

Some Results on Majorization of Matrices

Divya K. Udayan ^{*,†} and Kanagasabapathi Somasundaram [†] 

Department of Mathematics, Amrita School of Engineering, Amrita Vishwavidyapeetham, Coimbatore 641112, India; s_sundaram@cb.amrita.edu

* Correspondence: ku_divya@cb.students.amrita.edu

† These authors contributed equally to this work.

Abstract: For two $n \times m$ real matrices X and Y , X is said to be majorized by Y , written as $X \prec Y$ if $X = SY$ for some doubly stochastic matrix of order n . Matrix majorization has several applications in statistics, wireless communications and other fields of science and engineering. Hwang and Park obtained the necessary and sufficient conditions for X, Y to satisfy $X \prec Y$ for the cases where the rank of $Y = n - 1$ and the rank of $Y = n$. In this paper, we obtain some necessary and sufficient conditions for X, Y to satisfy $X \prec Y$ for the cases where the rank of $Y = n - 2$ and in general for rank of $Y = n - k$, where $1 \leq k \leq n - 1$. We obtain some necessary and sufficient conditions for X to be majorized by Y with some conditions on X and Y . The matrix X is said to be doubly stochastic majorized by Y if there is $S \in \Omega_m$ such that $X = YS$. In this paper, we obtain some necessary and sufficient conditions for X to be doubly stochastic majorized by Y . We introduced a new concept of column stochastic majorization in this paper. A matrix X is said to be column stochastic majorized by Y , denoted as $X \preceq^c Y$, if there exists a column stochastic matrix S such that $X = SY$. We give characterizations of column stochastic majorization and doubly stochastic majorization for $(0, 1)$ matrices.

Keywords: matrix majorization; doubly stochastic majorization; multivariate majorization



Citation: Udayan, D.K.;

Kanagasabapathi, S. Some Results on Majorization of Matrices. *Axioms* **2022**, *11*, 146. <https://doi.org/10.3390/axioms11040146>

Academic Editors: Hans J. Haubold and Hari Mohan Srivastava

Received: 1 February 2022

Accepted: 20 March 2022

Published: 23 March 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Let \mathbb{R}^n denote the set of all real column vectors with n coordinates. A real matrix A is called non-negative, denoted by $A \geq 0$, if all its entries are non-negative. Let Ω_n denote the set of all $n \times n$ doubly stochastic matrices, i.e., real non-negative matrices, with each row sum and column sum equal to 1. For a vector $(a_1, a_2, \dots, a_n)^T \in \mathbb{R}^n$, let $(a_{[1]}, a_{[2]}, \dots, a_{[n]})^T$ denote the vector obtained from $(a_1, a_2, \dots, a_n)^T$ by rearranging the coordinates in non-increasing order. For vectors $x = (x_1, x_2, \dots, x_n)^T, y = (y_1, y_2, \dots, y_n)^T \in \mathbb{R}^n$, x is said to be majorized by y , denoted by $x \prec y$, if $\sum_{i=1}^k x_{[i]} \leq \sum_{i=1}^k y_{[i]}$ for all positive integers k such that $1 \leq k \leq n$.

It is well known that for two vectors $x, y \in \mathbb{R}^n$, $x \prec y$ iff $x = Sy$ for some $S \in \Omega_n$. In [1], the polytope of doubly stochastic matrices D for which $x = Dy$ was investigated. The notion of vector majorization is naturally extended to that of matrix majorization as follows. For two matrices X and Y in $\mathbb{R}^{n \times m}$, X is said to be majorized by Y , denoted by $X \prec Y$, if $X = SY$ for some $S \in \Omega_n$. In [2], a new notion of matrix majorization was introduced, called weak matrix majorization, where $X \prec_w Y$, if there exists a row stochastic matrix R such that $X = RY$. Additionally, relations between this concept and strong majorization (usual majorization) are considered. In [3], the polytope of row-stochastic matrices R for which $X = RY$ was investigated and generalizations of the results for vector majorization were obtained. The notion of matrix majorization was referred to as multivariate majorization in [4]. In [4], it was proved that, if $X \prec Y$ for $X, Y \in \mathbb{R}^{n \times m}$, then $XC \prec YC$ for any real matrix C with m rows.

For $X = [x_1, x_2, \dots, x_m], Y = [y_1, y_2, \dots, y_m] \in \mathbb{R}^{n \times m}$, if $X \prec Y$, then certainly $x_j \prec y_j$ for all $j = 1, 2, \dots, m$. However, the converse does not hold in general, as is easily seen with the matrices $X = \begin{bmatrix} \frac{1}{2} & 1 \\ \frac{1}{2} & 0 \end{bmatrix}$ and $Y = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$.

For $x, y \in \mathbb{R}^n$, let $\Omega(x \prec y)$ denote the set of all $S \in \Omega_n$ satisfying $x = Sy$. The set $\Omega(x \prec y)$ is known to contain various special types of doubly stochastic matrices [5,6]. While quite a lot of progress has been made in the theory of vector majorization, very little is known about multivariate majorization. In [7], it is proved that X is majorized by Y if Xv is majorized by Yv for every real n vector v , under the assumption that $[X, e][Y, e]^+$ is nonnegative, where e denotes the m -vector of ones and $[Y, e]^+$ denotes the Moore–Penrose generalized inverse of X .

In [8], a new matrix majorization order for classes (sets) of matrices was introduced, which generalizes several existing notions of matrix majorization. Matrix majorization order has several applications in mathematical statistics. Some applications of the matrix majorization were discussed by Marshall, Olkinbarry and Arnold [4] and Tong [9]. Majorization of $(0,1)$ matrices have important applications in classification theory and principal/dominant component analysis. Eduard Jorswieck and Holger Boche [10] reviewed the basic definitions of majorization theory and matrix monotone functions, describing their concepts clearly with many illustrative examples and then proceeded to show their applications in wireless communications. In [11], an algorithm was developed for the problem of finding a low-rank correlation matrix nearest to a given correlation matrix. The algorithm was based on majorization. The problem of rank reduction of correlation matrices occurs when pricing a derivative dependent on a large number of assets, where the asset prices are modeled as correlated log-normal processes. Mainly, such an application concerns interest rates. Matrix majorization also has applications in the comparison of eigenvalues [12].

Hwang and Park [13] obtained some necessary and sufficient conditions for $X, Y \in \mathbb{R}^{n \times m}$ to satisfy $X \prec Y$ for the case that the rank of Y is any one of $1, n - 1$ or n and for the case that $n \leq 3$. In Section 2, we obtain some necessary and sufficient conditions for $X, Y \in \mathbb{R}^{n \times m}$ to satisfy $X \prec Y$ for the case that the rank of Y is $n - 2$ and for the general case that the rank of Y is $n - k, 1 \leq k \leq n - 1$. We also obtain some necessary and sufficient conditions for X to be majorized by Y with some conditions on X and Y . Additionally, we obtain some necessary and sufficient conditions for X to be doubly stochastic majorized by Y with some conditions on X and Y .

Dahl, Guterman and Shteyner [14] obtained several results concerning matrix majorizations of $(0, 1)$ matrices and characterizations for certain matrix majorization orders. We extend these results for $(0, 1)$ matrices in Section 3. We introduce a new concept of column stochastic majorization. A matrix X is said to be column stochastic majorized by Y , denoted as $X \preceq^c Y$, if there exists a column stochastic matrix S such that $X = SY$. We obtain some characterizations for column stochastic majorization and doubly stochastic majorization of $(0, 1)$ matrices.

2. Matrix Majorization

Let I_n denote the identity matrix of order n . For two real matrices A, B of the same size, let $A \geq B$ (resp. $A \leq B$) denote that each entry of A is bigger (resp. less) than or equal to the corresponding entry of B . Let e_k denote all 1 vectors in \mathbb{R}^k . For matrix A , let $\sigma(A), r_A$ and C_A denote the sum of all of the entries of A , the row sum vector of A and the column sum vector of A , respectively. A vector $z \in \mathbb{R}^n$ is called a stochastic vector if $z \geq 0$ and $z^T e_n = 1$.

Theorem 1. Let $X = [x_{ij}], Y = [y_{ij}] = [I_{n-2}, y_1, y_2]^T \in \mathbb{R}^{n \times n-2}$. Then $X \prec Y$ if the following hold.

1. $c_X = c_Y$.
2. There exist stochastic vectors $z_1, z_2 \in \mathbb{R}^n$ such that $e_n - r_X = (1 - \sigma(y_1))z_1 + (1 - \sigma(y_2))z_2$ and $X \geq z_1y_1^T + z_2y_2^T$.

Proof. Let us assume $X \prec Y$. Then there exists $S \in \Omega_n$ satisfying $X = SY$. It is easy to see that $c_X = c_Y$. Let $S = [A, z_1, z_2]$ with $z_1, z_2 \in \mathbb{R}^n$. Then $X = SY = A + z_1y_1^T + z_2y_2^T$. Since $A \geq 0$, we obtain that there exist stochastic vectors z_1 and z_2 such that $X \geq z_1y_1^T + z_2y_2^T$. $Se_n = Ae_{n-2} + z_1 + z_2 = Xe_{n-2} - z_1y_1^Te_{n-2} - z_2y_2^Te_{n-2} + z_1 + z_2 = r_X + (1 - \sigma(y_1))z_1 + (1 - \sigma(y_2))z_2$. Since $Se_n = e_n$, we obtain that there exist stochastic vectors z_1 and z_2 such that $e_n = r_X + (1 - \sigma(y_1))z_1 + (1 - \sigma(y_2))z_2$. Conversely, suppose (1) and (2) hold. For the vectors z_1, z_2 in (2), let $S = [A, z_1, z_2]$, where $A = X - z_1y_1^T - z_2y_2^T$. Then $SY = [A, z_1, z_2][I_{n-2}, y_1, y_2]^T = A + z_1y_1^T + z_2y_2^T = X$. It remains to show that $S \in \Omega_n$. We see from (2) that $S \geq 0$. From (1) and (2) we have $Se_n = e_n$. Additionally, $e_n^T S = [e_n^T X - e_n^T z_1y_1^T - e_n^T z_2y_2^T, e_n^T z_1, e_n^T z_2] = [c_X - y_1^T - y_2^T, 1, 1] = [e_{n-2}^T, 1, 1] = e_n^T$ since $c_X = c_Y$ and $c_Y = e_{n-2}^T + y_1^T + y_2^T$. This implies that $S \in \Omega_n$. \square

The above theorem can be extended to any $Y \in \mathbb{R}^{n \times n-k}$ of rank $n - k$.

Theorem 2. Let $X = [x_{ij}], Y = [I_{n-k}, y_1, y_2, \dots, y_k]^T \in \mathbb{R}^{n \times n-k}, 1 \leq k \leq n - 1$. Then $X \prec Y$ if the following hold.

1. $c_X = c_Y$.
2. There exist stochastic vectors z_1, z_2, \dots, z_k such that $e_n - r_X = \sum_{i=1}^k (1 - \sigma(y_i))z_i$ and $X \geq z_1y_1^T + z_2y_2^T + \dots + z_ky_k^T$.

Proof. Proof is similar to the proof of Theorem 1 extending to k vectors. \square

In the next theorem, we give a necessary condition for $X \preceq Y$, for any two matrices $X, Y \in \mathbb{R}^{n \times m}$.

Theorem 3. Let $X, Y \in \mathbb{R}^{n \times m}$. If $X \preceq Y$ then there exists $\delta_1, \delta_2, \dots, \delta_{n-1} \in \mathbb{R}$ such that $|\delta_i| \leq 1$ for $i = 1, 2, 3, \dots, n - 1, \delta_1 + \delta_2 + \dots + \delta_{n-1} \geq n - 3$ and $x_{1j} + x_{2j} + \dots + x_{n-1,j} - x_{n,j} = \delta_1y_{1j} + \delta_2y_{2j} + \dots + \delta_{n-1}y_{n-1,j} + y_{nj}(n - 2 - \delta_1 - \delta_2 - \dots - \delta_{n-1})$.

Proof. Suppose $X \preceq Y$ so that $X = SY$ for some $S = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \in \Omega_n$. Let

$[1, 1, \dots, -1]$ be a $1 \times n$ matrix. Then the j -th element in the row vector $[1, 1, \dots, -1]X = [1, 1, \dots, -1]SY$ is $x_{1j} + x_{2j} + \dots + x_{n-1,j} - x_{n,j} = (a_{11} + a_{21} + \dots - a_{n1})y_{1j} + (a_{12} + a_{22} + \dots - a_{n2})y_{2j} + \dots + (a_{1n} + a_{2n} + \dots - a_{nn})y_{nj} = (a_{11} + a_{21} + \dots - (1 - (a_{11} + a_{21} + \dots + a_{n1})))y_{1j} + (a_{12} + a_{22} + \dots - (1 - (a_{12} + a_{22} + \dots - a_{n2})))y_{2j} + \dots + (a_{1n-1} + a_{2n-1} + \dots - (1 - (a_{1n-1} + a_{2n-1} + \dots - a_{nn-1})))y_{nj} = (2(a_{11} + a_{21} + \dots + a_{n-11}) - 1)y_{1j} + (2(a_{1,2} + a_{2,2} + \dots + a_{n-1,2}) - 1)y_{2,j} + \dots + (2(a_{1,n} + a_{2,n-1} + \dots + a_{n-1,n-1}) - 1)y_{n,j} = (2(a_{1,1} + a_{2,1} + \dots + a_{n-1,1}) - 1)y_{1,j} + (2(a_{1,2} + a_{2,2} + \dots + a_{n-1,2}) - 1)y_{2,j} + \dots + (2(a_{1,n} + a_{2,n} + \dots + a_{n-1,n-1}) - 1)y_{n,j}$. $2(a_{1n} + a_{2n} + \dots + a_{n-1n}) - 1 = 2n - 3 - 2(a_{11} + a_{12} + \dots + a_{1n-1} + \dots + a_{n-11} + a_{n-12} + \dots + a_{n-1n-1})$. Let $2(a_{11} + a_{21} + \dots + a_{n-11}) - 1 = \delta_1, 2(a_{12} + a_{22} + \dots + a_{n-12}) - 1 = \delta_2, \dots, 2(a_{1n-1} + a_{2n-1} + \dots + a_{n-1n-1}) - 1 = \delta_{n-1}$.

This implies $x_{1,j} + x_{2,j} + \dots + x_{n-1,j} - x_{n,j} = \delta_1y_{1,j} + \delta_2y_{2,j} + \dots + \delta_{n-1}y_{n-1,j} + y_{nj}(2n - 3 - 2(\frac{1+\delta_1}{2} + \frac{1+\delta_2}{2} + \dots + \frac{1+\delta_{n-1}}{2})) = \delta_1y_{1j} + \delta_2y_{2j} + \dots + \delta_{n-1}y_{n-1,j} + (n - 2 - (\delta_1 + \delta_2 + \dots + \delta_n))y_{nj}$.

$2(a_{1,j} + a_{2,j} + \dots + a_{n-1,j}) - 1 = \delta_j$.
 Since $0 \leq a_{1,j} + a_{2,j} + \dots + a_{n-1,j} \leq 1, -1 \leq 2(a_{1,j} + a_{2,j} + \dots + a_{n-1,j}) - 1 \leq 1$.
 $\Rightarrow |\delta_j| \leq 1. \Rightarrow a_{11} + a_{21} + \dots + a_{n-11} + \dots + a_{1n-1} + a_{2n-1} + \dots + a_{n-1n-1} + 2 - n \geq$
 $0. \frac{1+\delta_1}{2} + \frac{1+\delta_2}{2} + \dots + \frac{1+\delta_{n-1}}{2} + 2 - n \geq 0. \delta_1 + \delta_2 + \dots + \delta_{n-1} \geq n - 3. \quad \square$

Example 1. Let $Y = \begin{bmatrix} 1 & 3 & 1 \\ 1 & 1 & 1 \\ 2 & 2 & 3 \end{bmatrix}, S = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0.5 & 0.5 \\ 0 & 0.5 & 0.5 \end{bmatrix}$ and $X = \begin{bmatrix} 1 & 3 & 1 \\ 1.5 & 1.5 & 2 \\ 1.5 & 1.5 & 2 \end{bmatrix}$. Then $\delta_1 = 1$ and $\delta_2 = 0$ are such that they satisfy the required conditions of the theorem and so this example validates Theorem 3.

The converse of Theorem 3 is not true.

Example 2. Let $X = \begin{bmatrix} 1.5 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 2 & 3 \end{bmatrix}$ and $Y = \begin{bmatrix} 1 & 3 & 1 \\ 1 & 1 & 1 \\ 2 & 2 & 3 \end{bmatrix}$.

Take $\delta_1 = 0.5$ and $\delta_2 = 0$. Then X, Y and δ_1, δ_2 satisfy the conditions of the above theorem but there exists no matrix S such that $X = SY$. Assume that there exists a matrix S such that $S = \begin{bmatrix} s_{11} & s_{12} & 1 - s_{11} - s_{12} \\ s_{21} & s_{22} & 1 - s_{21} - s_{22} \\ 1 - s_{11} - s_{21} & 1 - s_{12} - s_{22} & (s_{11} + s_{12} + s_{21} + s_{22} - 1) \end{bmatrix}$ Since $X = SY$ when we multiply the first row of S with the first column of Y , we obtain $x_{11} = 1.5 = s_{11} + s_{12} + (1 - s_{11} - s_{12})2$. This implies that $s_{11} + s_{12} = 0.5$. Similarly for x_{21} , we have $x_{21} = 0 = s_{21} + s_{22} + (1 - s_{21} - s_{22})2$. This implies that $s_{21} + s_{22} = 1$. Similarly for x_{31} , we have $(1 - s_{11} - s_{21}) + (1 - s_{12} - s_{22}) + (s_{11} + s_{21} + s_{12} + s_{22} - 1)2 = 0$. This implies that $s_{11} = 0, s_{12} = 0, s_{21} = 0$ and $s_{22} = 0$. However, this is a contradiction to $s_{11} + s_{12} = 0.5$. Therefore, there exists no such matrix S such that $X = SY$. In the next theorem, we show sufficient conditions for $X \preceq Y$, for any two matrices $X, Y \in \mathbb{R}^{n \times m}$.

Theorem 4. Let $X, Y \in \mathbb{R}^{n \times m}$. A sufficient condition for X to be majorized by Y is that $c_X = c_Y$ and $x_{1j} = x_{2j} = \dots = x_{n-1j}$ for $j = 1, 2, \dots, m$ and there exists $\delta_1, \delta_2, \dots, \delta_{n-1} \in \mathbb{R}$ such that $|\delta_i| \leq 1$ for $i = 1, 2, \dots, n - 1, \delta_1 + \delta_2 + \dots + \delta_{n-1} \geq n - 3$ and for $j = 1, 2, \dots, m$ $x_{1j} + x_{2j} + \dots + x_{n-1j} - x_{nj} = \delta_1 y_{1j} + \delta_2 y_{2j} + \dots + \delta_{n-1} y_{n-1j} + y_{nj}(n - 2 - \delta_1 - \delta_2 - \dots - \delta_{n-1})$.

Proof. Assume that $c_X = c_Y$ and $\delta_1, \delta_2, \dots, \delta_{n-1} \in \mathbb{R}$ such that the conditions given in the theorem hold. Let $S = \begin{bmatrix} \frac{1+\delta_1}{2(n-1)} & \frac{1+\delta_2}{2(n-1)} & \dots & \frac{1+\delta_{n-1}}{2(n-1)} & \frac{1}{2} - \frac{1}{2(n-1)}(\delta_1 + \delta_2 + \dots + \delta_{n-1}) \\ \frac{1+\delta_1}{2(n-1)} & \frac{1+\delta_2}{2(n-1)} & \dots & \frac{1+\delta_{n-1}}{2(n-1)} & \frac{1}{2} - \frac{1}{2(n-1)}(\delta_1 + \delta_2 + \dots + \delta_{n-1}) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{1+\delta_1}{2(n-1)} & \frac{1+\delta_2}{2(n-1)} & \dots & \frac{1+\delta_{n-1}}{2(n-1)} & \frac{1}{2} - \frac{1}{2(n-1)}(\delta_1 + \delta_2 + \dots + \delta_{n-1}) \\ \frac{1-\delta_1}{2} & \frac{1-\delta_2}{2} & \dots & \frac{1+\delta_{n-1}}{2} & 1 - \frac{(n-1)}{2} + \frac{\delta_1}{2} + \frac{\delta_2}{2} + \dots + \frac{\delta_{n-1}}{2} \end{bmatrix}$.

Then S is a doubly stochastic matrix. When we multiply the first row of S and the first column of Y , we will obtain $\frac{(1+\delta_1)}{2(n-1)}y_{11} + \frac{(1+\delta_2)}{2(n-1)}y_{21} + \dots + \frac{1+\delta_{n-1}}{2(n-1)}y_{n-11} + (\frac{1}{2} - \frac{1}{2(n-1)}(\delta_1 + \delta_2 + \dots + \delta_{n-1}))y_{n1} = \frac{y_{11}}{2(n-1)} + \frac{y_{21}}{2(n-1)} + \dots + \frac{y_{n-11}}{2(n-1)} + \frac{y_{n1}}{2(n-1)} + \frac{1}{2(n-1)}(\delta_1 y_{1,1} + \delta_2 y_{2,1} + \dots + \delta_{n-1} y_{n-1,1}) + y_{n,1}(\frac{1}{2} - \frac{1}{2(n-1)} - \frac{1}{2(n-1)}(\delta_1 + \delta_2 + \dots + \delta_{n-1})) = \frac{1}{2(n-1)}(y_{11} + y_{21} + \dots + y_{n1}) + \frac{1}{2(n-1)}(\delta_1 y_{11} + \delta_2 y_{21} + \dots + \delta_{n-1} y_{n-1,1}) + y_{n1}(\frac{1}{2} - \frac{1}{2(n-1)} - \frac{1}{2(n-1)}(\delta_1 + \delta_2 + \dots + \delta_{n-1})) = \frac{1}{2(n-1)}(x_{11} + x_{21} + \dots + x_{n1}) + \frac{1}{2(n-1)}(\delta_1 y_{11} + \delta_2 y_{21} + \dots + \delta_{n-1} y_{n-1,1}) + y_{n1}(n - 2 - (\delta_1 + \delta_2 + \dots + \delta_{n-1})) = \frac{1}{2(n-1)}(x_{11} + x_{21} + \dots + x_{n1}) + (x_{11} + x_{21} + \dots + x_{n-11} - x_{n1}) = x_{11}$. Hence, $X = SY$. This implies that X is majorized by Y . \square

Example 3. Let $X = \begin{bmatrix} 1 & 0.5 & 1 \\ 1 & 0.5 & 1 \\ 2 & 3 & 2 \end{bmatrix}$ and $Y = \begin{bmatrix} -4 & -2 & -1 \\ 4 & 8 & 4 \\ 4 & -2 & 1 \end{bmatrix} \in \mathbb{R}^{3 \times 3}$.

Then X and Y satisfy the conditions of Theorem 4 and so X is majorized by Y . Take $\delta_1 = 0.5$ and $\delta_2 = 0$. Then $S = \begin{bmatrix} 0.375 & 0.25 & 0.375 \\ 0.375 & 0.25 & 0.375 \\ 0.25 & 0.5 & 0.25 \end{bmatrix} \in \Omega_3$ and $X = SY$. So X is majorized by Y .

A matrix X is said to be doubly stochastic majorized by Y , denoted by $X \prec^{ds} Y$, when there is $S \in \Omega_m$ such that $X = YS$. In the next theorem, we prove a necessary condition for X to be doubly stochastic majorized by Y .

Theorem 5. Let $X, Y \in \mathbb{R}^{n \times m}$. If $X \prec^{ds} Y$ then there exists $\delta_1, \delta_2, \dots, \delta_{m-1} \in \mathbb{R}$ such that $|\delta_i| \leq 1$ for $i = 1, 2, \dots, m - 1, \delta_1 + \delta_2 + \dots + \delta_{m-1} \geq m - 3$ and $x_{i1} + x_{i2} + \dots + x_{im-1} - x_{im} = \delta_1 y_{i1} + \delta_2 y_{i2} + \dots + y_{im}(m - 2 - \delta_1 - \delta_2 - \dots - \delta_{m-1})$.

Proof. Suppose $X \prec^{ds} Y$ so that $X = YS$ for some $S = (a_{ij})_{m \times m} \in \Omega_m$. The j -th element of the vector $X[1, 1, \dots, -1] = YS[1, 1, \dots, -1]$ is $x_{i1} + x_{i2} + \dots - x_{im} = y_{i1}(2(a_{11} + a_{12} + \dots + a_{1m-1}) - 1) + y_{i2}(2(a_{21} + a_{22} + \dots + a_{2m-1}) - 1) + \dots + y_{im}(2(a_{m1} + a_{m2} + \dots + a_{mm-1}) - 1)$. $2(a_{m1} + a_{m2} + \dots + a_{mm-1}) - 1 = 2m - 3 - 2(a_{11} + \dots + a_{m-11} + a_{12} + a_{22} + \dots + a_{m-12} + \dots + a_{1m-1} + a_{2m-1} + \dots + a_{m-1m-1})$. Let $2(a_{11} + a_{21} + \dots + a_{m-11}) - 1 = \delta_1, 2(a_{12} + a_{22} + \dots + a_{m-12}) - 1 = \delta_2, \dots, 2(a_{1m-1} + a_{2m-1} + \dots + a_{m-1m-1}) - 1 = \delta_{m-1}$.

This implies, $x_{i1} + x_{i2} + \dots + x_{im-1} - x_{im} = \delta_1 y_{i1} + \delta_2 y_{i2} + \dots + y_{im}(2m - 3 - 2(\frac{1+\delta_1}{2} + \frac{1+\delta_2}{2} + \dots + \frac{1+\delta_{m-1}}{2})) = \delta_1 y_{i1} + \delta_2 y_{i2} + \dots + y_{im}(m - 2 - (\delta_1 + \delta_2 + \dots + \delta_{m-1}))$. Since $0 \leq a_{i1} + a_{i2} + \dots + a_{im-1} \leq 1, -1 \leq 2(a_{i1} + a_{i2} + \dots + a_{im-1}) - 1 \leq 1. \Rightarrow |\delta_i| \leq 1. a_{mm} \geq 0. \Rightarrow a_{11} + a_{12} + \dots + a_{1m-1} + \dots + a_{m-1m-1} + 2 - m \geq 0. \Rightarrow \frac{1+\delta_1}{2} + \frac{1+\delta_2}{2} + \dots + \frac{1+\delta_{m-1}}{2} + 2 - m \geq 0. \Rightarrow \delta_1 + \delta_2 + \dots + \delta_{m-1} \geq m - 3. \square$

Example 4. Let $X = \begin{bmatrix} 1 & 1.5 & 1.5 \\ 3 & 1.5 & 1.5 \\ 1 & 2 & 2 \end{bmatrix}, Y = \begin{bmatrix} 1 & 1 & 2 \\ 3 & 1 & 2 \\ 1 & 1 & 3 \end{bmatrix}$ and $S = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0.5 & 0.5 \\ 0 & 0.5 & 0.5 \end{bmatrix}$. Then $\delta_1 = 1$

and $\delta_2 = 0$ are such that they satisfy the required conditions of the theorem and so this example validates Theorem 5.

The converse of Theorem 5 is not true.

Example 5. $X = \begin{bmatrix} 1.5 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 2 & 3 \end{bmatrix}$ and $Y = \begin{bmatrix} 1 & 3 & 1 \\ 1 & 1 & 1 \\ 2 & 2 & 3 \end{bmatrix}$. Take $\delta_1 = 0.5$ and $\delta_2 = 0$.

Then X, Y and δ_1, δ_2 satisfy the conditions of the above theorem, but there exists no matrix such that $X = YS$. Assume that there exists a matrix S such that $S = \begin{bmatrix} s_{11} & s_{12} & 1 - s_{11} - s_{12} \\ s_{21} & s_{22} & 1 - s_{21} - s_{22} \\ 1 - s_{11} - s_{21} & 1 - s_{12} - s_{22} & (s_{11} + s_{12} + s_{21} + s_{22} - 1) \end{bmatrix}$. Since $X = YS$ when we multiply the first row of Y with the first column of S , we obtain $x_{11} = 1.5 = 2s_{21} + 1$. This implies that $s_{21} = 0.25$. Similarly for x_{21} we have $x_{21} = 1 = 2s_{22} + 1$. This implies that $s_{22} = 0$. Similarly, for x_{13} we have $x_{13} = 1 = 3 - 2s_{21} - 2s_{22}$. This implies that $s_{21} + s_{22} = 1$ which is a contradiction since $s_{21} = 0.25$ and $s_{22} = 0$. Therefore, there exists no such matrix S such that $X = YS$.

Theorem 6. Let $X, Y \in \mathbb{R}^{n \times m}$. A sufficient condition for X to be doubly stochastic majorized by Y is that $r_X = r_Y$ and $x_{i1} = x_{i2} = \dots = x_{im-1}$ for $i = 1, 2, \dots, m$ and there exist $\delta_1, \delta_2, \dots, \delta_{m-1} \in \mathbb{R}$ such that $|\delta_i| \leq 1$ for $i = 1, 2, \dots, m - 1, \delta_1 + \delta_2 + \dots + \delta_{m-1} \geq m - 3$ and $x_{i1} + x_{i2} + \dots + x_{im-1} - x_{im} = \delta_1 y_{i1} + \delta_2 y_{i2} + \dots + y_{im}(m - 2 - \delta_1 - \delta_2 - \dots - \delta_{m-1})$.

Proof. The proof is similar to the proof of Theorem 4. \square

Example 6. Let $X = \begin{bmatrix} 1 & 1 & 2 \\ 0.5 & 0.5 & 3 \\ 1 & 1 & 2 \end{bmatrix}$ and $Y = \begin{bmatrix} -4 & 4 & 4 \\ -2 & 8 & -2 \\ -1 & 4 & 1 \end{bmatrix} \in \mathbb{R}^{3 \times 3}$. Then X and Y satisfy the conditions of Theorem 6 and so X is doubly stochastic majorized by Y .

Take $\delta_1 = 0.5$ and $\delta_2 = 0$. Then $S = \begin{bmatrix} 0.375 & 0.375 & 0.25 \\ 0.25 & 0.25 & 0.5 \\ 0.375 & 0.375 & 0.25 \end{bmatrix} \in \Omega_3$ and $YS = X$. So X is doubly stochastic majorized by Y .

3. Majorization for (0, 1) Matrices

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. Secondly, (0, 1) matrices are essential to represent combinatorial objects, and therefore one may look at the meaning of such a matrix majorization order (each associated with a (0, 1) matrix). In [14] weak, directional and strong majorizations of (0, 1) matrices are characterized. Matrix majorization on (0, 1) matrices is investigated.

Definition 1. A matrix X is said to be column stochastic majorized by Y , denoted as $X \preceq^c Y$, if there exists a column stochastic matrix S such that $X = SY$.

Theorems 7 and 8 give characterizations of \preceq^{ds} and \preceq^c for (0, 1) matrices.

Theorem 7. Let $X, Y \in M_{m,n}(0, 1)$ and $X \preceq^{ds} Y$. Then for every $i = 1, 2, \dots, m$ the number of 1's in the i -th column of X is equal to the number of 1's in the i -th column of Y .

Proof. By assumption, $X = YS$ for some $S \in \Omega_n$. Then $eX = eYS = eY$ where $e = [1, 1, \dots, 1]^t$. Since both matrices are (0, 1), the i -th entry of eX is the number of 1's in the i -th column of A . The same holds is for Y . Hence the result follows. \square

Theorem 8. Let $X, Y \in M_{m,n}(0, 1)$. Assume $X \preceq^c Y$, but $X \not\preceq^{ds} Y$. Then

- (i) There exists $S \in \Omega_n^{column}$ satisfying $X = SY$ such that S contains a zero row, and for each row sum r_i of S it holds that either $r_i = 0$ or $r_i \geq 1$.
- (ii) If Y does not contain a zero row, then for any $S \in \Omega_n^{column}$ satisfying $X = SY$ it holds that S contains a zero row, and for each row sum r_i of R it holds that $r_i = 0$ or $r_i \geq 1$.
- (iii) X contains a zero row.

Proof. (i) Suppose $X = SY$ where S is column stochastic. Since $S \in \Omega_n^{column}$, the sum of all elements in S is n . Since by assumption $X \not\preceq^{ds} Y$, there is i such that the i th row sum of S , $r_i < 1$ and there is k such that the k th row sum of S , $r_k > 1$. If $r_i < 1$, then $X_{(i)}$ is a zero row. Indeed, each element of $X_{(i)}$ is less than or equal to $r_i < 1$. Hence, it is zero. We are going to modify S in order to construct $S' \in \Omega_n^{column}$ such that $X = S'Y$ and $S'_{(i)}$ is zero. Fix some k such that $r_k > 1$. Suppose $s_{ip} \neq 0$. Since $0 = X_{(i)} = \sum_{j=1}^n s_{ij}Y^{(j)} \geq s_{ip}Y^{(p)}$, we obtain $Y^{(p)} = 0$. Consider arbitrary $x_{fg} = \sum_{h=1}^n s_{fh}y_{hg} = \sum_{h \neq p} s_{fp}y_{pg} + s_{fp}y_{pg} = \sum_{h \neq p} s_{fp}y_{pg}$. We consider the matrix S' , which is obtained from S by changing s_{kp} to $s_{kp} + s_{ip}$ and s_{ip} to 0. We do the same for the rest of the nonzero elements in $S_{(i)}$. Finally, we obtain S' such that $X = S'Y$, S'_i is a zero row, $r'_k > 1$ and $S'_{(l)} = S'_{(l)}$ for $l \neq i, k$. We repeat this procedure for every q with $0 < r_q < 1$. After several such substitutions, we obtain S' such that for every q either $r'_q = 0$ or $r'_q \geq 1$. S' contains a zero row, as required. (ii) Suppose Y does not contain a zero row. Let i be such that $r_i < 1$. From (i), $X_{(i)}$ is zero. $X_{(i)} = \sum_{j=1}^n s_{ij}Y_{(j)}$ and

all summands are non-negative, and it follows that $s_{ij} = 0$ for all j . Finally if $r_i < 1$ then $r_i = 0$ and S contains a zero row. (iii) By (i) we obtain that X' contains a zero row, and, as a consequence, $A = X'B$ contains a zero row. \square

In [14], it was proved that if $A \preceq^{wm} B$ then $R(A) \subseteq R(B)$, but matrix majorization cannot be described in terms of row/column inclusion. The following examples show that strong majorization also cannot be described in terms of row/column inclusion. The first example shows that column inclusion does not follow from $X \preceq^m Y$, and the second one shows that the converse implication does not hold as well.

Example 7. Let $X = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 1 \end{bmatrix}, Y = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}, S = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}.$

Then $X = SY$, so $X \preceq^s Y$. $X^{(1)} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \notin C(Y)$. Also, $Y^{(1)} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \notin C(X)$.

Example 8. Let $X = \begin{bmatrix} 1 & 1 & 1 \\ +1 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}$ and $Y = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}$. Then $eX = eY$ and $R(X) \subseteq R(Y)$.

But it is easy to verify that $X \not\preceq^m Y$. Suppose that there exists $S = \begin{bmatrix} a & d & g & j \\ b & e & h & k \\ \dots & & & \\ \dots & & & \end{bmatrix}$ such

that $X = SY$. Since $X = SY$. Since $SY = \begin{bmatrix} a+g & a+d & a+d+g+j \\ b+h & b+e & b+h+e+k \\ \dots & & \\ \dots & & \end{bmatrix}$ 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.

4. Conclusions

In this paper, we proved some necessary and sufficient conditions for $X, Y \in \mathbb{R}^{n \times m}$, $1 \leq m \leq n - 1$ to satisfy $X \prec Y$. We obtain some necessary and sufficient conditions for X to be majorized by Y with some conditions on X and Y . We also obtained some necessary and sufficient conditions for X to be doubly stochastic majorized by Y with some conditions on X and Y . We also obtained some characterizations for majorization of $(0, 1)$ matrices. Finding necessary and sufficient conditions for general matrix majorization is difficult and still it is open. Finding necessary and sufficient conditions for doubly stochastic majorization and strong majorization are the scope of future studies.

Author Contributions: Both authors have contributed equally to this paper. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Brualdi, R.A. The doubly stochastic matrices of a vector majorization. *Linear Algebra Appl.* **1984**, *61*, 141–154. [[CrossRef](#)]
2. Pería, F.D.; Massey, P.G.; Silvestre, L.E. Weak matrix majorization. *Linear Algebra Appl.* **2005**, *403*, 343–368. [[CrossRef](#)]
3. Dahl, G. Majorization Polytopes. *Linear Algebra Appl.* **1999**, *297*, 157–175. [[CrossRef](#)]
4. Marshall, A.W.; Olkin, I. *Inequalities: Theory of Majorization and Its Application*; Academic Press: New York, NY, USA, 1979.
5. Brualdi, R.A.; Hwang, S.G. Vector majorization via Hessenberg matrices. *J. Lond. Math. Soc.* **1996**, *53*, 28–38. [[CrossRef](#)]
6. Chao, K.M.; Wong, C.S. Application of M-matrices to majorization. *Linear Algebra Appl.* **1992**, *169*, 31–40. [[CrossRef](#)]
7. Hwang, S.G.; Pyo, S.S. Matrix majorization via vector majorization. *Linear Algebra Appl.* **2001**, *332–334*, 15–21. [[CrossRef](#)]
8. Dahl, G. Alexander Guterman, and Pavel Shteyner, Majorization for matrix classes. *Linear Algebra Appl.* **2018**, *555*, 201–221. [[CrossRef](#)]
9. Tong, Y.L. Some recent developments on majorization inequalities in probability and statistics. *Linear Algebra Appl.* **1994**, *199*, 69–90. [[CrossRef](#)]
10. Jorswieck, E.; Boche, H. *Majorization and Matrix-Monotone Functions in Wireless Communications*; Now Publishers Inc.: Delft, The Netherlands, 2007.
11. Pietersz, R.; Groenen, P.J. Rank reduction of correlation matrices by Majorization. *Quant. Financ.* **2004**, *4*, 649–662. [[CrossRef](#)]
12. Ando, T. Majorization, Doubly stochastic matrices, and comparison of eigenvalues. *Linear Algebra Appl.* **1989**, *118*, 163–248. [[CrossRef](#)]
13. Hwang, S.G.; Park, J.-Y. A note on multivariate majorization. *Commun. Korean Math. Soc.* **1999**, *14*, 479–485.
14. Dahl, G. Alexander Guterman, and Pavel Shteyner, Majorization for $(0, 1)$ -matrices. *Linear Algebra Appl.* **2020**, *585*, 147–163. [[CrossRef](#)]