viewed as a subset of $\{1, 2, \dots, m\}$.

Corollary 4.23. *Let $A, B \in M_{m,n}(0,1)$. Then $A \preceq^m B$ if and only if*

$$\sum_j \sum_{i \in supp(A^{(j)})} y_{ij} \le \sum_j \max_k \sum_{i \in supp(B^{(j)})} y_{ik} \tag{3}$$

for all matrices $Y = [y_{ij}] \in M_{m,n}$.

Based on this corollary it is sometimes very easy to check that A is not majorized by B as the following example shows.

Example 4.24. Let

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}.$$

Then $B \preceq^m A$ as $B = A_\pi$ where $\pi = (\{1, 2\}, \emptyset)$. However, $A \not\preceq^m B$ by Corollary 4.23. Indeed, in this case (3) becomes

$$y_{11} + y_{22} \le \max\{y_{11} + y_{21}, y_{12} + y_{22}\}$$

which is violated for $y_{11} = y_{22} = 1$, $y_{12} = y_{21} = 0$. The same conclusion is obtained from Lemma 4.14 as the first row of $Z(A, B)$ is zero. □

Declaration of competing interest

There is no competing interest.

Acknowledgements

The authors wish to thank a referee for several useful comments and suggestions.

References

[1] R.A. Brualdi, Combinatorial Matrix Classes, Encyclopedia of Mathematics, Cambridge University Press, 2006.
[2] G. Dahl, Matrix majorization, Linear Algebra Appl. 288 (1999) 53–73.
[3] G. Dahl, Majorization polytopes, Linear Algebra Appl. 297 (1999) 157–175.
[4] G. Dahl, A. Guterman, P. Shteyner, Majorization for matrix classes, Linear Algebra Appl. 555 (2018) 201–221.
[5] G. Gour, D. Jennings, F. Buscemi, R. Duan, I. Marvian, Quantum majorization and a complete set of entropic conditions for quantum thermodynamics, Nat. Commun. 9 (2018) 5352.
[6] G.H. Hardy, J.E. Littlewood, G. Pólya, Inequalities, Cambridge University Press, 1934, 1952, 1978.
[7] G. Koshevoy, Multivariate Lorenz majorization, Soc. Choice Welf. 12 (1995) 93–102.
[8] G. Koshevoy, The Lorenz zonotope and multivariate majorizations, Soc. Choice Welf. 15 (1998) 1–14.
[9] S.-G. Hwang, S.-S. Pyo, Matrix majorization via vector majorization, Linear Algebra Appl. 332–334 (2001) 15–21.
[10] A.W. Marshall, I. Olkin, B.C. Arnold, Inequalities: Theory of Majorization and Its Applications, second ed., Springer, New York, 2011.
[11] F.D. Martínez Pería, Pedro G. Massey, Luis E. Silvestre, Weak matrix majorization, Linear Algebra Appl. 403 (2005) 343–368.
[12] R.T. Rockafellar, Convex Analysis, second ed., Princeton University Press, New Jersey, 1972.
[13] S. Sherman, On a conjecture concerning doubly stochastic matrices, Proc. Amer. Math. Soc. 3 (1952) 511–513.
[14] E. Torgersen, Comparison of Statistical Experiments, Encyclopedia of Mathematics and Its Applications, vol. 36, Cambridge University Press, Cambridge, 1991.