MONOTONICITY CONJECTURE ON PERMANENTS OF DOUBLY STOCHASTIC MATRICES

KO-WEI LIH AND EDWARD T. H. WANG

Abstract. A stronger version of the van der Waerden permanent conjecture asserts that if $J_n$ denotes the $n \times n$ matrix all of whose entries are $1/n$ and $A$ is any fixed matrix on the boundary of the set of $n \times n$ doubly stochastic matrices, then $\text{per}(\lambda A + (1 - \lambda)J_n)$ as a function of $\lambda$ is nondecreasing in the interval $[0, 1]$. In this paper, we elucidate the relation of this assertion to some other conjectures known to be stronger than van der Waerden's. We also show that this assertion is true when $n = 3$ and in the case when, up to permutations of rows and columns, either (i) $A = J_s \otimes J_t$, $0 < s, t, s + t = n$ or (ii) $A = \left[ Y Z \right]$ where $0$ is an $s \times s$ zero matrix, $Y$ is $s \times t$ with all entries equal to $1/r$, and $Z$ is $t \times t$ with all entries equal to $(t - s)/t^2$, $0 < s < t, s + t = n$.

1. Introduction and notation. A nonnegative matrix is called doubly stochastic if all its row sums and column sums equal 1. The set of all $n \times n$ doubly stochastic matrices, denoted by $\Omega_n$, forms a convex polytope with permutation matrices as vertices [2]. We denote by $I_n$ the $n \times n$ identity matrix, and by $J_n$ the $n \times n$ doubly stochastic matrix all of whose entries equal $1/n$. Let $\partial(\Omega_n)$ denote the boundary of $\Omega_n$.

The famed van der Waerden permanent conjecture [9], still unresolved, asserts that if $A \in \Omega_n$ and $A \neq J_n$, then $\text{per}(A) > \text{per}(J_n)$.

In [5] Friedland and Minc remarked that a somewhat stronger version of this conjecture asserts that the function

$$f_A(\lambda) = \text{per}(\lambda A + (1 - \lambda)J_n),$$

where $A$ is any fixed matrix on $\partial(\Omega_n)$, is strictly increasing in the interval $0 < \lambda < 1$. They showed that this assertion is true when $A$ is either a permutation matrix $P$ or $(nJ_n - P)/(n - 1)$. In [8, p. 158], Minc also proposed the problem of finding other matrices on $\partial(\Omega_n)$ for which this assertion is true.

In the present paper we study this stronger version and observe that being "strictly increasing" is actually equivalent to being "nondecreasing". We thereof pose a conjecture—the Monotonicity Conjecture—and elucidate its relation to some other conjectures known to be stronger than that of van der Waerden. We give evidence supporting this conjecture by finding two more classes of matrices $A$ on $\partial(\Omega_n)$ for which $f_A(\lambda)$ is strictly increasing in the interval $0 < \lambda < 1$.

Received by the editors February 4, 1980 and, in revised form, September 8, 1980; presented to the Society, January 9, 1981.

1980 Mathematics Subject Classification. Primary 15A15.

Key words and phrases. Doubly stochastic matrix, permanent.

Research supported by the National Research Council of Canada under grant NRC A-9121 and by the National Science Council of the Republic of China.

© 1981 American Mathematical Society
0002-9939/81/0000-0252/$02.50

License or copyright restrictions may apply to redistribution; see http://www.ams.org/journal-terms-of-use
2. The monotonicity conjecture. For \( A \in \Omega_n \), let \( M_A \) and \( M_A^* \) denote the statements that \( f_A(\lambda) \) is nondecreasing and strictly increasing in the interval \( 0 < \lambda < 1 \), respectively. Let \( M \) denote the following conjecture.

\[ M \text{(the Monotonicity Conjecture). For all } A \in \partial(\Omega_n), M_A \text{ is true.} \]

**Theorem 2.1.** If \( A \in \Omega_n \) and \( A \neq J_n \), then \( M_A \) and \( M_A^* \) are equivalent.

**Proof.** Being a polynomial in \( \lambda \), \( f_A(\lambda) \) cannot be a constant over a subinterval of \( [0, 1] \) without being a constant throughout. Therefore, if \( f_A(\lambda) = c \) over a subinterval \([\lambda_1, \lambda_2]\) of \([0, 1]\), \( \lambda_1 < \lambda_2 \), then \( f_A(\lambda) \equiv c \). However, this contradicts a result of Marcus and Newman [7] stating that the permanent function has a strict local minimum at \( J_n \).

**Theorem 2.2.** \( M \) is true for \( n = 3 \).

**Proof.** Let \( A_1 = \lambda_1 A + (1 - \lambda_1) J_3 \) and \( A_2 = \lambda_2 A + (1 - \lambda_2) J_3 \), where \( A \in \partial(\Omega_3) \) and \( 0 < \lambda_1 < \lambda_2 < 1 \). By a result of Foregger [4], \( \text{per}(J_3 + (1 - \theta) S) < \text{per}(S) \) for all \( S \in \Omega_3 \) and \( 0 < \theta < 3/2 \). In particular, \( \text{per}(\lambda S + (1 - \lambda) J_3) < \text{per}(S) \) for \( \lambda \in [0, 1] \). Choosing \( S = A_2 \) and \( \lambda = \lambda_1/\lambda_2 \) then yields

\[ \text{per}(A_2) > \text{per}(\lambda A_2 + (1 - \lambda) J_3) = \text{per}(\lambda A_2 + (1 - \lambda_2) J_3) \]
\[ = \text{per}(\lambda A + (1 - \lambda_1) J_3) = \text{per}(A_1). \]

We now derive some consequences of the Monotonicity Conjecture.

**Lemma 2.1.** Let \( A \in \Omega_n \) and \( A \neq J_n \). Then there exist \( A^0 \in \partial(\Omega_n) \) and \( \lambda_0 \in (0, 1) \) such that \( A = \lambda_0 A^0 + (1 - \lambda_0) J_n \). Furthermore, \( M_A \) holds if \( M_{A^0} \) holds.

**Proof.** The first part of the lemma is known (see the proof of Theorem 1.4 in [8, p. 78]). Since \( A \) is a convex combination of \( A^0 \) and \( J_n \), the second assertion follows.

**Corollary 2.1.** If \( M \) holds, then \( M_A \) holds for all \( A \in \Omega_n \).

**Corollary 2.2.** The following statements are all consequences of \( M \).

(i) The van der Waerden conjecture.

(ii) For all \( A \in \Omega_n \), \( f_A(1) > 0 \).

(iii) For \( A \in \Omega_n \), define \( \phi(A) = (nJ_n - A)/(n - 1) \) and in general \( \phi^{(m+1)}(A) = \phi(\phi^{(m)}(A)) \) by iteration. Then if \( A \in \Omega_n \) and \( A \neq J_n \), \( \text{per}(A) > \text{per}(\phi^{(2k)}(A)) \), \( k = 1, 2, \ldots \).

(iv) If \( A \in \Omega_n \) and \( A \neq J_n \), then \( \text{per}(A) > \text{per}((nJ_n + A)/(n + 1)) \).

**Proof.** (i) This follows from Theorem 2.1 and Corollary 2.1.

(ii) This follows from Corollary 2.1 since \( M_A \) implies that \( f_A'(\lambda) > 0 \) for \( \lambda \in (0, 1) \). But \( f_A'(\lambda) \) is a polynomial in \( \lambda \) and hence \( f_A'(1) > 0 \).

(iii) and (iv) both follow from Corollary 2.1 since it is easily checked that

\[ \phi^{(2)}(A) = \left(1/(n - 1)^2\right)A + \left(1 - (1/(n - 1)^2)\right)J_n \]

and

\[ (nJ_n + A)/(n + 1) = \left(1/(n + 1)\right)A + \left(1 - (1/(n + 1))\right)J_n. \]
We remark that (ii) is a special case of a conjecture of Djoković [3], (iii) is a weaker form of a conjecture of Marcus and Minc [6] stating that $\text{per}(A) > \text{per}(\phi(A))$, and (iv) is a conjecture due to E. T. H. Wang [10].

3. Evidence supporting $\mathcal{M}^*$. In this section we prove that $\mathcal{M}^*$ holds in the case when $A$ is permutable to either $J_s \oplus J_t$, $s, t > 0$, $s + t = n$ or $\left[ \begin{array}{cc} J_t & Y \\ Y^T & Z \end{array} \right]$ where $0$ is an $s \times s$ zero matrix, $Y$ is an $s \times t$ matrix with all entries equal to $1/t$, and $Z$ is a $t \times t$ matrix with all entries equal to $(t - s)/t^2$, $0 < s < t$, $s + t = n$. Our main tools will be a result of Baum and Sell [1] and an idea of Friedland and Minc used in [5]. Following the usual notation, for given $A = (a_{ij})$, $A(i_1, i_2, \ldots, i_k; j_1, j_2, \ldots, j_k)$ is the submatrix obtained from $A$ by deleting rows $i_1, i_2, \ldots, i_k$ and columns $j_1, j_2, \ldots, j_k$. For convenience we shall refer to $A(i; j)$ as the submatrix of the entry $a_{ij}$.

Lemma 3.1 [1]. For $A = (a_{ij}) \in \Omega_n$, let $r: \Omega_n \to \Omega_n$ be the transformation defined by $(r(A))_{ij} = a_{ij} \frac{\text{per}(A(i; j))}{\text{per}(A)}$. Then $\text{per}(A) < \text{per}(\lambda A + (1 - \lambda) r(A))$ for $\lambda \in [0, 1)$ and the equality holds if and only if $r(A) = A$.

Lemma 3.2 [5]. Let $A \in \partial(\Omega_n)$. Suppose that, for any $\lambda \in (0, 1)$, there exists $\lambda'$ such that $\lambda' < \lambda'$ and $\tau(\lambda A + (1 - \lambda) J_n) = \lambda' A + (1 - \lambda') J_n$; then $\mathcal{M}_A^*$ holds.

Assume $0 < s, t$. Let $X$ denote an $s \times s$ matrix with all entries equal to a variable $x$, $Y$ an $s \times t$ matrix with all entries equal to a variable $y$, and $Z$ a $t \times t$ matrix with all entries equal to a variable $z$. Let

$$L = L(x, y, z) = \left[ \begin{array}{cc} X & Y \\ Y^T & Z \end{array} \right].$$

Note that, for nonnegative $x$, $y$, and $z$, $L \in \Omega_{s+t}$ if and only if $sx + ty = sy + tz = 1$. We define $f_{s,t} = \text{per}(L)$, $g_{s,t} = \text{per}(L(1; 1))$, $h_{s,t} = \text{per}(L(1; s + t))$ and $f_{0,0} = g_{0,0} = h_{0,0} = 1$ for notational convenience.

Lemma 3.3. (i) $f_{s,t} = sx g_{s,t} + ty h_{s,t}$.
(ii) $g_{s,t} = (s - 1) y h_{s-1,t} + tz f_{s-1,t-1}$.
(iii) $h_{s,t} = (s - 1) x h_{s-1,t} + ty f_{s-1,t-1}$.
(iv) $h_{s-1,t} = (s - 1) y g_{s-1,t-1} + (t - 1) z h_{s-1,t-1}$.

Proof. (i) Expand $f_{s,t}$ along the first row of $L$.
(ii) Expand $g_{s,t}$ along the last column of $L(1; 1)$.
(iii) Expand $h_{s,t}$ along the first column of $L(1; s + t)$.
(iv) Expand $h_{s-1,t}$ along the last row of $L(1, 2; 1, s + t)$.

Lemma 3.4. For all $x, y, z$ such that $\text{per}(L) \neq 0$, there exist $x', y', z'$ such that $\tau(L(x, y, z)) = L(x', y', z')$.

Proof. It suffices to note that the $(s + t - 1) \times (s + t - 1)$ submatrix of an entry in the block $Y$ and that of an entry in the block $Y^T$ are transposes of each other and hence have the same permanent.

Lemma 3.5. Let $L(x, y, z) \in \Omega_{s+t}$, $0 < s, t$. Then there exists $\alpha$ such that $L(x, y, z) = \alpha (J_s \oplus J_t) + (1 - \alpha) J_{s+t}$.
Proof. Let \( n = s + t \) and \( \alpha = 1 - ny \). Since \( L(x, y, z) \in \Omega_n \), \( sx + ty = sy + tz = 1 \), we get \( s(x - y) = (1 - ty) - sy = \alpha = (1 - sy) - ty = t(z - y) \). Therefore
\[
\alpha(J_s \oplus J_t) + (1 - \alpha)J_n = L\left(\frac{\alpha}{s} + \frac{1 - \alpha}{n}, \frac{1 - \alpha}{n}, \frac{\alpha}{t} + \frac{1 - \alpha}{n}\right) = L(x, y, z).
\]

**Theorem 3.1.** If \( A = P(J_s \oplus J_t)Q \) for some permutation matrices \( P \) and \( Q \) and \( 0 < s, t \), then \( \per(\lambda A + (1 - \lambda)J_{s+t}) \) is strictly increasing in the interval \( 0 < \lambda < 1 \).

Proof. Clearly it suffices to prove the theorem for \( A = J_s \oplus J_t \). Without loss of generality, we may assume that \( s > t \). Let \( n = s + t \) and \( \lambda \in (0, 1] \). Then it is readily verified that \( \lambda(J_s \oplus J_t) + (1 - \lambda)J_{s+t} = L(x, y, z) \) where \( x = (s + t\lambda)/sn \), \( y = (1 - \lambda)/n \), and \( z = (t + s\lambda)/tn \). By Lemmas 3.4 and 3.5, there exists \( \lambda' \) such that \( \tau(\lambda(J_s \oplus J_t) + (1 - \lambda)J_n) = \lambda'(J_s \oplus J_t) + (1 - \lambda')J_n \). Our result would follow from Lemma 3.2 if \( \lambda < \lambda' \). Since \( \lambda' \) satisfies the equality
\[
\frac{s + t\lambda'}{sn} = \frac{s + t\lambda g_{s,t}}{sn} f_{s,t},
\]
it suffices to show that \( g_{s,t} > f_{s,t} \). Since \( sx + ty = 1 \), we get by Lemma 3.3 (i) that \( g_{s,t} - f_{s,t} = (1 - sx)g_{s,t} - tyh_{s,t} = ty(g_{s,t} - h_{s,t}) \). We use induction simultaneously on \( s \) and \( t \) to show that \( g_{s,t} > h_{s,t} \). We first establish the induction basis \( g_{n,1} > h_{n,1} \) for all \( s > 1 \). When \( t = 1 \), we have \( x = (s + \lambda)/sn \), \( y = (1 - \lambda)/n \), and \( z = (1 + s\lambda)/n \). Since it is readily verified that \( g_{n,1} = (s - 1)!x^{s-2}y(xz + (s - 1)y^2) \) and \( h_{n,1} = s!x^{s-2}y \), it suffices to show that \( xyz + (s - 1)y^2 > sxy \). Now straightforward computations show that \( xyz + (s - 1)y^2 - sxy = \lambda(1 + s)(1 + s\lambda)/sn^2 > 0 \), and hence the induction basis holds. Assume \( s > t > 1 \). Since \( y - x = -\lambda/s \) and \( z - y = \lambda/t \), we compute, using Lemma 3.3, that
\[
\frac{g_{s,t} - h_{s,t}}{f_{s,t}} = \frac{(s - 1)(y - x)h_{s-1,t-1} + t(z - y)f_{s-1,t-1}}{sn}\]
\[
= \frac{(\lambda/s)[sf_{s-1,t-1} - (s - 1)h_{s-1,t-1}]}{sn} = \frac{(\lambda/s)[s(s - 1)xg_{s-1,t-1} + s(t - 1)yh_{s-1,t-1} - (s - 1)^2yg_{s-1,t-1} - (s - 1)(t - 1)zh_{s-1,t-1}]}{sn} = \frac{(\lambda/s)[(s - 1)(sx - (s - 1)y)g_{s-1,t-1} + (t - 1)(sy - (s - 1)z)h_{s-1,t-1}]}{sn} = \frac{(\lambda/s)[((s - 1)g_{s-1,t-1} + (t - 1)h_{s-1,t-1}) + \lambda(n - 1)(s - 1)\left(\frac{g_{s-1,t-1} - \frac{s(t - 1)}{t(s - 1)}h_{s-1,t-1}}{f_{s-1,t-1}}\right)]}.\]

Since \( s > t \), we have \( s(t - 1)/t(s - 1) < 1 \). Furthermore \( g_{s-1,t-1} > h_{s-1,t-1} \) are positive and the induction hypothesis asserts that \( g_{s-1,t-1} > h_{s-1,t-1} \). It follows that \( g_{s,t} > h_{s,t} \) and the induction is complete.

**Theorem 3.2.** If \( A = P(L(0, 1/t, (t - s)/t^2))Q \) for some permutation matrices \( P \) and \( Q \) where \( t > s > 0 \), then \( \per(\theta A + (1 - \theta)J_{s+t}) \) is strictly increasing in the interval \( 0 < \theta < 1 \).
Proof. Clearly it suffices to prove the theorem for \( A = L(0, 1/t, (t-s)/t^2) \).
We note first that \( A = (-s/t)(J_s \oplus J_t) + (1 + (s/t))J_{s+t} \). Hence \( \theta A + (1 - \theta)J_{s+t} = (-s\theta/t)(J_s \oplus J_t) + (1 + (s\theta/t))J_{s+t} \). It remains to show that \( f_\theta(\lambda) \) is strictly decreasing in the interval \(-s/t < \lambda < 0\) where \( B = J_s \oplus J_t \). We proceed exactly as in the proof of Theorem 3.1. However, this time we need to show that \( \lambda' < \lambda \).
Accordingly, we use induction simultaneously on \( s \) and \( t \) to show that \( g_{s,t} < h_{s,t} \). To establish the induction basis \( g_{1,t} < h_{1,t} \) for all \( t > 1 \), we observe that \( y = (1-\lambda)/n \) and \( z = (t + \lambda)/tn \) when \( s = 1 \); \( g_{1,t} = t!z^t \) and \( h_{1,t} = t!yz^{t-1} \). Now \( z - y = \lambda(1+t)/tn < 0 \) since \( \lambda < 0 \). Therefore, \( g_{1,t} - h_{1,t} = t!z^{t-1}(z - y) < 0 \) and the induction basis holds. Furthermore, \( t > s \) implies that \( s(t - 1)/t(s - 1) > 1 \).
Therefore in the final expression for \( g_{s,t} - h_{s,t} \), we have
\[
g_{s-1,t-1} - (s(t - 1)/(t(s - 1)))h_{s-1,t-1} < 0
\]
by the induction hypothesis. It follows that \( g_{s,t} - h_{s,t} < 0 \).

4. Concluding remarks and an open question. For arbitrary \( A \in \partial(\Omega_n) \) it is in general difficult to verify that \( \mathcal{R}_A \) holds, if it is indeed true. One reason is that it is usually quite tedious to compute \( f_A(\lambda) \). But even when an explicit formula for \( f_A(\lambda) \) is known, it need not be easy to show that \( f_A(\lambda) \) is strictly increasing in the interval \( 0 < \lambda < 1 \). For example, a special case covered by Theorem 3.1 is the matrix \( A = J_m \oplus J_m \in \partial(\Omega_{2m}) \). For this matrix, it is not difficult to see that
\[
f_A(\lambda) = \frac{(m!)^2}{(2m)^{2m}} \sum_{k=0}^{m} \binom{m}{k}^2 (1 + \lambda)^{2m-2k}(1 - \lambda)^{2k}.
\]
Showing that \( f_A(\lambda) \) is increasing in \([0, 1]\), however, seems difficult. The \( \tau \) function introduced by Baum and Sell and the idea developed by Friedland and Mine avoid direct computation of \( f_A(\lambda) \). Nevertheless, the matrix \( A \) must be of some "special" form so that \( \tau(A) \) is of the same form, as can be seen from all the known examples. In this connection, it is natural to ask whether \( \mathcal{R}_A \) holds for all \( A = J_{n_1} \oplus J_{n_2} \oplus \cdots \oplus J_{n_k} \) where \( n_i > 0 \) such that \( \Sigma n_i = n \). Note that all the known examples are special cases of this direct sum. For small values of \( n \), the answer to this question can be easily shown to be in the affirmative. For example, if \( A = I_2 \oplus J_2 \), then \( f_A(\lambda) = (5\lambda^4 + 4\lambda^3 + 4\lambda^2 + 3)/32 \) and hence \( f_A(\lambda) = \lambda(5\lambda^2 + 3\lambda + 2)/8 \). However, we have been unable to settle this question in general.

Acknowledgement. This work was done while the second author was on sabbatical leave at the Institute of Mathematics, Academia Sinica, Republic of China. The hospitality of the Institute is greatly appreciated.

References

**Institute of Mathematics, Academia Sinica, Taipei, Taiwan, Republic of China**

**Department of Mathematics, Wilfrid Laurier University, Waterloo, Ontario, Canada**