JOURNAL OF RESEARCH of the National Bureau of Standards—B. Mathematical Sciences
Vol. 80B, No. 4, October–December 1976

An Inequality for Doubly Stochastic Matrices*

Charles R. Johnson and R. Bruce Kellogg****

Institute for Basic Standards, National Bureau of Standards, Washington, D.C. 20234

(June 30, 1976)

Interrelated inequalities involving doubly stochastic matrices are presented. For example, if B is an n by n doubly stochastic matrix, x any nonnegative vector and $y = Bx$, then $x_1x_2 \cdot \cdots \cdot x_n \leq y_1y_2 \cdot \cdots \cdot y_n$. Also, if A is an n by n *nonnegative matrix and* D and E are positive diagonal matrices such that $B = DAE$ is doubly stochastic, then $\det DE \geq \rho(A)^{-n}$, where $\rho(A)$ is the Perron-Frobenius eigenvalue of A. The relationship between these two inequalities is exhibited.

Key words: Diagonal scaling; doubly stochastic matrix; Perron-Frobenius eigenvalue.

An n by n entry-wise nonnegative matrix $B = (b_{ij})$ is called row (column) *stochastic* if $\sum_{j=1}^{n} b_{ij} = 1$ for all $i = 1,\cdots,n$ $\left(\sum_{i=1}^{n} b_{ij} = 1 \text{ for all } j = 1,\cdots,n\right)$. If B is simultaneously row and column stochastic then B is said to be *doubly stochastic*. We shall denote the Perron-Frobenius (maximal) eigenvalue of an arbitrary n by n entry-wise nonnegative matrix A by $\rho(A)$. Of course, if A is stochastic, $\rho(A) = 1$.

It is known precisely which n by n nonnegative matrices may be diagonally scaled by positive diagonal matrices D, E so that

$$B = DAE \tag{1}$$

is doubly stochastic. If there is such a pair D, E, we shall say that A has property (*). In this event it is our interest to obtain inequalities on D and E. In the process, certain related inequalities for doubly stochastic matrices are noticed.

It was first realized by Sinkhorn [4][1] that if A is entry-wise positive and square, then A has property (*). The proof amounts to showing that the process of alternately scaling A to produce a row stochastic matrix, and then a column stochastic matrix, and then continuing the process, actually converges to a doubly stochastic matrix. The hypothesis of positivity, however, can be weakened somewhat. If there exists no single permutation matrix P such that

$$P^TAP = \begin{pmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{pmatrix}$$

where A_{11} and A_{22} are square, then A is called irreducible. If there is a pair of permutation matrices P, Q such that $C = PAQ$, then we shall say that A and C are *equivalent*. If, further, A is equivalent to no matrix of the form

AMS Subject Classification: 15A45, 15A48, 15A51, 65F35.

* Work supported in part by NSF grant GP20555.

** Present address: Institute for Physical Science and Technology, University of Maryland, College Park, Md. 20742.

[1] Figures in brackets indicate the literature references at the end of this paper.

(2) $$\begin{pmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{pmatrix}$$

where A_{11} and A_{22} are square, then A is termed *completely irreducible.* It is an easy calculation to show that A has property (*) if and only if each matrix equivalent to A does, and it is equally clear that for A to have property (*) it must have the zero-nonzero sign pattern of a doubly stochastic matrix. For example, if A has property (*) and if A is of form (2) this means we must then have $A_{21} = 0$. It has further been shown [1] that when A is completely irreducible the alternate scaling process of Sinkhorn still converges and thus A has property (*). Since property (*) is preserved under direct summation, we may summarize as follows.

REMARK 1: *A square nonnegative matrix* A *has property* (*) *if and only if* A *is completely irreducible or* A *is equivalent to a direct sum of completely irreducible matrices.*
Thus, property (*) depends only on the zero pattern of A. It is also a straightforward calculation (following [4]) that

REMARK 2: *If* A *has property* (*), *then the product* DE *of* (1) *is unique.*

Our first observation is both necessary for later proofs and of interest by itself.

THEOREM 1: *If* $\mathrm{B} = (\mathrm{b_{ij}})$ *is an* n *by* n *doubly stochastic matrix and* $\mathrm{x} \geq 0$ *is any nonnegative vector, then, for* $\mathrm{y} = \mathrm{Bx}$, *we have*

(3) $$\prod_{i=1}^{n} x_i \leq \prod_{i=1}^{n} y_i.$$

If B *is completely irreducible, equality holds in* (3) *if and only if the right-hand side is* 0 *or all components of* x *are the same. Furthermore, among all irreducible nonnegative square matrices* B *satisfying* $\rho(\mathrm{B}) \leq 1$, *only those diagonally similar to doubly stochastic matrices satisfy* (3) *for all* $\mathrm{x} \geq 0$.

PROOF: From the arithmetic-geometric mean inequality [2]

(4) $$\prod_{i=1}^{n} x_i^{\gamma_i} \leq \sum_{i=1}^{n} \gamma_i x_i$$

where $x = (x_1,\cdots,x_n)^T$ is any nonnegative vector, and $\gamma = (\gamma_1,\cdots,\gamma_n)$ is a vector of nonnegative numbers satisfying $\sum_{i=1}^{n} \gamma_i = 1$. Equality holds in (4) if and only if the x_i's corresponding to nonzero γ_i's are all equal. Now, suppose $B = (b_{ij})$ is row stochastic and $y = Bx$, $x \geq 0$. It follows from (4) that

(5) $$\prod_{j=1}^{n} x_j^{b_{ij}} \leq \sum_{j=1}^{n} b_{ij} x_j = y_i, \qquad \text{for } i = 1,\cdots,n.$$

Taking a product over i of both sides, we arrive at

(6) $$\prod_{j=1}^{n} x_j^{\sum_{i=1}^{n} b_{ij}} \leq \prod_{i=1}^{n} y_i.$$

If B is doubly stochastic, $\sum_{i=1}^{n} b_{ij} = 1$ for each $j = 1,\cdots,n$, and it follows that (3) holds.

To analyze the case of equality, it is clear that equality holds in (3) if either x is a vector of equal components or the right-hand side of (3) is 0. On the other hand, if equality holds in (3) and the right-hand side of (3) is not 0, then equality must hold in (5) for each $i = 1,\cdots,n$. This means that for each i, the x_j's corresponding to nonzero b_{ij}'s are all equal. This, in turn, implies, by virtue of equality holding in (5) for all i, that $y = Q^T x$ for some permutation matrix Q. Since $BQQ^T x = y$, we have that BQ has $Q^T x = y$ as a Perron-Frobenius eigenvector (corresponding to $\rho(BQ) = 1$). If B is completely irreducible, then BQ is irreducible, and, since BQ is doubly stochastic, its Perron-Frobenius eigenspace is one-dimensional (because of the irreducibility) and is spanned by $(1,1,\cdots,1)^T$. Therefore all components of $Q^T x$, and thus of x, are the same. It should be noted that even in case B is completely irreducible it is possible that the right-hand side of (3) be 0 for a nonnegative nonzero vector x. For example, let

$$B = \begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 1/6 & 1/6 & 2/3 \\ 1/2 & 1/2 & 0 \end{pmatrix} \text{ and } x = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Thus, the second statement of the theorem is proven. It should also be noted that the case of equality in (3) may similarly be analyzed if B is equivalent to a direct sum of completely irreducibles. It is enough to assume B is equal to a direct sum of completely irreducibles. Then, in addition to equality occurring in (3) when the right-hand side is 0, equality occurs precisely when the components of x are equal within each piece corresponding to a direct summand of B.

For the third statement, it is enough to assume that B satisfies $\rho(B) = 1$. It then follows from the irreducibility of B that there is a positive diagonal matrix D such that DBD^{-1} is row stochastic (D^{-1} is obtained from the Perron-Frobenius eigenvector of B, which is positive). Since $Bx = y$ if and only if $(DBD^{-1})Dx = Dy$, (3) holds for DBD^{-1} if it holds for B, so that we may as well assume B is row stochastic. Then, if B is not doubly stochastic, some column sum is <1, so let $\theta = \sum_i b_{ij} < 1$. Let $x_i = 1, i \neq j$, and $x_j = 1 + \epsilon$. Then (3) becomes

$$1 + \epsilon \leq \prod_i (1 + b_{ij}\epsilon) = 1 + \theta\epsilon + 0(\epsilon^2),$$

which is impossible if $\epsilon > 0$ is small enough. This completes the proof of the theorem.

It should be noted that essential portions of theorem 1 may also be demonstrated by a maximization argument.

EXAMPLE: The assumption of irreducibility in the third statement of theorem 1 cannot, in general, be relaxed. If $B = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ then B satisfies (3) for all $x \geq 0$ and $\rho(B) = 1$, but B is not similar to a doubly stochastic matrix.

An alternate form of (3) is

COROLLARY 1: *If* $\mathrm{B} = (\mathrm{b}_{ij})$ *is an* n *by* n *doubly stochastic matrix, then for any* n *real numbers* $\mathrm{t}_1, \cdots, \mathrm{t}_n$, *satisfying* $\mathrm{t}_i \geq -1$, $\mathrm{i} = 1, \cdots, \mathrm{n}$, *we have*

$$\prod_{i=1}^{n} (1 + t_i) \leq \prod_{i=1}^{n} (1 + \sum_{j=1}^{n} t_j b_{ij}).$$

Our primary observation concerns row stochastic matrices with property (*).

THEOREM 2: *If* A *is a row stochastic matrix with property* (*) *and* D *and* E *are the positive diagonal matrices guaranteed by* (1) *then*

$$\det \mathrm{DE} \geq 1. \tag{7}$$

Furthermore, equality holds if and only if A *is actually doubly stochastic.*

PROOF: As $B = (b_{ij})$ runs through all n by n doubly stochastic matrices and $F = \operatorname{diag}\{f_1, \cdots, f_n\}$ runs through all positive diagonal matrices, then $A = D^{-1}BF$ runs through all row stochastic matrices with property (*) where $D = \operatorname{diag}\{\sum_{j=1}^{n} b_{1j}f_j, \cdots, \sum_{j=1}^{n} b_{nj}f_j\}$. Thus, since $B = DAE$, where $E = F^{-1}$, it suffices to show that $\det D \geq \det F$. If we denote $(f_1, \cdots, f_n)^T$ by f, this is equivalent to saying that the product of the entries of Bf is greater than or equal to that of f for any positive vector f. This, of course, follows from theorem 1. To analyze the case of equality in (7), it suffices to assume B is completely irreducible. In this event, it follows from theorem 1 and the fact that Bf has no 0 components that equality in (7) implies that all entries of f are the same. Thus $D = F$ and equality holds in (7) precisely when A is already doubly stochastic.

Note: A related but rather different inequality when A is symmetric appears in [3, theorem 3]. Also a portion of the proof of that result could be used to prove part of the first statement in our theorem 1.

It follows from theorem 2 that

COROLLARY 2: *If* A *is a row stochastic matrix with property* (*) *and* B *is related to* A *by* (1), *then* $|\det \mathrm{A}| \leq |\det \mathrm{B}|$.

We denote the eigenvalues of A by $\alpha_1,\cdots,\alpha_n$, ordered so that $|\alpha_1| \leq \cdots \leq |\alpha_n|$, and those of B by $\beta_1,\cdots,\beta_n$, ordered so that $|\beta_1| \leq \cdots \leq |\beta_n|$. Since $\alpha_n = 1 = \beta_n$, it follows from corollary 2 that

COROLLARY 3: *If* A *is a row stochastic matrix with property* (*) *and* B *is related to* A *by* (1), *then* $\prod_{i=1}^{n-1} |\alpha_i| \leq \prod_{i=1}^{n-1} |\beta_i|$.

We conjecture that in case A is row stochastic with property (*) and B is related to A by (1), then actually

$$|\alpha_i| \leq |\beta_i|, \, i = 1,\cdots,n.$$

The result of theorem 2 may be extended to all matrices with property (*) in the following way.

THEOREM 3: *If* A *is any nonnegative* n *by* n *matrix with property* (*) *and* D *and* E *are positive diagonal matrices guaranteed by* (1), *then* det DE $\geq \rho(\mathrm{A})^{-\mathrm{n}}$. *Furthermore, equality holds only if* DE $= \rho(\mathrm{A})^{-1}\mathrm{I}$.

PROOF: It is enough to assume A is completely irreducible (for, if not, it is equivalen to a direct sum of same) and then A is irreducible. In this event there is a positive vector x such that $Ax = \rho(A)x$ and, therefore, $\frac{1}{\rho(A)} X^{-1}AX$ is row stochastic, where $X = \mathrm{diag}\{x_1,\cdots,x_n\}$. Application of theorem 2 to $\frac{1}{\rho(A)} X^{-1}AX$ yields det $D'E' \geq 1$ where $D'(\frac{1}{\rho(A)} X^{-1}AX)E' = B$. Setting $D = \frac{1}{\rho(A)} D'X^{-1}$ and $E = XE'$, gives $B = DAE$ and det $DE \geq \rho(A)^{-n}$ as was to be shown. The case of equality also follows from theorem 2.

REMARK 3: The reader may wish to note the relationship between the present work and the notion of the *equilibrant*,

$$E(B) \equiv \inf \rho(FB)$$

(where the inf is taken over all positive diagonal matrices of determinant 1), of a nonnegative matrix mentioned in [5].

References

[1.] Brualdi, R., Parter, S., and Schneider, H., The diagonal equivalence of a nonnegative matrix to a stochastic matrix, J. Math. Anal. and Appl. **16**, 31–50 (1966).

[2.] Hardy, G., Littlewood, J., Pólya, G., Inequalities, (Cambridge University Press, 1959).

[3.] Marcus, M., Newman, M., Generalized functions of symmetric matrices, Proc. AMS **16**, 826–830, (1965).

[4.] Sinkhorn, R., A relationship between arbitrary positive matrices, and doubly stochastic matrices, Ann. Math. Stat. **35**, 876–879, (1964).

[5.] Hoffman, A., Linear G-functions, Lin. and Multilin. Alg. **3**, 45–52, (1975).

(Paper 80B4–454)