JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS 34, 648–652 (1971)

On Matrices with a Doubly Stochastic Pattern

DAVID LONDON

Technion-Israel Institute of Technology, Haifa

Submitted by Ky Fan

1. Introduction

In [7] Sinkhorn proved that if A is a positive square matrix, then there exist two diagonal matrices $D_1 = \{d_1^{(1)},..., d_n^{(1)}\}$ and $D_2 = \{d_1^{(2)},..., d_n^{(2)}\}$ with positive entries such that D_1AD_2 is doubly stochastic. This problem was studied also by Marcus and Newman [3], Maxfield and Minc [4] and Menon [5].

Later Sinkhorn and Knopp [8] considered the same problem for A nonnegative. Using a limit process of alternately normalizing the rows and columns sums of A, they obtained a necessary and sufficient condition for the existence of D_1 and D_2 such that D_1AD_2 is doubly stochastic. Brualdi, Parter and Schneider [1] obtained the same theorem by a quite different method using spectral properties of some nonlinear operators.

In this note we give a new proof of the same theorem. We introduce an extremal problem, and from the existence of a solution to this problem we derive the existence of D_1 and D_2. This method yields also a variational characterization for $\prod_{i=1}^n (d_i^{(1)} d_i^{(2)})$, which can be applied to obtain bounds for this quantity. We note that bounds for $\prod_{i=1}^n (d_i^{(1)} d_i^{(2)})$ may be of interest in connection with inequalities for the permanent of doubly stochastic matrices [3].

2. Preliminaries

Let $A = (a_{ij})$ be an $n \times n$ nonnegative matrix; that is, $a_{ij} \geqslant 0$ for $i, j = 1,..., n$. A is called *fully indecomposable* if and only if A does not contain an $s \times t$ zero submatrix such that $s + t = n$. A matrix A is called *doubly stochastic* if it is nonnegative and all its row sums and column sums are 1. A nonnegative matrix $A = (a_{ij})$ is said to *have a doubly stochastic pattern* if there exists a doubly stochastic matrix $B = (b_{ij})$ such that $a_{ij} = 0$ whenever $b_{ij} = 0$ and vice versa.

The following result is contained in [6, Theorem 1]: Let A be a nonnegative square matrix. A has a doubly stochastic pattern if and only if after independent permutations of rows and columns, A is a direct sum of fully indecomposable matrices.

3. Main Theorem

THEOREM. *Let A be an $n \times n$ nonnegative matrix. Then there exist diagonal matrices $D_1 = \{d_1^{(1)},..., d_n^{(1)}\}$ and $D_2 = \{d_1^{(2)},..., d_n^{(2)}\}$ with positive entries such that D_1AD_2 is doubly stochastic if and only if A has a doubly stochastic pattern.*

Proof. It is obvious that if positive and diagonal matrices D_1 and D_2 exist for which D_1AD_2 is doubly stochastic, then A has a doubly stochastic pattern.

Assume now that A has a doubly stochastic pattern. By the result stated above, there exist permutation matrices P and Q such that PAQ is the direct sum of fully indecomposable matrices $A_1,..., A_k$. It is obvious that it is enough to prove that for each A_j, $j = 1,..., k$, there exist positive and diagonal matrices $D_1^{(j)}$ and $D_2^{(j)}$ such that $D_1^{(j)}A_jD_2^{(j)}$ is doubly stochastic. Hence, it is enough to prove the existence of D_1 and D_2 under the assumption that A is fully indecomposable.

Let thus $A = (a_{ij})$ be fully indecomposable, let

$$F(x) = F(x_1,..., x_n) = \prod_{i=1}^{n} \left(\sum_{j=1}^{n} a_{ij}x_j \right)$$

and consider the following extremal problem[1]: $\inf F(x)$, where the infimum is taken over all $x = (x_1,..., x_n)$ satisfying $x_i > 0$, $i = 1,..., n$, and $\prod_{i=1}^n x_i = 1$. We denote the infimum by m. Obviously, $0 \leqslant m < \infty$.

We shall show that there exists a vector $x^{(0)} = (x_1^{(0)},..., x_n^{(0)})$ with positive components $x_i^{(0)} > 0$, $i = 1,..., n$, for which the infimum, hence the minimum, is obtained. If such a positive vector $x^{(0)}$ does not exist, then there exists a sequence of positive vectors $x^{(k)} = (x_1^{(k)},..., x_n^{(k)})$ such that $\prod_{i=1}^n x_i^{(k)} = 1$, $k = 1, 2,...$, $\lim_{k\to\infty} F(x^{(k)}) = m$, $\lim_{k\to\infty} x_i^{(k)} = 0$ for at least one i and (because of the condition $\prod_{i=1}^n x_i^{(k)} = 1$) $\lim_{k\to\infty} x_i^{(k)} = \infty$ for at least one i. After permuting the columns of A, if necessary, we thus obtain

$$\begin{aligned} &\lim_{k\to\infty} x_i^{(k)} = \infty, && i = 1,..., \ell,\ \ell > 1, \\ &\lim_{k\to\infty} x_i^{(k)} = 0, && i = \ell + 1,..., m,\ m > \ell, \\ &\lim_{k\to\infty} x_i^{(k)} = \alpha_i, && 0 < \alpha_i < \infty, \quad i = m + 1,..., n. \end{aligned} \tag{1}$$

[1] The author wishes to thank Dr. H. Sternin who suggested, in a private communication, the use of this extremal problem for positive matrices.

As A is fully indecomposable, it follows from the Frobenius-Konig theorem (see e.g. [2]) that it contains a positive diagonal; that is, there exists a permutation σ for which $a_{i\sigma(i)} > 0$, $i = 1,..., n$. By permuting the rows of A, we may assume that σ is the identity permutation. That is,

$$a_{ii} > 0, \qquad i = 1,..., n. \tag{2}$$

Using again the fact that A is fully indecomposable, it follows that there exists an element a_{pq} such that

$$a_{pq} > 0, \qquad \ell + 1 \leqslant p \leqslant n, \quad 1 \leqslant q \leqslant \ell. \tag{3}$$

Obviously,

$$F(x^{(k)}) \geqslant \frac{a_{pq} \prod_{i=1}^{n} a_{ii}}{a_{pp}} \frac{x_q^{(k)} \prod_{i=1}^{n} x_i^{(k)}}{x_p^{(k)}} = \frac{a_{pq} \prod_{i=1}^{n} a_{ii}}{a_{pp}} \frac{x_q^{(k)}}{x_p^{(k)}}. \tag{4}$$

From (1), (2), (3) and (4) it follows that

$$\lim_{k\to\infty} F(x^{(k)}) = \infty. \tag{5}$$

But

$$\lim_{k\to\infty} F(x^{(k)}) = m < \infty. \tag{6}$$

(5) contradicts (6), and the existence of a positive vector $x^{(0)}$ for which the minimum is obtained is established.

As $x^{(0)}$ is positive, the minimum obtained at $x^{(0)}$ is a local minimum. Hence, using Lagrange multipliers for

$$\log F = \sum_{i=1}^{n} \log \left(\sum_{j=1}^{n} a_{ij} x_j \right)$$

and the constraint $\sum_{i=1}^{n} \log x_i = 0$, we obtain:

$$\sum_{i=1}^{n} \left(\frac{a_{ij}}{\sum_{k=1}^{n} a_{ik} x_k^{(0)}} \right) = \frac{\lambda}{x_j^{(0)}}, \qquad j = 1,..., n. \tag{7}$$

Set

$$\frac{1}{\sum_{k=1}^{n} a_{ik} x_k^{(0)}} = y_i^{(0)}, \qquad i = 1,..., n, \tag{8}$$

and

$$b_{ij} = y_i^{(0)} a_{ij} x_j^{(0)}, \qquad i, j = 1,..., n. \tag{9}$$

(7), (8) and (9) imply

$$\sum_{j=1}^{n} b_{ij} = y_i^{(0)} \sum_{j=1}^{n} a_{ij} x_j^{(0)} = 1, \qquad i = 1,..., n, \tag{10}$$

and

$$\sum_{i=1}^{n} b_{ij} = x_j^{(0)} \sum_{i=1}^{n} a_{ij} y_i^{(0)} = \lambda, \qquad j = 1,..., n. \tag{11}$$

From (10) and (11) follows that $\lambda = 1$, and therefore $B = (b_{ij})$ is doubly stochastic. But $B = D_1 A D_2$, where

$$D_1 = \{d_1^{(1)},..., d_n^{(1)}\} \qquad \text{and} \qquad D_2 = \{d_1^{(2)},..., d_n^{(2)}\}$$

are the diagonal matrices given by $d_i^{(1)} = y_i^{(0)}$, $d_i^{(2)} = x_i^0$, $i = 1,..., n$. The proof of the theorem is thus completed.

We note that from our proof follows that

$$\frac{1}{\prod_{i=1}^{n} (d_i^{(1)} d_i^{(2)})} = \min_{x_j > 0} \frac{\prod_{i=1}^{n} \left(\sum_{j=1}^{n} a_{ij} x_j \right)}{\prod_{j=1}^{n} x_j}. \tag{12}$$

(12) can be applied to obtain lower bounds for the product $\prod_{i=1}^{n} (d_i^{(1)} d_i^{(2)})$.

For example, substituting in (12) $x = (1,... 1)$, we obtain

$$\frac{1}{\prod_{i=1}^{n} r_i} \leqslant \prod_{i=1}^{n} (d_i^{(1)} d_i^{(2)}),$$

where $r = (r_1,..., r_n)$ is the row sums vector of A.

If A is fully indecomposable, and if we let x in (12) be the positive characteristic vector of A corresponding to its (dominant) characteristic value α, we obtain

$$\frac{1}{\alpha^n} \leqslant \prod_{i=1}^{n} (d_i^{(1)} d_i^{(2)}).$$

From (12) follows also the following result: Let A_1 and A_2 be nonnegative matrices such that $A_2 - A_1$ is nonnegative. Denote by

$$D_1(A_i) = \{d_1^{(1)}(A_i), \ldots, d_n^{(1)}(A_i)\}$$

and

$$D_2(A_i) = \{d_1^{(2)}(A_i), \ldots, d_n^{(2)}(A_i)\}, \qquad i = 1, 2,$$

the D_1 and D_2 matrices corresponding, according to our theorem, to A_i. Then

$$\prod_{i=1}^{n} (d_i^{(1)}(A_2)\, d_i^{(2)}(A_2)) \leqslant \prod_{i=1}^{n} (d_i^{(1)}(A_1)\, d_i^{(2)}(A_1)).$$

References

1. R. Brualdi, S. Parter, and H. Schneider, The diagonal equivalence of a non-negative matrix to a stochastic matrix, *J. Math. Anal. Appl.* **16** (1966), 31–50.
2. M. Marcus and H. Minc, "A Survey of Matrix Theory and Matrix Inequalities," Allyn and Bacon, Boston, 1964.
3. M. Marcus and M. Neuman, Generalized functions of symmetric matrices, *Proc. Amer. Math. Soc.* **16** (1965), 826–830.
4. J. Maxfield and H. Minc, A doubly stochastic matrix equivalent to a given matrix, *Notices Amer. Math. Soc.* **9** (1962), 309.
5. M. V. Menon, Reduction of a matrix with positive elements to a doubly stochastic matrix, *Proc. Amer. Math. Soc.* **18** (1967), 244–247.
6. L. Mirsky and H. Perfect, The distribution of positive elements in doubly stochastic matrices, *J. London Math. Soc.* **40** (1965), 689–698.
7. R. Sinkhorn, A relationship between arbitrary positive matrices and doubly stochastic matrices, *Ann. Math. Statist.* **35** (1964), 876–879.
8. R. Sinkhorn and P. Knopp, Concerning nonnegative matrices and doubly stochastic matrices, *Pacific J. Math.* **21** (1967), 343–348.

JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS 34, 653–664 (1971)

On Roots of Normal Operators

Heydar Radjavi and Peter Rosenthal

Department of Mathematics, University of Toronto, Toronto, Canada

Submitted by Peter D. Lax

1. Introduction

If T is a bounded linear operator on Hilbert space, then T^2 can be normal without T being normal—for example, let T be any operator other than 0 such that $T^2 = 0$. Putnam [8] has given a sufficient condition that a square root of a normal operator be normal. Stampfli [10] has proven a result that shows that an nth root of an *invertible* normal operator is similar to a normal operator. Stampfli's result has been generalized in a number of papers: [7, 1, 2, 3, 4]. Stampfli implicitly found a representation of all roots of invertible normal operators—for an explicit discussion of this representation see [4].

In this paper we give a representation of all square roots of normal operators (including non-invertible operators). Our representation does not seem to be easily derivable from Stampfli's. Putnam's result, as we show below, follows immediately from our representation; Stampfli's representation does not seem to shed any light on Putnam's result. On the other hand, of course, Stampfli's representation is valid for nth roots (and even more general algebraic functions [4]) rather than just for square roots.

We also use our representation of square roots in order to obtain information about operators that have real spectrum and are nth roots of Hermitian operators.

2. Preliminaries

One of our basic tools is the Fuglede-Putnam Theorem: *if A_1 and A_2 are normal and $A_1B = BA_2$ then $A_1{}^*B = BA_2{}^*$*; (see [5] and the references given there).

We also need some other preliminary results. If $\mathscr{M}$ is a subspace of $\mathscr{H}$ and P is the orthogonal projection onto $\mathscr{M}$, then $PAP\,|_{\mathscr{M}}$ is the *compression* of A to $\mathscr{M}$.