PROOFS OF TWO THEOREMS ON DOUBLY-STOCHASTIC MATRICES

L. MIRSKY

1. All numbers considered below are real. We shall write

$$(y_1, \cdots, y_n) \prec (x_1, \cdots, x_n) \tag{1}$$

to indicate that

$$\bar{y}_1 + \cdots + \bar{y}_k \leqq \bar{x}_1 + \cdots + \bar{x}_k \qquad (k = 1, \cdots, n-1),$$
$$y_1 + \cdots + y_n = x_1 + \cdots + x_n;$$

here $\bar{x}_1, \cdots, \bar{x}_n$ denote the numbers $x_1, \cdots, x_n$ arranged in non-ascending order of magnitude, and similarly for the y's. All matrices we consider are of type $n \times n$, unless the contrary is obvious from the context; a matrix is called *doubly-stochastic* (d.s.) if its elements are non-negative numbers such that the sum of elements in each row and in each column is equal to 1. The set of all $n \times n$ d.s. matrices will be denoted by $\mathfrak{D}_n$. Further, we denote by $\mathfrak{S}_n$ the set of the $n!$ permutations of $1, \cdots, n$; and by π_0 the identical permutation in $\mathfrak{S}_n$. $H(x_1, \cdots, x_n)$ is the convex hull of the $n!$ vectors $(x_{\pi 1}, \cdots, x_{\pi n})$, $\pi \in \mathfrak{S}_n$. Finally, δ_{kj} denotes the Kronecker delta.

The object of the present note is to give simple proofs of two results on d.s. matrices. The first of these results is due to Hardy, Littlewood, and Pólya ([**3**, Theorem 46]; for an alternative proof see [**5**, §§5–8]); the second is due to G. Birkhoff [**2**, §1].

THEOREM I. *The relation* (1) *is necessary and sufficient for the existence of a doubly-stochastic matrix D such that*

$$(y_1, \cdots, y_n) = (x_1, \cdots, x_n)D.$$

THEOREM II. *The set of doubly-stochastic matrices is identical with the convex hull of the set of permutation matrices.*

I should like to express my thanks to the referee for pointing out an error in my original proof of Theorem II, and to Dr. H. Burkill for useful criticism.

2. We begin with a few preliminary results. Lemma 2 will be used in the proof of Theorem I, and Lemmas 1 and 4 in the proof of Theorem II.

Received by the editors September 4, 1957 and, in revised form, November 21, 1957.

LEMMA 1. *The vector* $(x_1, \cdots, x_m)$ *lies in the convex hull of the vectors* $(x_{k1}, \cdots, x_{km})$, $k=1, \cdots, r$, *if and only if, for any numbers* $a_1, \cdots, a_m$,

$$\sum_{j=1}^{m} a_j x_j \leqq \max_{1 \leqq k \leqq r} \sum_{j=1}^{m} a_j x_{kj}.$$

For details of this (intuitively obvious) result, see [**1**, pp. 23–24].

LEMMA 2. *If* (1) *is satisfied, then* $(y_1, \cdots, y_n) \in H(x_1, \cdots, x_n)$.

This was shown by R. Rado [**6**, §1] to be an almost immediate consequence of Lemma 1. Another derivation, given by A. Horn [**4**, p. 621], depends on the use of Theorem II.

LEMMA 3. *Let* (ξ_{kj}) *be a d.s. matrix other than the unit matrix. Then there exists a permutation* $\pi \in \mathfrak{S}_n$, $\pi \neq \pi_0$ *such that the numbers* $\xi_{k,\pi k}$ $(1 \leqq k \leqq n;\ k \neq \pi k)$ *are all different from zero.*

Assume, on the contrary, that, for every permutation $\pi \neq \pi_0$, the set of numbers $\xi_{k,\pi k}$ $(1 \leqq k \leqq n;\ k \neq \pi k)$ contains at least one zero. For $t \geqq 0$, write $\eta_{kj} = (\xi_{kj} + t\delta_{kj})/(1+t)$. Then (η_{kj}) is again a d.s. matrix and, for every $\pi \neq \pi_0$, the set of numbers $\eta_{k,\pi k}$ $(1 \leqq k \leqq n;\ k \neq \pi k)$ contains at least one zero. Hence

$$\det_{1 \leqq k, j \leqq n} (\eta_{kj}) = \eta_{11} \cdots \eta_{nn} + \sum_{\pi \neq \pi_0} (\pm \eta_{1,\pi 1} \cdots \eta_{n,\pi n}) = \eta_{11} \cdots \eta_{nn},$$

and so

$$\det_{1 \leqq k, j \leqq n} (\xi_{kj} + t\delta_{kj}) = \prod_{k=1}^{n} (\xi_{kk} + t). \tag{2}$$

This relation holds for $t \geqq 0$, and is therefore an identity in t. Hence the characteristic roots of (ξ_{kj}) are $\xi_{11}, \cdots, \xi_{nn}$. But every d.s. matrix has 1 as a characteristic root; thus $\xi_{kk} = 1$ for some value of k, say $\xi_{nn} = 1$. The identity (2) therefore implies the further identity

$$\det_{1 \leqq k, j \leqq n-1} (\xi_{kj} + t\delta_{kj}) = \prod_{k=1}^{n-1} (\xi_{kk} + t).$$

Hence $\xi_{11}, \cdots, \xi_{n-1,n-1}$ are the characteristic roots of the d.s. matrix (ξ_{kj}), $k, j = 1, \cdots, n-1$; and so at least one of them is equal to 1. Continuing in this manner, we see that $\xi_{11} = \cdots = \xi_{nn} = 1$, i.e. contrary to our hypothesis, the original matrix (ξ_{kj}) is the unit matrix.

LEMMA 4. *For any matrix* (a_{kj}), *we have*

$$(3) \qquad \sup_{\mathfrak{D}_n} \sum_{k,j=1}^{n} a_{kj}x_{kj} = \max_{\pi\in\mathfrak{S}_n} \sum_{k=1}^{n} a_{k,\pi k},$$

where the upper bound is taken with respect to all d.s. matrices (x_{kj}).

Since neither side of (3) is affected by a permutation of the columns of (a_{kj}), there is no loss of generality in supposing that the right-hand side of (3) is equal to $a_{11}+\cdots+a_{nn}$. It is then sufficient to prove that

$$(4) \qquad \sup_{\mathfrak{D}_n} \sum_{k,j=1}^{n} a_{kj}x_{kj} \leqq \sum_{k=1}^{n} a_{kk}.$$

Let $\epsilon>0$, and put $b_{kj}=a_{kj}+\epsilon\delta_{kj}$. Then

$$(5) \qquad \sum_{k=1}^{n} b_{kk} - \sum_{k=1}^{n} b_{k,\pi k} \geqq 2\epsilon \qquad (\pi \in \mathfrak{S}_n,\ \pi \neq \pi_0).$$

Further, let

$$(6) \qquad \sup_{\mathfrak{D}_n} \sum_{k,j=1}^{n} b_{kj}x_{kj} = \sum_{k,j=1}^{n} b_{kj}\xi_{kj},$$

where (ξ_{kj}) is a d.s. matrix. Assume that (ξ_{kj}) is not the unit matrix and denote by π $(\neq\pi_0)$ some permutation in $\mathfrak{S}_n$ which satisfies the conclusion of Lemma 3. Define η_{kj} as $\xi_{kj}+\delta$ if $j=k\neq\pi k$, as $\xi_{kj}-\delta$ if $j=\pi k\neq k$, and as ξ_{kj} in all other cases. Then, if δ is chosen as a sufficiently small positive number, (η_{kj}) is again a d.s. matrix; and, using (5), we have

$$\sum_{k,j=1}^{n} b_{kj}\eta_{kj} - \sum_{k,j=1}^{n} b_{kj}\xi_{kj} = \delta\left\{\sum_{k=1}^{n} b_{kk} - \sum_{k=1}^{n} b_{k,\pi k}\right\} \geqq 2\delta\epsilon > 0.$$

But this contradicts (6); hence (ξ_{kj}) is the unit matrix and, in view of (6), we have

$$\sup_{\mathfrak{D}_n} \sum_{k,j=1}^{n} a_{kj}x_{kj} \leqq \sup_{\mathfrak{D}_n} \sum_{k,j=1}^{n} b_{kj}x_{kj} = \sum_{k=1}^{n} b_{kk} = \sum_{k=1}^{n} a_{kk} + \epsilon n.$$

Since this holds for an arbitrarily small ϵ, (4) follows and the lemma is therefore proved.

3. We next come to the proof of Theorem I. The necessity of the condition (1) is almost immediate (see [**3**, l.c.]) and we may confine ourselves to the proof of sufficiency.

Denote by $K(x_1, \cdots, x_n)$ the set of vectors which can be represented in the form $(x_1, \cdots, x_n)D$, where D is some d.s. matrix. If

D_1, D_2 are d.s. matrices and $0 \leqq t \leqq 1$, then $tD_1+(1-t)D_2$ is again d.s. Therefore $K(x_1, \cdots, x_n)$ is a convex set; and, since

$$(x_{\pi 1}, \cdots, x_{\pi n}) \in K(x_1, \cdots, x_n) \qquad (\pi \in \mathfrak{S}_n),$$

it follows that

$$H(x_1, \cdots, x_n) \subseteq K(x_1, \cdots, x_n). \tag{7}$$

Suppose now that (1) is satisfied. Then, by Lemma 2 and (7), $(y_1, \cdots, y_n) \in K(x_1, \cdots, x_n)$. This establishes the required conclusion.

It may be noted that the following characterization of d.s. matrices is an easy consequence of Lemma 2 and the necessity part of Theorem I: *a matrix D is d.s. if and only if* $(x_1, \cdots, x_n)D \prec (x_1, \cdots, x_n)$ *for every choice of the x's.*

4. Finally, we give the proof of Theorem II. Any matrix contained in the convex hull of the set of permutation matrices is obviously d.s.; it remains to establish the converse.

Let (d_{kj}) be a d.s. matrix. If (a_{kj}) is any matrix, then, by Lemma 4,

$$\sum_{k,j=1}^{n} a_{kj} d_{kj} \leqq \max_{\pi \in \mathfrak{S}_n} \sum_{k=1}^{n} a_{k,\pi k}.$$

Hence, by Lemma 1, (d_{kj}) lies in the convex hull of the set of permutation matrices.

References

1. T. Bonnesen and W. Fenchel, *Theorie der konvexen Körper*, Berlin, 1934.
2. G. Birkhoff, *Tres observaciones sobre el algebra lineal*, Universidad Nacional de Tucumán, Revista. Serie A. vol. 5 (1946) pp. 147–151.
3. G. H. Hardy, J. E. Littlewood, and G. Pólya, *Inequalities*, Cambridge, 1934.
4. A. Horn, *Doubly stochastic matrices and the diagonal of a rotation matrix*, Amer. J. Math. vol. 76 (1954) pp. 620–630.
5. A. Ostrowski, *Sur quelques applications des fonctions convexes et concaves ausens de I. Schur*, J. Math. Pures Appl. (9) vol. 31 (1952) pp. 253–292.
6. R. Rado, *An inequality*, J. London Math. Soc. vol. 27 (1952) pp. 1–6.

University of Sheffield, England