Comm. Korean Math. Soc. 8 (1993), No. 4, pp. 569–578

MINIMUM PERMANENT ON CERTAIN FACE OF Ω_n

SUK-GEUN HWANG AND SUN-JEONG SHIN

I. Introduction

Let $X = [x_{ij}]$ be a nonnegative matrix of order n. X is called *doubly stochastic* if all of its row sums and column sums are equal to 1. The set of all $n \times n$ doubly stochastic matrices is denoted by Ω_n. It is well known that Ω_n is a convex polytope of dimension $n^2 - 2n + 1$ in the n^2-dimensional Euclidean space, of which the extreme points are the $n \times n$ permutation matrices [Birkhoff, 9].

The permanent per$X$ of an $n \times n$ matrix $X = [x_{ij}]$ is defined by

$$\mathrm{per}X = \sum_{\sigma \in S_n} \prod_{i=1}^{n} x_{i\sigma(i)},$$

where S_n stands for the symmetric group on $\{1, 2, \ldots, n\}$.

For an n-square $(0,1)$-matrix $A = [a_{ij}]$, let $\mathcal{F}(A) = \{X = [x_{ij}] \in \Omega_n | X \leq A\}$, where $X \leq A$ means that every entry of X is less than or equal to the corresponding entry of A. In order for $\mathcal{F}(A)$ to be nonempty it is necessary and sufficient that $\mathrm{per}\,A > 0$. For A with positive permanent, $\mathcal{F}(A)$ is a face of the polytope Ω_n, and every face of Ω_n is given in this fashion. As is a compact subset of a finite dimensional Euclidean space, $\mathcal{F}(A)$ contains a matrix $\overline{A}$ such that $\mathrm{per}\,\overline{A} \leq \mathrm{per}\,X$ for all $X \in \mathcal{F}(A)$. Such a matrix $\overline{A}$ is called a *minimizing matrix* on $\mathcal{F}(A)$.

One of the most famous problems in the theory of permanent was the van der Waerden conjecture appeared in 1926, which was proved in 1981 by Egoryĉev [3] and Falikman [4] independently.

Received April 13, 1993.

THEOREM (VAN DER WAERDEN-EGORYČEV-FALIKMAN). *For any* $A \in \Omega_n$,

$$\operatorname{per} A \geq \operatorname{per} J_n = \frac{n!}{n^n}$$

with equality if and only if $A = J_n$, *where* J_n *is the* $n \times n$ *matrix all of whose entries equal* $\frac{1}{n}$.

Since the affirmative resolution of the van der Waerden conjecture, there has been a lot of interest in determining the minimum permanent over various faces of Ω_n [1, 5, 6, 7, 8, 10].

Determination of the minimum permanent and permanent-minimizing matrices over an arbitrary face $\mathcal{F}(A)$ is an extremely hard problem. However solutions to this problem have been achieved for several $(0,1)$-matrices A.

For example, let

$$B_n = \begin{bmatrix} 0 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{bmatrix}.$$

Knopp and Sinkhorn [7] showed that

$$\min\{\operatorname{per} X | X \in \mathcal{F}(B_n)\} = (n-2)! \frac{(n-2)^{n-2}}{(n-1)^{2n-4}}.$$

And Brualdi [1] proved that

$$\min\{\operatorname{per} X | X \in \mathcal{F}(H_n)\} = (\frac{1}{2})^{n-1},$$

where H_n is the lower Hessenberg matrix of order n given by

$$\begin{bmatrix} 1 & 1 & 0 & \cdots & 0 & 0 \\ 1 & 1 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 1 & 1 & \cdots & 1 & 0 \\ 1 & 1 & 1 & \cdots & 1 & 1 \\ 1 & 1 & 1 & \cdots & 1 & 1 \end{bmatrix}.$$

A square matrix X is called *partly decomposable* if there exist permutation matrices P and Q such that

$$PXQ = \begin{bmatrix} X_1 & 0 \\ X_2 & X_3 \end{bmatrix}$$

where X_1 and X_3 are square matrices of order ≥ 1. If X is not partly decomposable, it is called *fully indecomposable.* Since a partly decomposable doubly stochastic matrix can always be written as a direct sum of two doubly stochastic matrices(after interchanging rows and columns, if necessary) of strictly smaller orders, it is sufficient to look at only those matrices A which are fully indecomposable when considering problems concerning the permanent on $\mathcal{F}(A)$.

The face $\mathcal{F}(A)$ is *barycentric* provided the minimum permanent on $\mathcal{F}(A)$ is achieved at its barycenter, that is, at the matrix

$$\frac{1}{\operatorname{per} A} \sum_{P \leq A} P$$

where the summation extends over all permutation matrices P with $P \leq A$.

Let A be a $(0,1)$-matrix of order n and let k be an integer with $1 \leq k \leq n$. Let B be the $(0,1)$-matrix of order $n+1$ obtained from A by appending on the right a new column equal to the k-th column of A and then appending on the bottom a new row whose only 1's are in positions k and $n+1$. We say that the matrix B is obtained from A by *copying the k-th column* . The matrix obtained from A by *copying the k-th row* is defined in a similar way. If there is a sequence of matrices $A = A_1, A_2, \ldots, A_p = C$ such that A_i can be obtained by copying a column or a row of A_{i-1} $(i = 2, \ldots, p)$, then we say that C is obtained from A by *copying rows and columns.* The Hessenberg matrix H_n can be obtained from $I_1 = [1]$ by successively copying the last row. A matrix which can be obtained from I_1 by copying rows and columns is called a *generalized Hessenberg* matrix [2]. It is easy to verify that each generalized Hessenberg matrix is fully indecomposable.

The generalized Hessenberg faces of Ω_n are the faces $\mathcal{F}(PAQ)$ where P and Q are permutation matrices and A is a generalized Hessenberg matrix of order n. In [2], Brualdi and Shader investigated the generalized Hessenberg matrices. And, for each generalized Hessenberg face $\mathcal{F}(A)$ of Ω_n, they showed that the minimum permanent equals $(\frac{1}{2})^{n-1}$.

In this paper, we determine the minimum permanent on the face $\mathcal{F}(A)$ of Ω_n for A, an $n \times n$ matrix which can be obtained from I_1 by multicopying rows and columns.

Let A be a $(0,1)$-matrix of order n and let l be an integer with $1 \leq l \leq n$. Let B be the $(0,1)$-matrix of order $n+k$ obtained from A by appending on the right k new columns equal to the l-th column of A and then appending on the bottom k new rows whose only 1's are in positions $\{(n+j, n+r), 1 \leq j, r \leq k\}$ and $\{(n+j, l), 1 \leq j \leq k\}$.

The matrix B will be called the matrix obtained from A by *k-copying the l-th column* . The matrix obtained from A by *k-copying the l-th row* is defined in a similar way. If there is a sequence of matrices $A = A_1, A_2, \ldots, A_p = C$ such that A_i can be obtained by k_{i-1}-copying a column or a row of $A_{i-1} (i = 2, \ldots, p)$, then we will say that C can be obtained from A by *multicopying rows and columns*. It can be easily proved that this matrix C is fully indecomposable. Notice that C is a generalized Hessenberg matrix if k_j=1, for all j.

EXAMPLE.

$$K_7 = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix}.$$

K_7 is a 7×7 matrix obtained from I_1 by alternately 2-copying the last column, 1-copying the last row and 3-copying the last column.

2. Some Preliminary Lemmas and Results

For a matrix A, $A(i_1, \ldots, i_s | j_1, \ldots, j_t)$ will denote the matrix obtained from A by striking out the rows numbered $i_1, \ldots, i_s$ and the columns numbered $j_1, \ldots, j_t$ and $A[i_1, \ldots, i_s | j_1, \ldots, j_t]$ is the $s \times t$ matrix whose (p, q)-entry is the same as the (i_p, j_q)-entry of A. If A is a $(0,1)$-matrix of order n, let $\mathcal{P}(A)$ denote the minimum permanent of the matrices in $\mathcal{F}(A)$, that is, $\mathcal{P}(A) = \min\{\text{per}\, X : X \in \mathcal{F}(A)\}$, and let $\mathcal{F}^{\min}(A)$ denote the set of matrices in $\mathcal{F}(A)$ with minimum permanent.

LEMMA 2.1 [FOREGGER, 5]. *Let $A = [a_{ij}]$ be an n-square fully indecomposable $(0,1)$-matrix and let $X = [x_{ij}]$ be a minimizing matrix on $\mathcal{F}(A)$. Then X is fully indecomposable, and moreover, for (i,j) such that a_{ij}=1,*

$$\begin{aligned} \text{per}\, X(i|j) &= \text{per}\, X \quad \text{if} \quad x_{ij} > 0, \\ \text{per}\, X(i|j) &\geq \text{per}\, X \quad \text{if} \quad x_{ij} = 0. \end{aligned}$$

LEMMA 2.2 [MINC, 8]. *Let $X = [x_{ij}]$ be a minimizing matrix on $\mathcal{F}(A)$, where $A = [a_1, \ldots, a_n]$ is an n-square $(0,1)$-matrix. If, for some $k \leq n$, $a_{j_1} = \cdots = a_{j_k}$, and if, for some i, $x_{ij_1} + \cdots + x_{ij_k} \neq 0$, then*

$$\text{per}\, X(i|j_t) = \text{per}\, X, \quad \text{for} \quad t = 1, \ldots, k.$$

LEMMA 2.3 [MINC, 8]. *Let $X = [x_{ij}]$ be a minimizing matrix on $\mathcal{F}(A)$, where $A = [a_1, \ldots, a_n]$ is an n-square $(0,1)$-matrix. If, for some $k \leq n$, $a_1 = \cdots = a_k$, then for any $p \leq k$, $X(J_p \bigoplus I_{n-p}) \in \mathcal{F}(A)$ and $\text{per}\, X(J_p \bigoplus I_{n-p}) = \text{per}\, X$, where $J_p = [\frac{1}{p}]_{p\times p}$ and I_{n-p} is the identity matrix of order $n-p$, i.e., the matrix obtained from X by replacing each of its first p columns by their average remains a minimizing matrix on $\mathcal{F}(A)$. A similar statement holds for rows.*

For a nonnegative matrix $X = [x_1, \ldots, x_n]$, let $\tilde{X}$ denote the n-square matrix obtained from X by replacing each of its last k columns by their average, and let X^* denote the $(n-1)$-square matrix obtained from $\tilde{X}(n|n)$ by multiplying each of the last $k-1$ columns by $\frac{k}{k-1}$, i.e., $\tilde{X} = [x_1, \ldots, x_{n-k}, s, \ldots, s]_{n\times n}$, where

$$s = \frac{1}{k}(x_{n-k+1} + \cdots + x_n).$$

and $X^* = [x_1, \ldots, x_{n-k}, \frac{k}{k-1}(s, \ldots, s)](n|n)$.

The matrix X^* will be called the (n,n)-*contraction* (or just a *contraction*, if no ambiguity arises) of X. And we define, inductively,

$$X^{[r]} = (X^{[r-1]})^*, \quad X^{[1]} = X^*, \quad r = 2, \dots, n-1.$$

For an n-square matrix X, X' will denote the matrix $X(n|n)$, and inductively, we define

$$X^{(r)} = (X^{(r-1)})', \quad X^{(1)} = X', \quad r = 2, \dots, n-1.$$

LEMMA 2.4. *Let $A = [a_{ij}]$ be a fully indecomposable $(0,1)$-matrix of order n, let $B = [b_{ij}]$ be the matrix obtained by k-copying the last column of A and let $Y \in \mathcal{F}(B)$. If the last $k+1$ columns of Y are identical, then for each $j = n, \dots, n+k$,*

$$\operatorname{per} Y(n+k|j) = \operatorname{per} Y,$$

and hence

$$\operatorname{per} Y^* = (\frac{k+1}{k})^k \operatorname{per} Y.$$

Proof.

$$\begin{aligned}\operatorname{per} Y &= \sum_{j=n}^{n+k} y_{n+k,j} \operatorname{per} Y(n+k|j) \\ &= \operatorname{per} Y(n+k|j), \quad j = n, \dots, n+k.\end{aligned}$$

And

$$\begin{aligned}\operatorname{per} Y^* &= (\frac{k+1}{k})^k \operatorname{per} Y(n+k|n+k) \\ &= (\frac{k+1}{k})^k \operatorname{per} Y.\end{aligned}$$

LEMMA 2.5. *Let $A = [a_{ij}]$ be a fully indecomposable $(0,1)$-matrix of order n and let $B = [b_{ij}]$ be the matrix obtained by k-copying the last column of A. Then*

$$\begin{aligned}\mathcal{P}(B) &= (\frac{k}{k+1})^k (\frac{k-1}{k})^{k-1} \cdots (\frac{2}{3})^2 (\frac{1}{2}) \mathcal{P}(A) \\ &= \prod_{i=1}^{k} (\frac{i}{i+1})^i \mathcal{P}(A).\end{aligned}$$

Proof. Let Z be a matrix in $\mathcal{F}^{\min}(B)$. Then Z^* is in $\mathcal{F}(B')$ and by Lemma 2.3 we may assume, without loss of generality, that the last $k+1$ columns of Z are identical. Thus by Lemma 2.4,

$$\operatorname{per} Z^* = (\frac{k+1}{k})^k \operatorname{per} Z.$$

In particular,

$$\mathcal{P}(B) \geq (\frac{k}{k+1})^k \mathcal{P}(B').$$

And inductively,

$$\mathcal{P}(B^{(r)}) \geq (\frac{k-r}{k-r+1})^{k-r} \mathcal{P}(B^{(r+1)}), \quad r = 1, \ldots, k-1.$$

Hence

$$\mathcal{P}(B) \geq (\frac{k}{k+1})^k (\frac{k-1}{k})^{k-1} \cdots (\frac{1}{2}) \mathcal{P}(A).$$

Now, let U be a matrix in $\mathcal{F}^{\min}(A)$. Let the n-th column of U be the vector u and let $u = u^1 + u^2$, where u^1 and u^2 are nonnegative vectors. Let $\hat{U}$ be a matrix obtained from U by replacing the n-th column by u^1, appending on the right a new column equal to u^2, and appending on the bottom a new row whose only nonzero entries are $\frac{1}{2}$'s in the n-th column and $(n+1)$-th column. Then $\hat{U}$ belongs to $\mathcal{F}(B^{(k-1)})$ and

$$\operatorname{per} \hat{U} = (\frac{1}{2}) \operatorname{per} U.$$

Thus

$$\mathcal{P}(B^{(k-1)}) \leq (\frac{1}{2}) \mathcal{P}(A).$$

Next, let U_1 be a matrix in $\mathcal{F}^{\min}(B^{(k-1)})$. And let the n-th column vector u of U be also represented by $u = v^1 + v^2 + v^3$, where v^1,v^2 and v^3 are nonnegative vectors. Let $\hat{U}_1$ be a matrix of order $n+2$ obtained from U by replacing the n-th column by v^1, appending two new columns $n+1$ and $n+2$ equal to v^2 and v^3 respectively, and appending on the bottom two new rows whose only nonzero entries are $\frac{1}{3}$'s in the n-th, $(n+1)$-th and $(n+2)$-th columns. Then $\hat{U}_1$ belongs to $\mathcal{F}(B^{(k-2)})$ and

$$\operatorname{per}\hat{U}_1 = (\frac{2}{3})^2 \operatorname{per} U_1.$$

Hence

$$\mathcal{P}(B^{(k-2)}) \leq (\frac{2}{3})^2 \mathcal{P}(B^{(k-1)}).$$

Continue this process until $\hat{U}_{k-1}$ belongs to $\mathcal{F}(B)$ and

$$\operatorname{per}\hat{U}_{k-1} = (\frac{k}{k+1})^k \operatorname{per} U_{k-1}.$$

Then

$$\mathcal{P}(B) \leq (\frac{k}{k+1})^k \mathcal{P}(B^{(1)}).$$

Therefore

$$\mathcal{P}(B) \leq (\frac{k}{k+1})^k (\frac{k-1}{k})^{k-1} \cdots (\frac{2}{3})^2 (\frac{1}{2}) \mathcal{P}(A)$$

and hence the Lemma is proved.

THEOREM 2.6. *Let B be an n-square matrix obtained from I_1 by k_j-copying rows and columns successively, where $j = 1, \ldots, m$, and $\sum_{j=1}^m k_j + 1 = n$. Then*

$$\mathcal{P}(B) = \prod_{j=1}^m \frac{k_j!}{(k_j+1)^{k_j}}.$$

Proof. It follows from Lemma 2.5 by induction that

$$\begin{aligned}\mathcal{P}(B) &= \prod_{j=1}^m \prod_{i=1}^{k_j} (\frac{i}{i+1})^i \\ &= \prod_{j=1}^m \frac{\prod_{i=1}^{k_j} i}{(k_j+1)^{k_j}} \\ &= \prod_{j=1}^m \frac{k_j!}{(k_j+1)^{k_j}}.\end{aligned}$$

Letting $k_j = 1$ for all $j = 1, \ldots, m$, in the equality in Theorem 2.6, we obtain the following result.

COROLLARY [BRUALDI AND SHADER, 4]. *For H_n a generalized Hessenberg matrix of order $n \geq 2$,*

$$\mathcal{P}(H_n) = (\frac{1}{2})^{n-1}.$$

THEOREM 2.7. *Let $A = [a_{ij}]$ be a fully indecomposable $(0,1)$-matrix of order n and let $B = [b_{ij}]$ be the matrix obtained by k-copying the last column of A. Then*

(1) "(a)" $Z = [z_{ij}]$ is a minimizing matrix on $\mathcal{F}(B)$ if and only if $Z^{[k]}$ is a minimizing matrix on $\mathcal{F}(A)$.

(2) "(b)" $\mathcal{F}(B)$ is barycentric if and only if $\mathcal{F}(A)$ is barycentric.

Proof. (a): $Z = [z_{ij}]$ is a minimizing matrix on $\mathcal{F}(B)$ if and only if Z^* is a minimizing matrix on $\mathcal{F}(B')$, more generally, $Z^{[r]}$ is a minimizing matrix on $\mathcal{F}(B^{(r)})(r = 1, \ldots, k)$. In fact, let $Z \in \mathcal{F}^{\min}(B)$. Then

$$Z^* = [z_1, \ldots, z_{n-k-1}, \frac{k+1}{k}s, \ldots, \frac{k+1}{k}s](n|n),$$

$$\text{where} \quad s = \frac{1}{k+1}(z_{n-k} + \cdots + z_n),$$

and

$$\operatorname{per} Z^* = (\frac{k+1}{k})^k \operatorname{per} Z.$$

Thus $Z^* \in \mathcal{F}^{\min}(B')$ because

$$\mathcal{P}(B') = (\frac{k+1}{k})^k \mathcal{P}(B)$$

as was seen in the proof of Lemma 2.5. Conversely, let $Z^* \in \mathcal{F}^{\min}(B')$. Since

$$\operatorname{per} Z^* = (\frac{k+1}{k})^k \operatorname{per} Z$$

and

$$\mathcal{P}(B') = (\frac{k+1}{k})^k \mathcal{P}(B),$$

it follows that $Z \in \mathcal{F}^{\min}(B)$. Repeating this process, we get that the second statement follows.

(b): The barycenter of $\mathcal{F}(B)$ is obtained from the barycenter X of $\mathcal{F}(A)$ by appending on the right k new columns equal to the n-th column of X, by appending on the bottom k rows whose only nonzero entries are 1's in the n-th,$(n+1)$-th,... , and $(n+k)$-th columns and then multiplying the n-th, $(n+1)$-th, ... , and $(n+k)$-th columns by $\frac{1}{k+1}$. Hence the result follows from (a).

COROLLARY. *Let B be an n-square matrix obtained from I_1 by k_j-copying rows and columns successively, $j = 1, \ldots, m$, and $\sum_{j=1}^{m} k_j + 1 = n$. Then $\mathcal{F}(B)$ is barycentric.*

References

1. R. A. Brualdi, *An interesting face of the polytope of doubly stochastic matrices*, Linear and Multilinear Algebra **17** (1985), 5–18.
2. R. A. Brualdi and B. L. Shader, *Minimum permanents on special faces of the polytope of Doubly Stochastic Matrices*, Linear Algebra Appl. to appear.
3. G. P. Egoryĉev, *The solution of the van der Waerden problem for permanents*, Dokl. Akad. Nauk SSSR. **258** (1981), 1041–1044.
4. D. I. Falikman, *A proof of van der Waerden's conjecture on the permanent of a doubly stochastic matrix*, Mat. Zametki **29** (1981), 931–938.
5. T. H. Foregger, *On the minimum value of the permanent of a nearly decomposable doubly stochastic matrix*, Linear Algebra Appl. **32** (1980), 75–85.
6. S.-G. Hwang, *Minimum permanent on faces of staircase type of the polytope of doubly stochastic matrices*, Linear and Multilinear Algebra **18** (1985), 271–306.
7. P. Knopp and R. Sinkhorn, *Minimum permanents of doubly stochastic matrices with at least one zero entry*, Linear and Multilinear Algebra **11** (1982), 351–355.
8. H. Minc, *Minimum permanents of doubly stochastic matrices with prescribed zero entries*, Linear and Multilinear Algebra **15** (1984), 225–243.
9. ———, *Permanents*, Encyclopedia Math. Appl., Addison-Wesley, Reading, Mass., 1978.
10. S.-Z. Song, *Minimum permanents on certain faces of matrices containing an identity submatrix*, Linear Algebra Appl. **108** (1988), 263–280.

Department of Mathematics
Kyungpook National University
Taegu 702-701, Korea