PROBLEMS INVOLVING DIAGONAL PRODUCTS IN NONNEGATIVE MATRICES(1)

BY

RICHARD SINKHORN AND PAUL KNOPP

Introduction and definitions. A conjecture of B. L. van der Waerden states that the minimal value of the permanent of the $n \times n$ doubly stochastic matrices is $n!/n^n$ and is uniquely achieved at the matrix J_n in which every element is $1/n$. In view of this problem and the fact that the permanent of a matrix is the sum of the diagonal products, this paper investigates the question of just how well the diagonal products of a matrix characterize that matrix. The results are stated for nonnegative matrices, but many of the theorems hold in a more general setting.

The main result is the following. If A is an $n \times n$ nonnegative fully indecomposable matrix whose positive diagonal products are equal, there exists a unique matrix B of rank one which is positive and is such that $b_{ij} = a_{ij}$ when $a_{ij} > 0$. As a consequence of this it is shown that no two doubly stochastic matrices have corresponding diagonal products equal.

Two of the main tools used in obtaining these results are well known. Proofs may be found in [3, pp. 97–98].

FROBENIUS-KÖNIG THEOREM. *Every diagonal of an $n \times n$ matrix A contains a zero element if and only if A has an $s \times t$ zero submatrix with $s + t = n + 1$.*

BIRKHOFF'S THEOREM. *The set of all $n \times n$ doubly stochastic matrices forms a convex polyhedron with the permutation matrices as vertices.*

We shall make use of the following notions and definitions.

A (0, 1)-matrix is a matrix in which every element is either 0 or 1.

A diagonal of a square matrix is a collection of entries from the matrix, one from each row and one from each column. If σ is a permutation of $\{1, 2, \ldots, n\}$ then the diagonal associated with σ is $a_{1\sigma(1)}, a_{2\sigma(2)}, \ldots, a_{n\sigma(n)}$. Every diagonal corresponds to a permutation. A positive diagonal is a diagonal in which every $a_{i\sigma(i)} > 0$. A diagonal product is the product of the elements on a diagonal.

A nonnegative square matrix A has doubly stochastic pattern if there is a doubly stochastic matrix B such that $a_{ij} = 0$ if and only if $b_{ij} = 0$. A consequence of Birkhoff's theorem is that a square matrix A has doubly stochastic pattern if and only if each positive entry lies on a positive diagonal.

Received by the editors September 25, 1967.

(1) This research was partially supported by NASA grants NGR-44-005-037 and NGR-44-005-021.

A square matrix A is fully indecomposable if there do not exist permutation matrices P and Q such that PAQ has the form

$$\begin{pmatrix} B & 0 \\ C & D \end{pmatrix},$$

where B and D are square matrices. Otherwise A is partly decomposable.

A square matrix A is chainable if for each pair of nonzero entries $a_{i_1 j_1}$ and $a_{i_k j_k}$ there is a sequence of nonzero entries $a_{i_1 j_1}, \ldots, a_{i_k j_k}$ where, for $r=1, \ldots, k-1$, either $i_r = i_{r+1}$ or $j_r = j_{r+1}$. This may be described by saying that one may move from $a_{i_1 j_1}$ to $a_{i_k j_k}$ by a sequence of rook moves on the nonzero entries. The set of elements $a_{i_1 j_1}, \ldots, a_{i_k j_k}$ will be called a chain with $a_{i_1 j_1}$ and $a_{i_k j_k}$ its end points.

Let $A=(a_{ij})$ be an $m \times n$ matrix, and let u and v be positive integers such that $1 \leqq u \leqq m$, $1 \leqq v \leqq n$. Let α denote a strictly increasing sequence of u integers $(i_1, \ldots, i_u)$ chosen from $1, \ldots, m$, and let β denote a strictly increasing sequence of v integers $(j_1, \ldots, j_v)$ chosen from $1, \ldots, n$. Then $A[\alpha|\beta]$ is that submatrix of A with rows indexed by α and columns indexed by β. $A[\alpha|\beta)$ is that submatrix of A with rows indexed by α and columns indexed by the complement of β in $\{1, 2, \ldots, n\}$. $A(\alpha|\beta]$ and $A(\alpha|\beta)$ are defined analogously.

E_{ij} will denote the (0, 1)-matrix having a one only in the (i, j) position.

Results and consequences.

LEMMA 1. *A nonnegative matrix A is fully indecomposable if and only if it is chainable and has doubly stochastic pattern.*

Proof. Suppose that A has doubly stochastic pattern. Then there is a doubly stochastic matrix B with the same zero pattern as A. If A is partly decomposable so is B. In such a case there would exist permutation matrices P and Q such that

$$PBQ = \begin{pmatrix} X & 0 \\ W & Y \end{pmatrix},$$

where X and Y are square matrices. Since B is doubly stochastic so is PBQ. Since X and Y are square, it readily follows that the sum of the elements in W is zero. Hence $W=0$. But then certainly B, and therefore A, is not chainable.

Now suppose that A is fully indecomposable. Suppose further that some positive a_{ij} does not lie on a positive diagonal. By the Frobenius-König theorem $A(i|j)$ contains an $s \times (n-s)$ zero submatrix. The presence of this zero submatrix in A would make A partly decomposable. Hence A has doubly stochastic pattern.

If A is not chainable there are positive elements $a_{i_0 j_0}$ and $a_{i_1 j_1}$ which are not end points of a chain. Let $H_1 = \{i_0\}$, $K_1 = \{j \mid a_{i_0 j} > 0\}$ and define

$$H_s = \left\{ i \notin \bigcup_{k=1}^{s-1} H_k \mid a_{ij} > 0 \text{ for some } j \in K_{s-1} \right\},$$

$K_s=\{j\notin \bigcup_{k=1}^{s-1} K_k \mid a_{ij}>0$ for some $i\in H_s\}$, $s=2, 3, \ldots$. Then set $H=\bigcup_s H_s$ and $K=\bigcup_s K_s$. There exist permutation matrices P and Q such that

$$PAQ = \begin{pmatrix} A[H|K] & A[H|K) \\ A(H|K] & A(H|K) \end{pmatrix}.$$

$A[H|K]$ contains $a_{i_0j_0}$ and $A(H|K)$ contains $a_{i_1j_1}$ and thus each is nonvoid. But then $A[H|K)=0$ and $A(H|K]=0$, and it follows that A is partly decomposable.

LEMMA 2. *Suppose A is an $n\times n$ nonnegative matrix of the form*

$$A = \begin{pmatrix} A_1 & 0 & 0 & \cdots & 0 & 0 & E_1 \\ E_2 & A_2 & 0 & \cdots & 0 & 0 & 0 \\ 0 & E_3 & A_3 & \cdots & 0 & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & A_{s-2} & 0 & 0 \\ 0 & 0 & 0 & \cdots & E_{s-1} & A_{s-1} & 0 \\ 0 & 0 & 0 & \cdots & 0 & E_s & A_s \end{pmatrix},$$

where $s>1$, and for $k=1, 2, \ldots, s$, A_k is fully indecomposable and E_k has exactly one positive entry. Then (1) *A has doubly stochastic pattern*, (2) *A is chainable*, (3) *A is fully indecomposable, and* (4) *if A has the property that for each positive entry a_{ij}, $A-a_{ij}E_{ij}$ is partly decomposable, then each A_k has the same property.*

Proof. (1) Clearly any element in any A_k, $k=1, \ldots, s$, lies on a positive diagonal. Consider any E_k and the positive element in that E_k. The position of an entry (not necessarily positive) of A_k is determined by the row of the positive entry in E_k and the column of the positive entry in $E_{k+1(\mathrm{mod}\, s)}$. If the row and column of A_k containing that element in A_k are removed from A_k, the resulting submatrix of A_k will have a positive diagonal, for otherwise A_k would be partly decomposable by the Frobenius-König theorem. Let δ_k be a positive diagonal in the submatrix of A_k. The elements on all the δ_k and the positive elements of every E_k constitute a positive diagonal in A. This diagonal, of course, accounts for the positive elements in every E_k.

(2) This is obvious.

(3) This follows from (1) and (2) via Lemma 1.

(4) Suppose there is a positive element in some A_k such that its replacement by 0 transforms A_k into a fully indecomposable matrix. Then the matrix obtained by replacing that element in A by 0 satisfies the hypotheses of Lemma 2 and consequently it is fully indecomposable. This is a contradiction.

LEMMA 3. *Suppose that A is an $n\times n$ nonnegative matrix whose positive diagonal products are equal, and suppose that A has the form*

$$A=\begin{pmatrix} A_1 & 0 & 0 & \cdots & 0 & 0 & E_1 \\ E_2 & A_2 & 0 & \cdots & 0 & 0 & 0 \\ 0 & E_3 & A_3 & \cdots & 0 & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & E_{s-1} & A_{s-1} & 0 \\ 0 & 0 & 0 & \cdots & 0 & E_s & A_s \end{pmatrix},$$

where each A_k is positive and has rank one, and each E_k has exactly one positive element. Then there exists an $n\times n$ positive matrix B of rank one such that $b_{ij}=a_{ij}$ when $a_{ij}>0$. (Assume $s>1$.)

Proof. Denote the positive element in E_k by e_k. Denote by a_k the element in A_k which lies on the row of e_k and on the column of $e_{k+1(\bmod s)}$. Let d_k denote the product of the elements on any diagonal of A_k which contains a_k except for a_k. If any A_k is 1×1 use $d_k=1$.

Since the positive diagonal products of A are equal, $\prod_{k=1}^s d_k\cdot\prod_{k=1}^s a_k=\prod_{k=1}^s d_k\cdot\prod_{k=1}^s e_k$. Thus in particular, $e_1=\prod_{k=1}^s a_k/\prod_{k=2}^s e_k$.

In the $s\times s$ submatrix

$$A'=\begin{pmatrix} a_1 & 0 & 0 & \cdots & 0 & 0 & e_1 \\ e_2 & a_2 & 0 & \cdots & 0 & 0 & 0 \\ 0 & e_3 & a_3 & \cdots & 0 & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & e_{s-1} & a_{s-1} & 0 \\ 0 & 0 & 0 & \cdots & 0 & e_s & a_s \end{pmatrix}$$

replace the zero in the i,jth position by $\prod_{k=i}^j a_k/\prod_{k=i+1}^j e_k$ when $i<j$ and $(i,j)\neq(1,s)$ and by $\prod_{k=j+1}^i e_k/\prod_{k=j+1}^{i-1} a_k$ when $i>j+1$. The resulting $s\times s$ submatrix will then have rank one, for the $(i+1)$th row equals e_{i+1}/a_i times the ith row, $i=1,\ldots,s-1$.

By permuting the rows and columns of A it may be assumed that a_k is in the (1, 1) position in A_k. After the zero elements of A corresponding to the zero elements of A' have been replaced as indicated, A has the form

$$A''=\begin{pmatrix} A_1 & X_{12} & \cdots & X_{1s} \\ X_{21} & A_2 & \cdots & X_{2s} \\ \cdots & \cdots & \cdots & \cdots \\ X_{s1} & X_{s2} & \cdots & A_s \end{pmatrix},$$

where only the (1, 1) position of each X_{uv} is nonzero. Denote that element by x_{uv}.

Since each A_k has rank one the i, jth element in A_k has the form $r_{ik}c_{jk}$. The matrix B may now be obtained by replacing the zero in the i, jth position in X_{uv} by $r_{iu}c_{jv}x_{uv}/r_{1u}c_{1v}$.

THEOREM 1. *Suppose A is an $n\times n$ nonnegative fully indecomposable matrix whose positive diagonal products are equal. Then there exists a unique positive $n\times n$ matrix B of rank one such that $b_{ij}=a_{ij}$ when $a_{ij}>0$. Thus a_{ij} is of the form r_ic_j when $a_{ij}>0$. (Assume $n>1$.)*

Proof. We first show that B is unique. For any i, j, $A(i|j)$ contains a positive diagonal, for otherwise A would be partly decomposable by the Frobenius-König theorem. Since the diagonal products in any B are equal to the common value of the positive diagonal products in A, any b_{ij} replacing a zero in A would be completely determined from a positive diagonal in $A(i|j)$. Thus there can be at most one B.

The proof of the existence is by induction on the number of positive elements in A. According to [2], A must have at least $2n$ positive elements. If A has exactly $2n$ positive elements, there must be exactly two positive elements in each row and each column. By a permutation of rows, A may be assumed to have a positive main diagonal $a_1, a_2, \ldots, a_n$. If the positive element other than that on the main diagonal is denoted by some e_k, the rows and columns of A may be simultaneously permuted to put A in the form

$$A' = \begin{pmatrix} a_1 & 0 & 0 & \cdots & 0 & 0 & e_1 \\ e_2 & a_2 & 0 & \cdots & 0 & 0 & 0 \\ 0 & e_3 & a_3 & \cdots & 0 & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & e_{n-1} & a_{n-1} & 0 \\ 0 & 0 & 0 & \cdots & 0 & e_n & a_n \end{pmatrix}.$$

Otherwise A would be partly decomposable. For such an A, the B exists as was seen in the proof of Lemma 3.

Suppose such a B exists when A has m positive elements where $m\geqq 2n$, and consider an A with $m+1$ positive elements. Two cases arise. Either (1) there is a positive a_{ij} so that $A-a_{ij}E_{ij}$ is fully indecomposable or (2) for every positive a_{ij}, $A-a_{ij}E_{ij}$ is partly decomposable.

If case (1) holds, $A'=A-a_{i_0j_0}E_{i_0j_0}$ is fully indecomposable for some positive $a_{i_0j_0}$. By the induction hypothesis there is a B of rank one corresponding to A'. Since A has doubly stochastic pattern $a_{i_0j_0}$ lies on a positive diagonal $a_{1\sigma(1)}, a_{2\sigma(2)}, \ldots, a_{i_0j_0}, \ldots, a_{n\sigma(n)}$. If all the positive diagonal products in A are κ, $a_{i_0j_0}=\kappa/\prod_{i\neq i_0} a_{i\sigma(i)}$. Since A is fully indecomposable there is a positive diagonal in A which does not contain $a_{i_0j_0}$. This diagonal is in A', and thus the positive diagonal products in A' and therefore in B are equal to κ. Hence

$$b_{i_0 j_0} = \kappa \Big/ \prod_{i \neq i_0} a'_{i\sigma(i)} = \kappa \Big/ \prod_{i \neq i_0} a_{i\sigma(i)} = a_{i_0 j_0}.$$

Thus B also corresponds to A.

Suppose case (2) holds. Since A has doubly stochastic pattern, there is a doubly stochastic matrix C with the same zero pattern as A. Let c be the minimal positive element in C. By assumption $C - cE_{ij}$ is partly decomposable if i, j is such that $c_{ij} = c$. Since C itself is fully indecomposable, the rows and columns of C may be permuted to the form

$$\begin{pmatrix} C_1 & F_1 \\ C_{21} & C_2 \end{pmatrix},$$

where C_1 and C_2 are square and where exactly one element of F_1 equals c, with the remaining elements equal to zero. Since C is doubly stochastic, the row sums are one and thus the sum of the elements in C_1 is $r - c$, where r is the number of rows in C_1. But the column sums of C are also one, and thus the sum of the elements in C_{21} is c. This means that $C_{21} = F_2$ where exactly one element in F_2 equals c and all other elements are zero.

If C_1 is partly decomposable, a further permutation of rows and columns of C brings C into the form

$$\begin{pmatrix} C'_1 & 0 & F'_1 \\ C'_{21} & C'_2 & 0 \\ 0 & F'_3 & C'_3 \end{pmatrix},$$

where the C'_k are square and F'_1 and F'_3 have one element equal to c and other elements zero. Since C is doubly stochastic $C'_{21} = F'_2$ also has one element equal to c and other elements zero. Thus the form above becomes

$$\begin{pmatrix} C'_1 & 0 & F'_1 \\ F'_2 & C'_2 & 0 \\ 0 & F'_3 & C'_3 \end{pmatrix}.$$

The same form results if C_2 is partly decomposable.

If we continue inductively and recall that A and C have the same zero pattern we finally conclude that there are permutation matrices P and Q such that

$$(*) \qquad PAQ = \begin{pmatrix} A_1 & 0 & 0 & \cdots & 0 & 0 & E_1 \\ E_2 & A_2 & 0 & \cdots & 0 & 0 & 0 \\ 0 & E_3 & A_3 & \cdots & 0 & 0 & 0 \\ \multicolumn{7}{c}{\dotfill} \\ 0 & 0 & 0 & \cdots & E_{s-1} & A_{s-1} & 0 \\ 0 & 0 & 0 & \cdots & 0 & E_s & A_s \end{pmatrix},$$

where each A_k is fully indecomposable and each E_k has one positive element and all other elements equal to zero. The positive diagonal products in any particular A_k are equal.

By part (4) of Lemma 2 and the reasoning above, each A_k can be brought into the form (*). This process may be continued until matrices of the form (*) are obtained which have positive (1×1) blocks on the main diagonal. Repeated use of Lemma 3 will yield the desired positive rank one matrix B.

This completes the proof of Theorem 1.

THEOREM 2. *Suppose A is a nonnegative fully indecomposable $n \times n$ matrix such that $a_{ij} = r_i c_j$ when $a_{ij} > 0$. The nth term of the sequence of matrices obtained by alternately dividing the elements in each row by the maximal element in the row and then the elements in each column by the minimal positive element in the column is a* (0, 1)-*matrix.*

Proof. Denote the sequence of matrices by $A_1, A_2, \ldots$.

Since A is fully indecomposable there are at least two positive elements in every row and every column. Thus there are at least two rows where $a_{ij} > 0$ and c_j is maximal. Whence A_1 will have at least one column of zeros and ones, and every such column will contain at least two ones. Suppose that for one such column the ones occur at the $i_1, i_2, \ldots, i_s$ positions, where in the i_kth row of A_1, $a_{i_k j} = c_j / \max c_j$ if $a_{i_k j} > 0$. Thus the $i_1, i_2, \ldots, i_s$ rows of A_2 are zeros and ones, with at least two of the elements ones. A_2 has at least one column of zeros and ones and is such that for every one in that column, the row containing that one is a row of zeros and ones. There are at least two such rows.

We may now proceed inductively, generating at least two additional columns of zeros and ones in $A_3, A_5, A_7, \ldots$ and at least two additional rows of zeros and ones in $A_4, A_6, A_8, \ldots$ until the iteration is completed. If n is even, the iteration will be completed by the time we reach A_n, since A_n will have n rows of zeros and ones. If n is odd the iteration will still be completed at A_n since then A_n will have n columns of zeros and ones.

The examples

$$A = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} \quad \text{and} \quad A = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 2 & 3 \\ 1 & 0 & 3 \end{pmatrix}$$

show that all n steps of the iteration may be required.

COROLLARY 1. *If A is a nonnegative fully indecomposable matrix whose positive diagonal products are equal, there exist diagonal matrices D_1 and D_2 with positive main diagonals such that $D_1 A D_2$ is a* (0, 1)-*matrix.*

COROLLARY 2. *If A is a nonnegative matrix with doubly stochastic pattern whose positive diagonal products are equal, there exist diagonal matrices D_1 and D_2 with positive main diagonals such that D_1AD_2 is a* (0, 1)*-matrix.*

Proof. There is a doubly stochastic matrix with the same zero pattern as A. For convenience we suppose that it is A. If A is partly decomposable, there exist permutation matrices P and Q such that

$$PAQ = \begin{pmatrix} X & 0 \\ W & Y \end{pmatrix},$$

where X and Y are square. Since A is doubly stochastic, $W=0$. X and Y are thus each doubly stochastic, and if either is partly decomposable, it may be decomposed similarly. This process may be continued until A is decomposed into a direct sum of fully indecomposable doubly stochastic matrices. Hence there exist permutation matrices P_1 and Q_1 such that $P_1AQ_1 = A_1 \oplus A_2 \oplus \cdots \oplus A_s$, where each A_k, $k=1,\ldots,s$, is fully indecomposable. By Corollary 1, there are diagonal matrices D_{1k} and D_{2k} so that $D_{1k}A_kD_{2k}$ is a (0, 1)-matrix for $k=1,\ldots,s$. Then let $D_1 = P_1^T(D_{11} \oplus \cdots \oplus D_{1s})P_1$ and $D_2 = Q_1(D_{21} \oplus \cdots \oplus D_{2s})Q_1^T$.

The example

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

shows that the assumption of doubly stochastic pattern may not be dropped.

COROLLARY 3. *If A is a nonnegative matrix with doubly stochastic pattern whose positive diagonal products are equal, then every square submatrix of A has its positive diagonal products equal.*

COROLLARY 4. *Distinct $n \times n$ doubly stochastic matrices A and B do not have proportional corresponding diagonal products, i.e. there is no $k>0$ such that for each permutation σ, $\prod_{i=1}^n a_{i\sigma(i)} = k \prod_{i=1}^n b_{i\sigma(i)}$.*

Proof. Let A and B be $n \times n$ doubly stochastic matrices with corresponding diagonal products proportional. Note that this implies that A and B have the same zero pattern. For suppose for some i, j, $a_{ij}=0$ while $b_{ij}>0$. Since B has doubly stochastic pattern b_{ij} lies on a positive diagonal with positive product. The corresponding diagonal product in A must have a positive product. This is impossible since $a_{ij}=0$. Thus $a_{ij}=0 \Rightarrow b_{ij}=0$. Likewise, $b_{ij}=0 \Rightarrow a_{ij}=0$.

Let $C=(c_{ij})$ be defined as follows. Put $c_{ij}=0$ if $a_{ij}=b_{ij}=0$ and put $c_{ij}=a_{ij}/b_{ij}$ if $a_{ij}>0$ and $b_{ij}>0$. Then C has doubly stochastic pattern, and the positive diagonal products of C are equal. By Corollary 2, there exist diagonal matrices D_1 and D_2 with positive main diagonals such that D_1CD_2 is a (0, 1)-matrix. This means that $D_1AD_2=B$. But by a result in [1] and [4] no two doubly stochastic matrices are diagonally equivalent. Thus $A=B$.

References

1. Richard A. Brualdi, Seymour V. Parter and Hans Schneider, *The diagonal equivalence of a nonnegative matrix to a stochastic matrix*, J. Math. Anal. Appl. **16** (1966), 31–50.

2. Marvin Marcus and Henryk Minc, *Disjoint pairs of sets and incidence matrices*, Illinois J. Math. **7** (1963), 137–147.

3. ———, *A survey of matrix theory and matrix inequalities*, Allyn and Bacon, Boston, Mass., 1964.

4. Richard Sinkhorn and Paul Knopp, *Concerning nonnegative matrices and doubly stochastic matrices*, Pacific J. Math. **21** (1967), 343–348.

University of Houston,
Houston, Texas