Kyungpook Mathematical Journal
Vol.33, No.2, 173-180, December 1993

BEHAVIOR OF THE PERMANENT FUNCTION ON SOME CLASSES OF DOUBLY STOCHASTIC MATRICES

Sun-Jeong Shin

1. Introduction

Let Ω_n denote the convex polytope consisting of all $n \times n$ doubly stochastic matrices. Let J_n denote the $n \times n$ matrix all of whose entries equals $\frac{1}{n}$. It is well known that the permanent function attains the global minimum $\frac{n!}{n^n}$ over Ω_n uniqualy at J_n. This fact was conjectured in 1926 by van der Waerden [12] and solved by Egoryĉev [3] in 1981. A strong version of the van der Waerden-Egorycev's theorem is the following property:

For $A \in \Omega_n$, the permanent function is monotone increasing on the line segment joining J_n and A . This property is referred to as the *monotonicity of permanent* (abb. MP) for A.

It is not yet known whether MP holds for all $A \in \Omega_n$. In [10] the problem of finding matrices in Ω_n for which MP holds are proposed. Up to present time several classes of doubly stochastic matrices are proved to satisfy MP [4,6,8] .

In [11], Sinkhorn proved MP for $A = \frac{1}{2}(I_n + P_n)$ where P_n denotes the $n \times n$ permutation matrix with 1's in positions $(2,1), (3,2), \ldots, (n, n-1), (1, n)$ and I_n denotes the identity matrix of order n.

In this paper, we prove MP for $\frac{1}{2r}\{(I_m + P_m) \otimes K_r\}$, $m = 3, 4$, $r \geq 1$,where K_r denotes the $r \times r$ matrix of 1's and $\otimes$ stands for the Kronecker product.

Besides the monotonicity one other interesting property of the permanent function is its convexity on a certain subclass of Ω_n. It is known [10] that the permanent function is strictly convex near J_n.

In this paper, we also prove the convexity of $\mathrm{per}[(1-t)J_n + tA]$ for $A = \frac{1}{2}(I_n + P_n), n \geq 3$, over some subinterval of $[0, 1]$.

Received May 19, 1993.

2. MP for $\frac{1}{2r}\{(I_m + P_m) \otimes K_r\}$

For a matrix A, let $A(i|j)$ denote the matrix obtained from A by deleting the row i and column j. Let D be a real valued function of Ω_n defined by

$$D(A) = \mathrm{per}A - \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} \mathrm{per}A(i|j), \quad \text{for} \quad A \in \Omega_n.$$

For $A \in \Omega_n$, and for a real number t, let $A_t = (1-t)J_n + tA$ and let $f_A(t) = \mathrm{per}A_t$, the permanent of A_t. It is proved in [4] that MP holds for $A \in \Omega_n$ if and only if $D(A_t) \geq 0$ for all t, $0 \leq t \leq 1$.

Lemma 1[Sinkhorn-Bapat, 1]. *Let $A \in \Omega_n$ be such that $\mathrm{per}A(i|j) = \mathrm{per}A$ for all $i, j = 1, \cdots, n$, then either $A = J_n$ or $A = \frac{1}{2}(I_n + P_n)$,up to permutations of rows and columns.*

Let A be an $n \times n$ real matrix. We say that the positions (i,j) and (k,l) of A are *equivalent* if there exist permutation matrices P, Q such that $PAQ = A$ and the transformation $X \to PXQ$ takes the (i,j)-entry of X onto the (k,l)-entry of PXQ for all $n \times n$ matrix X.

It is clear that, if the positions (i,j) and (k,l) of A are equivalent, then $\mathrm{per}A(i|j) = \mathrm{per}A(k|l)$.

Lemma 2. *Let $m = 3, 4, r \geq 1$ be integers such that $n = mr$. For $(I_m + P_m) \otimes K_r := [x_{ij}]$, let $Z = \{(i,j) | x_{ij} = 1\}$. Let $A = [a_{ij}]$ be an $n \times n$ matrix such that all $a_{ij}, (i,j) \in Z$, are the same and all $a_{kl}, (k,l) \notin Z$, are the same. Then all of the Z-positions of A are equivalent and all of the off Z-positions of A are equivalent.*

Proof. For $p, q = 1, \cdots, m$, let $Z_{pq} = \{(i,j) | (p-1)r+1 \leq i \leq pr, (q-1)r+1 \leq j \leq qr\}$. Then $Z = Z_{11} \cup Z_{22} \cup \cdots \cup Z_{mm} \cup Z_{21} \cup Z_{32} \cup \cdots \cup Z_{mm-1} \cup Z_{1m}$. Clearly all the positions of A in a single Z_{pq} are equivalent. To show the equivalence of all the elements of positions in $Z_{11} \cup Z_{21}$ of A, we can use the transformation $X \to PXQ$, where

$$P = \left[\begin{array}{cc|cccc} 0 & I_r & & & & \\ & & & & \mathbf{O} & \\ I_r & 0 & & & & \\ \hline & & & & & I_r \\ & & & \mathbf{O} & \cdot^{\cdot^{\cdot}} & \\ & \mathbf{O} & & & & \\ & & I_r & & & \mathbf{O} \end{array}\right],$$

$$Q = \left[\begin{array}{c|ccccc} I_r & & & & & \\ & & & \mathbf{O} & & \\ \hline & & & & & I_r \\ & & \mathbf{O} & & I_r & \\ \mathbf{O} & & & \cdot^{\cdot^{\cdot}} & & \\ & & & & & \mathbf{O} \\ & I_r & & & & \end{array}\right].$$

In addition to that, this transformation shows also that all the position of A in $Z_{12} \cup Z_{2m}$ are equivalent. And the transformation $X \to RXS$ with

$$R = \left[\begin{array}{cccccc} I_r & \cdot & & \mathbf{O} & & \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ & \cdot & & & & I_r \\ & \cdot & \mathbf{O} & & I_r & \\ \mathbf{O} & \cdot & & \cdot^{\cdot^{\cdot}} & & \\ & \cdot & & & & \mathbf{O} \\ & \cdot & I_r & & & \end{array}\right], \quad S = \left[\begin{array}{ccccc} & & & & I_r \\ & & & I_r & \\ & \mathbf{O} & \cdot^{\cdot^{\cdot}} & & \\ & & & \mathbf{O} & \\ I_r & & & & \end{array}\right],$$

shows the equivalence of all the $(Z_{11} \cup Z_{1m})$-positions, of all the $(Z_{22} \cup Z_{mm-1})$-positions and of all the $(Z_{m1} \cup Z_{2m})$-positions of A. Repeating similar arguments a finite number of times, we see that all of the Z-positions of A are equivalent and all of the off Z-positions of A are equivalent.

Corollary 2.1. *Let $m = 3, 4, r \geq 1$ be integers such that $n = mr$ and let $A = \frac{1}{2r}(I_m + P_m) \otimes K_r$. Then , for each $t \in [0,1]$, there exist real numbers, λ_t, μ_t such that*

$$\mathrm{per}A_t(i|j) - \mathrm{per}A_t = \begin{cases} \mu_t & \text{if } (i,j) \in Z, \\ \lambda_t & \text{otherwise.} \end{cases} \tag{1}$$

Expanding per A_t along the first row, we get

$$\mathrm{per}A_t = \mathrm{per}A_t + \frac{1}{n}\{nt + 2r(1-t)\}\mu_t + \frac{1}{n}\{(n-2r)(1-t)\}\lambda_t.$$

Thus we get the following relationship between λ_t and μ_t.

Corollary 2.2. *Let $m = 3, 4, r \geq 1$ be integers such that $n = mr$ and let $A = \frac{1}{2r}(I_m + P_m) \otimes K_r$. Then, for each $t \in [0,1]$, $\{nt + 2r(1-t)\}\mu_t + \{(n-2r)(1-t)\}\lambda_t = 0$, where μ_t and λ_t are the numbers defined by (1). And, if $t \neq 1$, we have*

$$\lambda_t = -\frac{nt + 2r - 2rt}{n - 2r - nt + 2rt}\mu_t. \tag{2}$$

Lemma 3. *Let $m = 3, 4, r \geq 1$ be integers such that $n = mr$ and let $A = \frac{1}{2r}(I_m + P_m) \otimes K_r$. Then, for each $t \in [0,1)$,*

$$D(A_t) = \frac{t}{1-t}\mu_t,$$

where μ_t is the number defined in (1).

Proof. From (1), we have

$$\begin{aligned} \sum_{i=1}^{n}\sum_{j=1}^{n} \mathrm{per}A_t(i|j) &= \Big(\sum_{(i,j)\in Z} + \sum_{(i,j)\notin Z}\Big)\mathrm{per}A_t(i|j) \\ &= n^2\mathrm{per}A_t + (n^2 - 2rn)\lambda_t + 2rn\mu_t. \end{aligned}$$

Therefore by Corollary 2.2 it follows that

$$\begin{aligned} D(A_t) &= \mathrm{per}A_t - \frac{1}{n^2}\sum_i\sum_j \mathrm{per}A_t(i|j) \\ &= -\frac{2r}{n}\mu_t - \frac{n-2r}{n}\lambda_t \\ &= \frac{(n-2r)t}{(n-2r)(1-t)}\mu_t \\ &= \frac{t}{1-t}\mu_t. \end{aligned}$$

Theorem 4. *Let $m = 3, 4, r \geq 1$ be integers such that $n = mr$. Then MP holds for $\frac{1}{2r}(I_m + P_m) \otimes K_r$.*

Proof. Let $A = \frac{1}{2r}(I_m + P_m) \otimes K_r$. To prove MP for A, it sufficies to show that $\mu_t \geq 0$ for all $t \in (0,1)$. Suppose that $\mu_d < 0$ for some $d \in (0,1)$. Since J_n is the unique matrix in Ω_n with the minimum permanent, MP holds for matrices in a sufficiently small neighbourhood of J_n in Ω_n. Hence we can take a $c \in (0,d)$ such that $D(A_c) > 0$ so that $\mu_c > 0$. Then, by the Intermediate Value Theorem, there exists a $t \in (c,d)$ such that $\mu_t = 0$. And hence also that $\lambda_t = 0$ by Corollary 2.2. Then $\mathrm{per}A_t(i|j) = \mathrm{per}A_t$ for all $i,j = 1,\cdots,n$. This contradicts to Lemma 1 because $A_t \neq J_n$ and $A_t \neq \frac{1}{2}(I_n + P_n)$. Thus $\mu_t \geq 0$ for all $t \in (0,1)$.

3. Convexity of the permanent for $\frac{1}{2}(I_n + P_n)$

In [2], Brualdi and Newman showed that for any $A \in \Omega_n$, $\mathrm{per}[(1-t)A + tI_n] \leq (1-t)\mathrm{per}A + t$ for all $t \in [0,1]$, and pointed out that it is not in general true that $\mathrm{per}[(1-t)J_n + tA] \leq (1-t)\mathrm{per}J_n + t\mathrm{per}A$ for all $t \in [0,1]$ and for all $A \in \Omega_n$. But, since the permanent function is strictly convex on a neighbourhood of J_n, we see that there exists some $\epsilon > 0$ such that the above inequality holds for all $A \in \Omega_n$ and for all $t \in [0,\epsilon]$.

In [7], Lih and Wang conjectured that the inequality $\mathrm{per}[(1-t)J_n + tA] \leq (1-t)\mathrm{per}J_n + t\mathrm{per}A$ holds for all $A \in \Omega_n$ and for all $t \in [0,\frac{1}{2}]$, and proved their conjecture for the case of $n = 3$.

In [5], Hwang proposed a similar conjecture asserting that the permanent function is convex on the straight line segment joining J_n and $\frac{J_n+A}{2}$ for all $A \in \Omega_n$, and proved it for the case of $n = 3$.

But the above two assertions are slightly different and so far any evidence that Lih and Wang's conjecture and Hwang's conjecture are equivalent has not been found.

In this section, we prove that the function $\mathrm{per}[(1-t)J_n + tA]$ for $A = \frac{1}{2}(I_n + P_n)$, $n \geq 3$, is convex in the interval $0 \leq t \leq \frac{1}{n-1}$.

For an $n \times n$ matrix A, let $p_k(A)$, $k = 1,\cdots,n$, be the sum of the permanents of all $\binom{n}{k}^2$ $k \times k$ submatrices of A and define $p_o(A) = 1$. Note that $p_n(A) = \mathrm{per}A$.

Let $f_{A,k}(t) = p_k((1-t)J_n + tA), \quad k = 1,\cdots,n.$

First, we shall prove that $f_{A,k}(t)$, for $A = \frac{1}{2}(I_n + P_n)$ is a convex function of t on $[0,\frac{1}{k-1}]$. For this, it suffices to show that $f''_{A,k}(t) \geq 0$ for $A = \frac{1}{2}(I_n + P_n)$ and for all $t \in [0,\frac{1}{k-1}]$.

Lemma 5. [Marcus and Minc,9] *For $A \in \Omega_n$,*

$$\begin{aligned} f_{A,k}(t) &= p_k((1-t)J_n + tA) \\ &= p_k(J_n)\sum_{i=0}^{k}\binom{k}{i}(1-t)^{k-i}t^i\frac{p_i(A)}{p_i(J_n)}, \quad k = 1,\cdots,n. \end{aligned}$$

Differentiating $f_{A,k}(t)$ twice with respect to t, we get the following.

Lemma 6. *For $A = \frac{1}{2}(I_n + P_n), n \geq 3$,*

$$\begin{aligned} f''_{A,k}(t) &= k(k-1)p_k(J_n)[\sum_{i=1}^{k-2}\{\binom{k-2}{i-1}(1-t)^{k-1-i}t^{i-1} \\ &\quad -\binom{k-2}{i}(1-t)^{k-2-i}t^i\}(\frac{p_{i+1}(A)}{p_{i+1}(J_n)} - \frac{p_i(A)}{p_i(J_n)}) \\ &\quad +t^{k-2}(\frac{p_k(A)}{p_k(J_n)} - \frac{p_{k-1}(A)}{p_{k-1}(J_n)})], \quad k = 3,\cdots,n. \end{aligned}$$

Lemma 7. [London,8] *Let $n \geq 2$. Then*

$$p_k(I_n + P_n) = \begin{cases} \frac{n}{n-k}\binom{2n-k-1}{k}, & k = 0,\cdots,n-1, \\ 2, & k = n. \end{cases}$$

Using the fact that

$$p_k(J_n) = \binom{n}{k}^2\frac{k!}{n^k}, \quad k = 0,\cdots,n,$$

and Lemma 7, we get the following theorem.

Theorem 8. *For $A = \frac{1}{2}(I_n + P_n)$, $n \geq 3$, $p_k[(1-t)J_n + tA]$, $k = 3,\cdots,n$, is a convex function of t on $[0, \frac{1}{k-1}]$.*

Proof. It sufficies to show that

$$(i) \qquad \frac{2p_i(I_n + P_n)}{p_i(J_n)} \leq \frac{p_{i+1}(I_n + P_n)}{p_{i+1}(J_n)}, \quad i = 1,\cdots,n-1.$$

and

$$(ii) \qquad \binom{k-2}{i-1}(1-t)^{k-1-i}t^{i-1} - \binom{k-2}{i}(1-t)^{k-2-i}t^i \geq 0,$$

for $t \in [0, \frac{1}{k-1}]$, $\quad k = 3, \cdots, n, \quad i = 1, \cdots, k-2.$

In fact,

(i) since

$$\frac{p_{i+1}(I_n + P_n)}{p_{i+1}(J_n)} = \frac{[(n-i-1)!]^2 n^{i+2}(2n-i-2)!}{(n!)^2(n-i-1)(2n-2i-3)!}$$

we have

$$\begin{aligned}
&\frac{p_{i+1}(I_n + P_n)}{p_{i+1}(J_n)} - \frac{2p_i(I_n + P_n)}{p_i(J_n)} \\
&= \frac{[(n-i-1)!]^2 n^{i+2}(2n-i-2)!}{(n!)^2(n-i-1)(2n-2i-3)!} - \frac{2[(n-i)!]^2 n^{i+1}(2n-i-1)!}{(n!)^2(n-i)(2n-2i-1)!} \\
&= \frac{2in^{i+1}(2n-i-2)![(n-i-1)!]^2(n-i-1)}{(n!)^2(2n-2i-1)!} \geq 0.
\end{aligned}$$

$$\begin{aligned}
(ii)\quad &\binom{k-2}{i-1}(1-t)^{k-1-i}t^{i-1} - \binom{k-2}{i}(1-t)^{k-2-i}t^i \\
&= \frac{(k-2)!}{(i-1)!(k-i-1)!}(1-t)^{k-1-i}t^{i-1} - \frac{(k-2)!}{i!(k-i-2)!}(1-t)^{k-2-i}t^i \\
&= \frac{(k-2)!}{i!(k-i-1)!}(1-t)^{k-2-i}t^{i-1}\{i-(k-1)t\} \geq 0.
\end{aligned}$$

for $t \in [0, \frac{1}{k-1}]$, $\quad k = 3, \cdots, n, \quad i = 1, \cdots, k-2$. Therefore, it follows that $f''_{A,k}(t) \geq 0$ for $t \in [0, \frac{1}{k-1}]$ from Lemma 6.

Corollary. *For $A = \frac{1}{2}(I_n + P_n)$, $n \geq 3$, per$[(1-t)J_n + tA]$ is a convex function of t on $[0, \frac{1}{n-1}]$.*

References

[1] R. Bapat, *Doubly stochastic matrices with equal subpermanents.* Lin. Alg. Applics., 51:1-8, 1983.

[2] R.A. Brualdi. and M.Newman, *Inequalities for permanents and permanental minors.* Proc. Cambridge Philos. Soc., 61:741-746, 1965.

[3] G.P.Egoryĉev, *The solution of the van der Waerden problem for permanents,* Dokl. Akad. Nauk SSSR. 258:1041-1044, 1981.

[4] S.-G. Hwang, *On the monotonicity of the permanent,* Proc. Amer. Math. Soc., 106:59-63, 1989.

[5] S.-G. Hwang, *Convexity of the permanent for Doubly Stochastic Matrices*, Lin. Multilin. Alg., 30:129-134, 1991.

[6] K.-W. Lih. and E.T.H. Wang, *Monotonicity Conjecture on permanents of doubly stochastic matrices*, Proc. Amer. Math. Soc., 82:173-178, 1981.

[7] K.-W. Lih. and E.T.H. Wang, *A convexity inequality on the permanent of doubly stochastic matrices*, Congressus Numerantium, 36:189-198, 1982.

[8] D. London, *Monotonicity of permanents of certain doubly stochastic matrices*, Pacific J. Math., 95:125-131, 1981.

[9] M. Marcus. and H. Minc, *Extensions of classical matrix inequalities*, Lin. Alg. Applics., 1:421-444, 1968.

[10] H. Minc, Permanents, Encyclopedia of Math. and Its Appl.,Addison-Wesley, Reading, Mass., 1978.

[11] R. Sinkhorn, *Concerning the question of monotonicity of the permanent on doubly stochastic matrices*, Lin. Multilin. Alg., 8:323-328, 1980.

[12] B.L. van der Waerden, Aufgabe 45, Jber. Deutsch. Math. Verein, 35:117, 1926.

DEPARTMENT OF MATHEMATICS, KYUNGPOOK NATIONAL UNIVERSITY, TAEGU 702-701, KOREA.