Sankhyā : The Indian Journal of Statistics
1985, Volume 47, Series A, Pt. 3, pp. 357–365.

TWO CHARACTERIZATIONS OF DOUBLY SUPER-STOCHASTIC MATRICES

By SUBIR KUMAR BHANDARI
and
SOMESH DAS GUPTA
Indian Statistical Institute

SUMMARY. Two characterizations of doubly superstochastic matrices are given in this paper. These have originated from some open problems cited in Marshall and Olkin (1979).

1. Introduction

Marshall and Olkin (1979) called a $t \times t$ matrix $\boldsymbol{P}$ doubly superstochastic (d.s.s.) if there exists a doubly stochastic matrix (d.s.) $\boldsymbol{D}$ such that $\boldsymbol{P} \geqslant \boldsymbol{D}$, where $\geqslant$ signifies elementwise inequality. They stated two necessary conditions (to be mentioned as $C2$ and $C3$) in Proposition 2.D.3 (page 31) for a matrix $\boldsymbol{P}$ to be d.s.s., and posed the open question whether any of these conditions is sufficient for a matrix $\boldsymbol{P}$ to be d.s.s. A more interesting unsettled question stated in Marshall and Olkin (page 31) is whether $y\boldsymbol{P} \prec^{w} y$ for all $y \in R^t_+$ implies that $\boldsymbol{P}$ is d.s.s. Let us recall that $x \in R^t$ is said to be weakly supermajorized by $y \in R^t$ (written as $x \prec^{w} y$) if

$$\sum_{i=1}^{k} x_{(i)} \geqslant \sum_{i=1}^{k} y_{(i)}, \quad k = 1, \ldots, t \qquad \ldots \quad (1.1)$$

where $x_{(1)} \leqslant \ldots \leqslant x_{(t)}$ and $y_{(1)} \leqslant \ldots \leqslant y_{(t)}$ are ordered sets of components of x and y, respectively. Here R^t_+ denotes the non-negative orthant of the t-dimensional real space R^t.

In this paper we have resolved both the problems stated above. We have introduced a condition ($C1$) which is shown to be equivalent to either $C2$ or $C3$, and proved that a non-negative matrix $\boldsymbol{P}$ ($\geqslant 0$) is d.s.s. iff it satisfies $C1$. Moreover, we have shown that $y\boldsymbol{P} \prec^{w} y$ for all $y \in R^t_+$ iff $\boldsymbol{P}$ is d.s.s. Thus these results give two characterizations of d.s.s. matrices.

AMS (1980) *subject classification :* 15A99.

Key words and phrases : Doubly superstochastic matrix, characterization, weakly supermajorized, δ-transform.

2. Preliminaries

(i) For a matrix $A = [a_{ij}]$ we define

$$(A) = \sum_i \sum_j a_{ij} \qquad \dots \quad (2.1)$$

and

$$(A)_+ = \sum_i \sum_j \max(a_{ij}, 0) \qquad \dots \quad (2.2)$$

(ii) Consider a 2×2 submatrix

$$\begin{bmatrix} d_{ij} & d_{ik} \\ d_{sj} & d_{sk} \end{bmatrix} \qquad \dots \quad (2.3)$$

of a $t\times t$ d.s. $D = [d_{ij}]$ such that $d_{ij} > 0$, and $d_{sk} > 0$. By δ-transform of D with respect to this submatrix we mean another doubly stochastic matrix which has all the elements the same as those in D except for the elements in the above submatrix which are transformed to

$$\begin{bmatrix} d_{ij}-\delta & d_{ik}+\delta \\ d_{sj}+\delta & d_{sk}-\delta \end{bmatrix} \qquad \dots \quad (2.4)$$

where $\delta > 0$.

Given a $t\times t$ matrix P we say that a δ-transformation of a d.s. matrix D is invariant with respect to P if $(D-P)_+$ remains unchanged when D is replaced by its δ-transform. For simplicity, we shall use the same notation for a d.s. matrix or any of its invariant δ-transform.

(iii) Consider the class $\mathcal{D}$ of all $t\times t$ d.s. matrices. It is easy to see that $\mathcal{D}$ is compact in R^{t^2}. Note that for any $t\times t$ matrix P, the function $(D-P)_+$ is continuous in the elements of D. Hence there exists a d.s. matrix D such that

$$(D-P)_+ = \inf_{S\in\mathcal{D}} (S-P)_+. \qquad \dots \quad (2.5)$$

Such a matrix D will be called a minimizer with respect to P.

(iv) By a permutational transform of a $t\times t$ matrix P we mean the matrix P with some of its rows interchanged and/or some of its columns interchanged; i.e., P is transformed to $\pi_1 P \pi_2$, where π_1 and π_2 are permutation matrices. Note that a permutational transformation of a matrix P keeps (P) or $(P)_+$ unchanged.

Suppose D is a minimizer with respect to P. Then $\pi_1 D \pi_2$ is a minimizer with respect to $\pi_1 P \pi_2$. Note that the double super-stochastic property of a matrix is invariant under permutational transformation. For simplicity, we shall use the same notation for a matrix and any of its permutational transform whenever any condition imposed on that matrix is also satisfied by any of its permutational transform.

3. The basic lemma and the main results

The following lemma is the key to all the results in this paper. Its proof will be given later.

Basic Lemma : *Let* $\boldsymbol{P}$ *be a* $t\times t$ *matrix with all non-negative elements such that* $\inf_{S\in\mathcal{D}} (\boldsymbol{S}-\boldsymbol{P})_+ > 0$. *There exists a minimizer* $\boldsymbol{D} \in \mathcal{D}$ *with respect to* $\boldsymbol{P}$ *such that* $\boldsymbol{D}$ *and* $\boldsymbol{A} = \boldsymbol{D}-\boldsymbol{P}$, *after being subjected to a suitable permutational transformation, can be partitioned as*

$$\boldsymbol{D} = \left[\begin{array}{c|c|c} \boldsymbol{D}_{11} & \boldsymbol{D}_{12} & \boldsymbol{D}_{13} \\ \hline \boldsymbol{D}_{21} & \boldsymbol{D}_{22} & \boldsymbol{D}_{23} \\ \hline \boldsymbol{D}_{31} & \boldsymbol{D}_{32} & \boldsymbol{D}_{33} \end{array}\right]\begin{array}{c} p \\ q \\ r \end{array}, \qquad \boldsymbol{A} = \left[\begin{array}{c|c|c} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} & \boldsymbol{A}_{13} \\ \hline \boldsymbol{A}_{21} & \boldsymbol{A}_{22} & \boldsymbol{A}_{23} \\ \hline \boldsymbol{A}_{31} & \boldsymbol{A}_{32} & \boldsymbol{A}_{33} \end{array}\right]\begin{array}{c} p \\ q \\ r \end{array}$$
$$\qquad l \qquad m \qquad n \qquad\qquad\qquad\qquad l \qquad m \qquad n \qquad \ldots \quad (3.1)$$

where $pl > 0$, *and*

(i) $\boldsymbol{A}_{11} > \boldsymbol{0}$ *and all other elements of* $\boldsymbol{A}$ *are non-positive,*

(ii) $\boldsymbol{A}_{12} = \boldsymbol{0}$, $\boldsymbol{A}_{21} = \boldsymbol{0}$,

(iii) *each row of* $\boldsymbol{A}_{31}$ *contains at least one negative element, and each column of* $\boldsymbol{A}_{13}$ *contains at least one negative element,*

(iv) $\boldsymbol{D}_{32} = \boldsymbol{0}$, $\boldsymbol{D}_{33} = \boldsymbol{0}$.

Note : If any of q, r, m, n is zero, the corresponding row and/or column of both $\boldsymbol{D}$ and $\boldsymbol{A}$ will be absent in the above partitions.

Theorem 1 : *A* $t\times t$ *matrix* $\boldsymbol{P} \geqslant \boldsymbol{0}$ *is d.s.s. iff it satisfies the following condition* $C1$:

Condition $C1$: For $1 \leqslant k, l \leqslant t$, and any $k\times l$ submatrix $\boldsymbol{B}$ of $\boldsymbol{P}$, $(\boldsymbol{B}) \geqslant k+l-t$.

Proof : First note that $\boldsymbol{P}$ is d.s.s. iff any of its permutational transform is d.s.s. Moreover, $\boldsymbol{P}$ satisfies the condition $C1$ iff any of its permutational transform satisfies $C1$.

If $\inf_{S\in\mathcal{D}} (\boldsymbol{S}-\boldsymbol{P})_+ = 0$, we are done. Next we apply our Basic Lmmma to $\boldsymbol{P}$.

Suppose $\inf_{S \in \mathcal{S}} (S-P)_+ > 0$ and P satisfies the condition $C1$, Partition P as in the basic lemma. Then

$$\begin{pmatrix} P_{11} \\ P_{21} \end{pmatrix} < \begin{pmatrix} D_{11} \\ D_{21} \end{pmatrix} = l-(D_{31})$$
$$= l-r$$
$$= l-(t-p-q) = l+(p+q)-t.$$

This contradicts the condition $C1$. Hence P is d.s.s.

Suppose now P is d.s.s. Then there exists a d.s. matrix D such that $P \geqslant D$. Consider a $k \times l$ submatrix P_1 of P, and without loss of generality suppose

$$P = \left[\begin{array}{c|c} P_1 & P_2 \\ \hline P_3 & P_4 \end{array}\right]\begin{matrix} k \\ t-k \end{matrix}, \qquad D = \left[\begin{array}{c|c} D_1 & D_2 \\ \hline D_3 & D_4 \end{array}\right]\begin{matrix} k \\ t-k \end{matrix}.$$
$$\qquad\; l \quad t-l \qquad\qquad\qquad\quad l \quad t-l$$

Then

$$(P_1) \geqslant (D_1) = k-(D_2)$$
$$= k-[(t-l)-(D_4)]$$
$$= k+l-t+(D_4) \geqslant k+l-t.$$

Theorem 2: *A non-negative matrix* P: $t \times t$ *is d.s.s. iff* $yP \prec^{w} y$ *for all* $y \in R^t_+$.

Proof: Suppose P is d.s.s. Then there exists a d.s. matrix D such that $P \geqslant D$. Hence for any $y \in R^t_+$

$$yP \geqslant yD.$$

Thus

$$yP \prec^{w} yD.$$

Since $yD \prec y$ we have $yP \prec^{w} y$.

Suppose $yP \prec^{w} y$ for all $y \in R^t_+$. We shall show that P satisfies the condition $C1$, and hence, by virtue of Theorem 1, P is d.s.s.

Consider a $k \times l$ submatrix P_1 of P. Without any loss of generality, suppose

$$P = \left[\begin{array}{c|c} P_1 & P_2 \\ \hline P_3 & P_4 \end{array}\right]\begin{matrix} k \\ t-k \end{matrix}$$
$$\; l \quad t-l$$

Let $\delta_k = (l_k, 0) : 1\times t$, where 1_k is the $1\times k$ vector with all elements equal to 1. Since $\delta_k P \prec^w \delta_k$, we have

$$\sum_{i=1}^{l} x_{(i)} \geqslant \max(0, l+k-t),$$

where $\delta_k P = (1_k P_1 \quad 1_k P_2) \equiv (x_1, \ldots, x_t)$.

But
$$(P_1) = (1_k P_1) \geqslant \sum_{i=1}^{l} x_{(i)}.$$

Hence
$$(P_1) \geqslant \max(0, k+l-t) \geqslant k+l-t.$$

Marshall and Olkin have shown (2.D.4, page 31) that if $P \geqslant 0$ is d.s.s. then P satisfies the following condition :

Condition C2 : For $1 \leqslant k, l \leqslant t$, and $t\times t$ matrix P

(the sum of elements in any k columns of P)$-k$

$\geqslant$ (the sum of elements in the intersection of the k columns

and any l rows of P)$-l$.

It is easily seen that the above condition is equivalent to the condition $C1$. In this connection, Marshall and Olkin (1979) have introduced the following condition which is also equivalent to the condition $C1$:

Condition C3 : For a $t\times t$ matrix P and for $1 \leqslant k, l \leqslant t$

(the sum of elements in any l rows of P)$-l$

$\geqslant$ (the sum of elements in the intersection of the l rows and

any k columns)$-k$.

4. Proof of the basic lemma

Let D be a minimizer with respect to P. Then $(D-P)_+ > 0$. By suitable permutational transformation of $A = D-P$ it is possible to get a left-hand upper corner block of this matrix such that

(a) each row of this block has at least one positive element,

(b) each column of this block has at least one positive element, and

(c) all elements of the matrix outside this block are non-positive.

We shall show that all elements of this block are positive, or can be made to be positive by applying suitable invariant δ-transformations on D. Such a block will be called the "positive block" A_{11}. It is clear that $pl > 0$.

If the above block has only one row and/or only one column (i.e., $p = 1$ and/or $l = 1$) the block is trivially the positive block. Otherwise, consider an element x of this block which is not positive. Then there exists a 2×2 submatrix of this block which can be expressed, after suitable rearrangements of its rows and columns, as follows :

$$B = \begin{bmatrix} u & v \\ x & w \end{bmatrix}, u > 0, w > 0. \qquad \dots \quad (4.1)$$

Next we use a δ-transformation on D such that the above submatrix B of A is changed to

$$B_\delta \equiv \begin{bmatrix} u-\delta & v+\delta \\ x+\delta & w-\delta \end{bmatrix}, \qquad \dots \quad (4.2)$$

while all other elements of A are unchanged.

If $x < 0$, then $\delta > 0$ can be suitably chosen so that $x+\delta < 0$, and

$$(B_\delta)_+ < (B)_+. \qquad \dots \quad (4.3)$$

This contradicts the assumption that D is a minimizer. Thus $x \geqslant 0$. Similarly $v \geqslant 0$.

If $x = 0 = v$, δ can be so chosen that both $u-\delta$ and $w-\delta$ are positive, and

$$(B_\delta)_+ = (B)_+.$$

Thus this δ-transformation is invariant, and the resulting d.s. matrix D is also a minimizer. In this way all non-positive elements of this block can be changed to positive elements. We shall denote such a block by A_{11}.

It is possible to partition A and corresponding D, (by suitable permutational transformations, if necessary) as in the basic lemma, so that (i), (ii) and (iii) hold. If $r = 0$ there is noting else to prove.

Next we shall show that $D_{33} = 0$ when $r > 0$, $n > 0$. Suppose there is an element d_{ij} of D lying in the block D_{33} which is not zero. Then there exist an element a_{ik} in A_{31} and an element a_{sj} in A_{13} such that both a_{ik} and a_{sj} are negative. Note that $a_{sk} > 0$ and $a_{ij} \leqslant 0$. Now consider the 2×2 submatrix

$$B = \begin{bmatrix} a_{sk} & a_{sj} \\ a_{ik} & a_{ij} \end{bmatrix}.$$

It is possible to find a δ-transformation on D such that the above submatrix of $A = D-P$ is changed to

$$B_\delta = \begin{bmatrix} a_{sk}-\delta & a_{sj}+\delta \\ a_{ik}+\delta & a_{ij}-\delta \end{bmatrix}.$$

while the other elements of A remain unchanged. It is possible to choose $\delta > 0$ so that $a_{tk}+\delta < 0$, $a_{sj}+\delta < 0$ and $a_{sk}-\delta > 0$. Hence

$$(B_\delta)_+ < (B)_+,$$

which contradicts that D is a minimizer. Hence $D_{33} = 0$.

Suppose $r > 0$, $m > 0$, and there exists an element d_{tj} of D lying in the block D_{32} which is not zero. Then there exists a 2×2 submatrix of A given by

$$B = \begin{bmatrix} a_{sk} & a_{sj} \\ a_{tk} & a_{tj} \end{bmatrix}$$

such that $a_{tk} < 0$ is in A_{31}, $a_{sj} = 0$ is in A_{12}, and $a_{sk} > 0$ is in A_{11}. There exists a δ-transformation on D such that the above submatrix of $A = D-P$ is changed to

$$B_\delta = \begin{bmatrix} a_{sk}-\delta & a_{sj}+\delta \\ a_{tk}+\delta & a_{tk}-\delta \end{bmatrix}$$

while the other elements of A remain unchanged, and $a_{tk}+\delta < 0$, $a_{sk}-\delta > 0$. Then

$$(B_\delta)_+ = (B)_+. \qquad \ldots \quad (4.4)$$

Such a δ-transformation is invariant and it keeps the structure of A_{31} (satisfying (iii)) unchanged, while changing a_{sj} to a positive element. In this way, all the elements in A_{12} lying in the j-th column can be changed to positive elements by suitable invariant δ-transformations. Then this entire column can be annexed to the block

$$\begin{bmatrix} A_{11} \\ A_{21} \\ A_{31} \end{bmatrix},$$

thereby extending the positive block A_{11} by one more column. This process is continued until the remaining elements of D_{32} are zero. This process leads to the following structure of A :

$$A = \left[\begin{array}{cc|c|c} A_{11} & \boxed{\text{annexed columns}} & A_{12} & A_{13} \\ \hline A_{21} & \boxed{\phantom{\text{annexed columns}}} & A_{22} & A_{23} \\ \hline A_{31} & \boxed{\phantom{\text{annexed columns}}} & A_{32} & A_{33} \end{array}\right] \qquad \ldots \quad (4.5)$$

Although the new A_{31} would satisfy (iii), the new A_{21} may not be $\mathbf{0}$. If there is any row in (new) A_{21} which contains at least one negative element, then that entire row of A would be annexed to $[A_{31}\ A_{32}\ A_{33}]$. We then would get the following structure of A :

$$A = \left[\begin{array}{l|c|c} A_{11} & A_{12} & A_{13} \\ \hline A_{21} & A_{22} & A_{23} \\ \hline \boxed{\text{annexed rows}} & \boxed{(1)} & \boxed{(2)} \\ A_{31} & A_{32} & A_{33} \end{array}\right] \qquad \ldots \quad (4.6)$$

The new A_{31} would still satisfy (iii). Proceeding as before, it can be shown that the elements of D corresponding to the block (2), as indicated in (4.6), are all zero. However, the elements of D corresponding to the block (1), given in (4.6), may not be all zero.

At this stage, we repeat the above entire process until $D_{32} = \mathbf{0}$. This can be accomplished since the above process reduces the number of columns in A_{32} while possibly increasing its number of rows subsequently. But at some stage, there may not be any columns left in A_{32} so that no new rows may be annexed. The final partitions of A and D would then satisfy all the conditions (i), (ii), (iii) and (iv) in the basic lemma.

Remark 1 : Consider the matrix

$$P = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & -1 \end{bmatrix}$$

although P satisfies condition $C1$, P is not non-negative.

Remark 2 : As pointed out by Dr. Rahul Mukerjee, our basic lemma and consequently Theorem 3.1 could easily be generalised to the case when P and $D \geqslant 0$ are rectangular $m \times n$ matrices with specified row sums and column sums of D. In that case our condition $C1$ needs to be modified accordingly, and our proofs would go through except for trivial changes. Furthermore, the above problem could be seen as a problem in transportation theory.

The above problem was solved by Mirsky (1968) when the elements of p and D are all integers. However, Mirsky's complete proof would be exceedingly long. Mirsky has also pointed out in (Mirsky, 1971, p. 211) that his result for integral matrices could be extended to real matrices. It has been noted by Mirsky (1971, p. 213) that a slightly more general result was obtained by Kellerer (1961, 1964) from measure-theoretic viewpoint.

Theorem 1 again appears in a paper by Cruse (1975) where the author has also mentioned the generalisation to rectangular matrices. Our proof of Theorem 1 is entirely different from the proof given in Cruse (1975). Moreover, our Basic Lemma provides a new characterisation of matrices which are not d.s.s. Although Theorem 2 follows from Theorem 1, the statement of this theorem along with a proof is not available in the existing literature.

References

Cruse, A. B. (1975): A proof of Fulkerson's characterization of permutation matrices. *Linear Algebra and Its Applications* **12**, 21-28.

Kellerer, H. G. (1961): Funktionen auf Produktraumen mit vorgegebenen marginal-funktionen. *Math. Annalen*. **144**, 323-344.

——— (1964): Allgemeine systeme von repräsentaten. *Z. Wahrsch. Theorie*, **2**, 306-309.

Marshall, A. W. and Olkin, I. (1979): *Inequalities: Theory of Majorization and Its Applications*, Academic Press, New York.

Mirsky, L. (1968): Combinatorial theorems and integral matrices. *J. Combinatorial Theory*, **5**, 30-44.

——— (1971): *Transversal Theory*, Academic Press, New York.

Paper received : December, 1983.

Revised : December, 1984.