ON A RESULT OF S. SHERMAN CONCERNING DOUBLY STOCHASTIC MATRICES

SHMUEL SCHREIBER

I. Doubly stochastic (d.s.) matrices are defined as $n\times n$ real matrices $P(p_{ij})$ having non-negative entries and unit row—and columns sums, thus:

(1) $p_{ij} \geqq 0,\ 1 \leqq i, j \leqq n;$ (2) $\sum_{j=1}^{n} p_{ij} = 1,\ i \leqq i \leqq n;$

(3) $$\sum_{i=1}^{n} p_{ij} = 1,\ 1 \leqq j \leqq n.$$

S. Sherman [1] introduces a partial ordering of d.s. matrices defining, for two d.s. matrices P_1 and P_3 of the same order,

(4) $$P_1 < P_3$$

if, and only if, there exists a d.s. matrix P such that

(5) $$P_1 = P_2P_3.$$

He also defines, following Hardy, Littlewood and Polya [3] a partial ordering of n-dimensional real vectors: $\boldsymbol{a}<\boldsymbol{b}$ if, and only if, there exists a d.s. matrix P such that $a=P\boldsymbol{b}$. In the above named article, the author investigates a conjecture of S. Kakutani to the effect that, if two d.s. matrices are such that $P_1\boldsymbol{a}<P_3\boldsymbol{a}$ for every real vector $\boldsymbol{a}$, then $P_1<P_3$. For this purpose he constructs a linear mapping of all vectors of the form $P_3\boldsymbol{a}$ onto vectors $P_1\boldsymbol{a}$, and extends this mapping to a mapping ψ of the whole euclidean n-space onto the range of P_1. However, A. Horn (as quoted in [2]) points out, with the aid of a counter-example, that such an extension does not always preserve the properties of a mapping effected by a d.s. matrix, and thus Kakutani's conjecture is not true without restriction. On the other hand, if every vector is of the form $P_3\boldsymbol{a}$, i.e. if the matrix P_3 is regular, it is readily verified that Sherman's construction has the required properties and thus in this case, Kakutani's conjecture holds true. It is the object of the present note to give an elementary proof of this fact, as well as to derive a more or less intuitive description of Sherman's partial ordering of d.s. matrices.

Received by the editors December 2, 1957.

II. Let A denote the class of matrices satisfying (1), B the class of matrices satisfying (2), C the class of matrices satisfying (3). $A\cap B$ is the class of stochastic matrices occurring in the theory of probability (Markoff chains). Let $\mathbf{e}$ denote the column vector, all of whose components are unity and $\mathbf{f}=\mathbf{e}^T$ the corresponding row vector. Then $P\in B$ if, and only if $P\mathbf{e}=\mathbf{e}$. From $P_1\mathbf{e}=\mathbf{e}$ and $P_2\mathbf{e}=\mathbf{e}$ we have $P_1P_2\mathbf{e}=P_2\mathbf{e}=\mathbf{e}$. Moreover, if P is regular, then from $P\mathbf{e}=\mathbf{e}$ it follows that $e=P^{-1}\mathbf{e}$. Thus

LEMMA 1. *The matrices of class B form a semigroup under multiplication; the regular matrices of class B form a group.*

LEMMA 2. *The matrices of class C form a semigroup under multiplication; the regular matrices of class C form a group.*

The intersection of the two groups above is obviously formed by the regular matrices of $B\cap C$. This group contains the semigroup of all regular d.s. matrices.

Let now P be a matrix, $\mathbf{v}$ a column vector, and $\mathbf{w}=P\mathbf{v}$. Then $\mathbf{w}$ is a linear combination of the column vectors of P (the coefficient being the components of $\mathbf{v}$). If $\mathbf{v}$ has non-negative components with unit sum, in which case we shall call $\mathbf{v}$ a *stochastic vector*, then $\mathbf{w}$ is a *convex* linear combination of the column vectors of P. If moreover P is regular, $\mathbf{v}$ is uniquely determined from the equation $P\mathbf{v}=\mathbf{w}$; in this case, if $\mathbf{w}$ is a *convex* linear combination of the column vectors of P, $\mathbf{v}$ is a stochastic vector. This argument can be applied to each column vector of the right-hand factor in the product of two matrices and yields

LEMMA 3. *Let P and Q be matrices and $R=PQ$. Then each column of R is a linear combination of the columns of P. If $Q\in A\cap C$, each column of R is a convex linear combination of the columns of P. If P is regular, Q is uniquely determined from the equation $R=PQ$; in this case, if each column of R is a convex linear combination of the columns of P, then $Q\in A\cap C$.*

By a similar argument (or by setting $R^T=Q^TP^T$ in Lemma 3) one obtains

LEMMA 4. *Let P and Q be matrices and $R=PQ$. Then each row of R is a linear combination of the rows of Q. If $P\in A\cap B$, each row of R is a convex linear combination of the rows of Q. If Q is regular, P is uniquely determined from the equation $R=PQ$; in this case, if each row of R is a convex linear combination of the rows of Q, then $P\in A\cap B$.*

In particular, for d.s. matrices, one has

LEMMA 5. *Let P_2 and P_3 be d.s. matrices and $P_1 = P_2P_3$. Then P_1 (by Lemma 2.3) is a d.s. matrix; each column of P_1 is a convex linear combination of the columns of P_2 and each row of P_1 is a convex linear combination of the rows of P_3.*

We now have the following

PROPOSITION. *Let P_1 and P_3 be d.s. matrices and let P_3 be regular. Then, for $P_1 < P_3$, it is necessary and sufficient that each row vector of P_1 be a convex linear combination of the row vectors of P_3.*

PROOF. Necessity follows from Lemma 5. P_3 being regular, $P_3^{-1} \in C$ by Lemma 2. Again by Lemma 2, the matrix $P_2 = P_1P_3^{-1}$ belongs to C. But P_2 belongs to $A \cap B$ as well, by Lemma 4, thus P_2 is a d.s. matrix and the condition is sufficient. Observe next that if $\mathbf{u}_k$ is the kth column vector of the unit matrix U, $\mathbf{a} < \mathbf{u}_k$ if and only if $\mathbf{a}$ is stochastic. Indeed, if P is any d.s. matrix, then $P\mathbf{u}_k$ is the kth column vector of P. We can now prove Kakutani's conjecture, with the restriction to regular matrices P.

THEOREM (S. SHERMAN). *Let P_1 and P_3 be d.s. matrices, and let P_3 be regular. If $P_1\mathbf{a} < P_3\mathbf{a}$ for every $\mathbf{a}$, then $P_1 < P_3$.*

PROOF. P_3^{-1} exists by hypothesis. Let $\mathbf{v}_k$ denote the kth column vector of P_3^{-1}. Then $P_1\mathbf{v}_k < P_3\mathbf{v}_k = \mathbf{u}_k$ in the above notation, thus $P_1\mathbf{v}_k$ is stochastic. Therefore the matrix $P_2 = P_1P_3^{-1}$ is of class $A \cap C$. On the other hand, by Lemma 1, $P_3^{-1} \in B$ and $P_1P_3^{-1} \in B$, therefore P_2 is a d.s. matrix satisfying $P_1 = P_2P_3$, q.e.d.

III. Geometrically, things may be described as follows: Let π be the hyperplane in euclidean n-space given by $\sum_{i=1}^{n} x_i = 1$. Call the $n-1$ dimensional simplex on π, determined by the end points of the n positive unit vectors, S. Then the row vectors of every matrix $P \in B$ have their end points on π, and if $P \in A \cap B$, those end points are the vertices of a (possibly degenerate) simplex on S. The same is true of the column vectors of a matrix belonging to C or to $A \cap C$, respectively. If, apart from $P \in A \cap B$, we have $P \in C$ as well, then the row vectors of P add up to $f(1, 1, \cdots, 1)$. Therefore the centroid of the simplex, whose vertices are the end points of the row vectors of a d.s. matrix, coincides with the centroid of S, i.e. $(1/n, 1/n, \cdots, 1/n)$. The same is true of the simplex corresponding to the column vectors.

Following a suggestion of Dr. S. N. Afriat, we may define the *right stochastic* range of a matrix P to be the set of all vectors $P\mathbf{a}$, where $\mathbf{a}$ is a stochastic column vector, and analogously the *left stochastic range*. It is readily seen that the end points of all vectors from the *left* stochastic range of a matrix $P \in A \cap B$, and of the *right* stochastic range of a matrix $P \in A \cap C$ form the convex closure of the simplices described above. It is seen, moreover, that if the (left or right) stochastic range of a sequence of d.s. matrices, such as the powers of a given d.s. matrix, converges to a single vector, that vector has to be the one whose end point is the centroid. This fact has various applications; see for instance Dvoretzky and Wolfowitz [**4**], Feller [**5**, p. 327 etc.].

We observe now that the partial order relation (4)–(5) can be extended to simply stochastic matrices, of class $A \cap B$. Then Lemma 4 gives:

Let R and Q be two stochastic matrices. Then there exists a stochastic matrix P such that $R = PQ$ if, and only if the simplex corresponding to the row vectors of R is contained in the one corresponding to the row vectors of Q; or, alternatively, if the left stochastic range of R is contained in the left stochastic range of Q.

Thus, to the partial order relation, set up by S. Sherman in [**1**], there corresponds a partial order relation by inclusion of simplices.

Now the above partial order relation is invariant under matrix multiplication from the right. If $R_1 = R_2R_3$, then $R_1R_4 = R_2R_3R_4$. In the same way one could define a *left invariant* partial ordering of d.s. matrices, putting $R_1 < R_2$ if, and only if, there exists a d.s. matrix R_3 such that $R_1 = R_2R_3$. Such a definition could then be extended to matrices of class $A \cap C$; the partial order thus set up would correspond to partial ordering by inclusion of simplices from end points of column vectors, or by inclusion of right stochastic ranges.

References

1. S. Sherman, *On a conjecture concerning doubly stochastic matrices*, Proc. Amer. Math. Soc. vol. 3 (1952) pp. 511–513.

2. ———, *A correction to "On a conjecture concerning doubly stochastic matrices"*, Proc. Amer. Math. Soc. vol. 5 (1954) pp. 998–999.

3. G. H. Hardy, J. E. Littlewood and G. Polya, *Inequalities*, Cambridge, 1934.

4. A. Dvoretzky and J. Wolfowitz, *Sums of random integers reduced modulo m*, Duke Math. J. vol. 18 (1951) pp. 501–507.

5. W. Feller, *An introduction to probability theory and its applications*, New York, 1950.

The Hebrew University, Jerusalem