Linear Algebra and its Applications 436 (2012) 791–801

Contents lists available at SciVerse ScienceDirect

Linear Algebra and its Applications

journal homepage: www.elsevier.com/locate/laa

Some results towards the Dittert conjecture on permanents

Gi-Sang Cheon [a], Ian M. Wanless [b,*,1]

[a] Department of Mathematics, Sungkyunkwan University, Suwon 440-746, Republic of Korea
[b] School of Mathematical Sciences, Monash University, Vic 3800, Australia

ARTICLE INFO

Article history:
Received 4 March 2010
Accepted 30 August 2010
Available online 19 September 2010

Submitted by R.A. Brualdi

AMS classification:
15A15

Keywords:
Dittert conjecture
Permanent
Doubly superstochastic

ABSTRACT

Let K_n denote the convex set consisting of all real nonnegative $n \times n$ matrices whose entries have sum n. For $A \in K_n$ with row sums $r_1, \ldots, r_n$ and column sums $c_1, \ldots, c_n$, define $\phi(A) = \prod_{i=1}^n r_i + \prod_{j=1}^n c_j - \mathrm{per}(A)$. Dittert's conjecture asserts that the maximum of ϕ on K_n occurs uniquely at $J_n = [1/n]_{n\times n}$. In this paper, we prove:

(i) if $A \in K_n$ is partly decomposable then $\phi(A) < \phi(J_n)$;
(ii) if the zeroes in $A \in K_n$ form a block then A is not a ϕ-maximising matrix;
(iii) $\phi(A) < \phi(J_n)$ unless $\delta := \mathrm{per}(J_n) - \mathrm{per}(A) \leqslant \mathcal{O}(n^4 e^{-2n})$ and

$$\left| k - \sum_{i\in\alpha} r_i \right| < \sqrt{2\delta k},$$

$$\left| k - \sum_{i\in\beta} c_i \right| < \sqrt{2\delta k} \quad \text{and} \quad \sum_{i\in\alpha, j\in\beta} a_{ij} < k + \sqrt{\delta k}$$

for all sets α, β of k integers chosen from $\{1, 2, \ldots, n\}$.

1. Introduction

A square real nonnegative matrix is called *row* (resp. *column*) *stochastic* if all its row (resp. column) sums are equal to 1. A matrix which is both row stochastic and column stochastic is called *doubly*

* Corresponding author.
E-mail addresses: gscheon@skku.edu (G.-S. Cheon), ian.wanless@monash.edu (I.M. Wanless).
1 The second author's research is supported by ARC Discovery Grant DP1093320.

0024-3795/$ - see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.laa.2010.08.041

stochastic. As usual, the set of all $n \times n$ doubly stochastic matrices is denoted by Ω_n, and the $n \times n$ matrix all of whose entries equal $\frac{1}{n}$ is denoted by J_n.

For an $n \times n$ matrix $A = [a_{ij}]$, the *permanent* of A, $\mathrm{per}(A)$, is defined by

$$\mathrm{per}(A) = \sum_{\pi} a_{1\pi(1)} \dots a_{n\pi(n)},$$

where the sum is over all permutations π of $\{1, 2, \dots, n\}$.

The problem of determining the minimum of the permanent function on Ω_n was the subject of the so-called *van der Waerden conjecture* [14]. The conjecture remained open for over half a century until Egoryĉev [5] and Falikman [6] proved it.

Theorem 1.1 (Egoryĉev-Falikman). *If $A \in \Omega_n$ then*

$$\mathrm{per}(A) \geqslant \frac{n!}{n^n},$$

with equality if and only if $A = J_n$.

The subject of this paper is a conjecture generalising Theorem 1.1, that is known as the *Dittert conjecture*. Let K_n denote the set of all real nonnegative $n \times n$ matrices whose entries have sum n, and let ϕ denote a real valued function on K_n defined by

$$\phi(A) = \prod_{i=1}^{n} \sum_{j=1}^{n} a_{ij} + \prod_{j=1}^{n} \sum_{i=1}^{n} a_{ij} - \mathrm{per}(A)$$

for $A = [a_{ij}] \in K_n$. Since K_n is a compact subset of a finite dimensional Euclidean space, it contains a matrix A such that $\phi(A) \geqslant \phi(X)$ for all $X \in K_n$. Such a matrix A will be called a *ϕ-maximising matrix* on K_n.

The following conjecture due to E. Dittert is still open for $n \geq 4$. It is Conjecture 28 in Minc's well-known survey [12] of open problems in the theory of permanents (see [3] for the most recent update on those problems).

Dittert conjecture. If $A \in K_n$ then

$$\phi(A) \leqslant 2 - \frac{n!}{n^n},$$

with equality if and only if $A = J_n$.

The Dittert Conjecture asserts that J_n is the unique ϕ-maximising matrix on K_n. If A is restricted to Ω_n, then clearly the conjecture is true from Theorem 1.1. Sinkhorn [13] has proved that a ϕ-maximising matrix on K_n has a positive permanent, and that the conjecture is true for $n = 2$. In [8,9], Hwang obtained several partial results. Specifically, he proved *inter alia* that (i) the only possible positive ϕ-maximising matrix on K_n is J_n; (ii) the only possible row stochastic ϕ-maximising matrix on K_n is J_n; (iii) every row sum and column sum of a ϕ-maximising matrix on K_n lies between $1 - \sqrt{2 \cdot n!/n^n}$ and $1 + (n-1)\sqrt{2 \cdot n!/n^n}$; (iv) the function ϕ attains a strict local maximum on K_n at J_n; (v) the conjecture is true for $n = 3$. Cheon and Hwang [1] posed a problem generalising the Dittert conjecture and Cheon and Yoon [2] obtained some sufficient conditions for which the Dittert conjecture holds. Recently, Cheon and Wanless [4] gave a graph theoretic interpretation of the Dittert conjecture.

Until now, the literature on Dittert's conjecture has not recognised a parallel investigation being undertaken in the theory of random access communication, perfect hashing and graph entropy. In particular, Hajek [7] makes a conjecture that specialises in the 2-dimensional case to exactly Dittert's conjecture. A counterexample to the 4-dimensional case of Hajek's conjecture has been given by Körner and Marton [10].

An $n \times n$ nonnegative matrix with $n \geqslant 2$ is called *partly decomposable* if there exist permutation matrices P and Q of the same order as A such that

$$PAQ = \begin{bmatrix} O & X \\ Y & Z \end{bmatrix}$$

where X and Y are square matrices of order ≥ 1 and O is a zero matrix. A matrix which is not partly decomposable is said to be *fully indecomposable*. Hwang [9] writes "One hopes to prove that if ϕ is maximal at $A \in K_n$, then $A = J_n$. It seems that one would like to show here that even if $A \neq J_n$, at the very least, A is fully indecomposable. But at the present time I am not able to prove it except for the case $n = 3$." In this paper we achieve Hwang's goal, by showing that the Dittert conjecture is true for partly decomposable matrices.

Throughout this paper, $\mathcal{O}$ will denote "big O" notation for asymptotics as $n \to \infty$. Our main results are:

(i) if $A = [a_{ij}] \in K_n$ satisfies $\phi(A) \geqslant \phi(J_n)$ and A has i-th row sum r_i and j-th column sum c_j, then $\mathrm{per}(A) \geqslant \mathrm{per}(J_n) - \mathcal{O}(n^4 e^{-2n})$ and for all $\alpha, \beta \in Q_{k,n}$

$$\left| k - \sum_{i\in\alpha} r_i \right| \leqslant \mathcal{O}\left(\sqrt{k} n^2 e^{-n}\right),$$

$$\left| k - \sum_{i\in\beta} c_i \right| \leqslant \mathcal{O}\left(\sqrt{k} n^2 e^{-n}\right), \quad \text{and}$$

$$\sum_{i\in\alpha, j\in\beta} a_{ij} < k + \mathcal{O}\left(\sqrt{k} n^2 e^{-n}\right),$$

where $Q_{k,n}$ is the set of all subsets of $\{1, 2, \ldots, n\}$ of cardinality k.

(ii) If $A \in K_n$ is partly decomposable then $\phi(A) < \phi(J_n)$.

(iii) The zeroes in a ϕ-maximising matrix do not form a single block.

Result (i) is in Section 2, while results (ii) and (iii) are in Section 3.

2. The conjecture is almost true!

Informally speaking, the aim of this section is to show that any counterexample to the Dittert conjecture is very similar to J_n. In particular its row and column sums must be very close to 1 and its permanent must be extremely close to $\mathrm{per}(J_n)$. It will follow that ϕ cannot exceed its value at J_n by much, hence the title of the section.

Our first theorem significantly extends a result of Hwang [8, Theorem 3] bounding row and column totals.

Theorem 2.1. *Let $A = [a_{ij}] \in K_n$ have row sum vector $(r_1, \ldots, r_n)$ and column sum vector $(c_1, \ldots, c_n)$. Define $\delta = \mathrm{per}(J_n) - \mathrm{per}(A)$ and suppose $1 \leqslant k \leqslant n$. If $\phi(A) \geqslant \phi(J_n)$ then $0 \leqslant \delta \leqslant n!/n^n$ and also, for all $\alpha, \beta \in Q_{k,n}$,*

$$\left| k - \sum_{i\in\alpha} r_i \right| \leqslant \sqrt{2\delta k}, \tag{1}$$

$$\left| k - \sum_{i\in\beta} c_i \right| \leqslant \sqrt{2\delta k}, \tag{2}$$

and

$$\sum_{i\in\alpha, j\in\beta} a_{ij} \leqslant k + \sqrt{\delta k}. \tag{3}$$

Equality is achieved in (1) and (2) only when $\delta = 0$.

Proof. By definition, A is non-negative so $\delta \leqslant \mathrm{per}(J_n) = n!/n^n$. On the other hand, $\delta \geqslant 0$ follows from the observation that $\mathrm{per}(A) > \mathrm{per}(J_n)$ would imply $\phi(A) \leqslant 2 - \mathrm{per}(A) < 2 - \mathrm{per}(J_n) = \phi(J_n)$.

The case $k = n$ is trivial, so henceforth we assume that $1 \leqslant k \leqslant n - 1$. The case $\delta = 0$ can also be ignored, since it implies that A must be doubly stochastic.

For any $\alpha \in Q_{k,n}$, let $\varepsilon = \varepsilon(\alpha) := k - \sum_{i\in\alpha} r_i$ and suppose that $\varepsilon^2 \geqslant 2\delta k$. By the arithmetic-geometric mean inequality, we have

$$\Psi(n, k, \varepsilon) := \left(1 - \frac{\varepsilon}{k}\right)^k \left(1 + \frac{\varepsilon}{n-k}\right)^{n-k} \geqslant \prod_{i=1}^{n} r_i.$$

Note that

$$\frac{\partial}{\partial\varepsilon}\Psi(n, k, \varepsilon) = \frac{-\varepsilon n \Psi(n, k, \varepsilon)}{(k-\varepsilon)(n-k+\varepsilon)}$$

so $\Psi(n, k, \varepsilon)$ decreases as $|\varepsilon|$ increases (provided $|\varepsilon| < \min\{k, n-k\}$). Thus we get an upper bound on $\Psi(n, k, \varepsilon)$ by assuming that $\varepsilon^2 = 2\delta k$. If we also assume that $n \geqslant 10$ then $|\varepsilon| < 9/(n^2 + 9) < 1$, and in that case

$$\begin{aligned}
\left(1 - \frac{\varepsilon}{k}\right)^k &< 1 - \varepsilon + \frac{k-1}{2k}\varepsilon^2 + \sum_{i=3}^{k} \frac{1}{k^i}\binom{k}{i}|\varepsilon|^i \\
&< 1 - \varepsilon + \frac{k-1}{2k}\varepsilon^2 + \frac{1}{6}\sum_{i=3}^{k}|\varepsilon|^i \\
&< 1 - \varepsilon + \frac{1}{2}\left(1 - \frac{1}{k} + \frac{|\varepsilon|}{3(1-|\varepsilon|)}\right)\varepsilon^2 \\
&< 1 - \varepsilon + \frac{1}{2}\left(1 - \frac{1}{k} + \frac{3}{n^2}\right)\varepsilon^2.
\end{aligned}$$

With a similar bound for $(1 + \varepsilon/(n-k))^{n-k}$ we find that

$$\begin{aligned}
\Psi(n, k, \varepsilon) &< \left(1 - \varepsilon + \frac{1}{2}\left(1 - \frac{1}{k} + \frac{3}{n^2}\right)\varepsilon^2\right)\left(1 + \varepsilon + \frac{1}{2}\left(1 - \frac{1}{n-k} + \frac{3}{n^2}\right)\varepsilon^2\right) \\
&< 1 - \frac{1}{2}\left(\frac{1}{k} + \frac{1}{n-k} - \frac{6}{n^2}\right)\varepsilon^2 + \frac{1}{2}|\varepsilon|^3 + \frac{1}{4}\varepsilon^4 \\
&< 1 - \frac{1}{2k}\varepsilon^2 \qquad (4) \\
&\leqslant 1 - \delta.
\end{aligned}$$

Having proved (4) for all $n \geqslant 10$, it is a simple matter to check by solving $\Psi(n, k, \varepsilon) = 1 - \varepsilon^2/(2k)$ numerically for each pair of integers (n, k) in the range $1 < k < n < 10$ that (4) also holds in each of these cases, and hence is true without restriction. Thus we get

$$\phi(A) \leqslant \prod_{i=1}^{n} r_i + 1 - \mathrm{per}(A) < 2 - \delta - \mathrm{per}(A) = 2 - \mathrm{per}(J_n) = \phi(J_n),$$

which proves (1). The result (2) for columns follows by transposition.

It remains to show (3). Let $\varepsilon = \varepsilon(\alpha, \beta) := k - \sum_{i\in\alpha,\, j\in\beta} a_{ij}$. If $\varepsilon \geqslant 0$ then (3) is immediate, so we assume that $\varepsilon < 0$. In that case we can use

$$\sum_{i\in\alpha,\, j\in\beta} a_{ij} \leqslant \sum_{i\in\alpha} r_i \quad \text{and} \quad \sum_{i\in\alpha,\, j\in\beta} a_{ij} \leqslant \sum_{j\in\beta} c_j,$$

to deduce that

$$\prod_{i=1}^{n} r_i + \prod_{j=1}^{n} c_j \leqslant 2\left(1-\frac{\varepsilon}{k}\right)^k \left(1+\frac{\varepsilon}{n-k}\right)^{n-k} = 2\Psi(n,k,\varepsilon).$$

Then the required $\varepsilon > -\sqrt{\delta k}$ can be proved using (4). □

A square nonnegative matrix $S = [s_{ij}]$ is said to be *doubly superstochastic* if there exists a doubly stochastic matrix $D = [d_{ij}]$ such that $s_{ij} \geqslant d_{ij}$ for all i, j. Since $S = D + T$ for some nonnegative matrix T, if S is an $n \times n$ doubly superstochastic matrix then $\mathrm{per}(S) \geqslant \mathrm{per}(J_n)$ from Theorem 1.1. A characterisation of doubly superstochastic matrices was obtained by Li [11].

Lemma 2.2. *An $n \times n$ nonnegative matrix $S = [s_{ij}]$ is doubly superstochastic if and only if for all $I, J \subset \{1, 2, \ldots, n\}$,*

$$\sum_{i\in I,\, j\in J} s_{ij} \geqslant |I| + |J| - n.$$

Henceforth, let $s(M)$ denote the sum of the entries in a matrix M. A corollary is:

Lemma 2.3. *Suppose M is any non-negative $m \times m$ matrix with $\mathrm{per}(M) > 0$. Then there exists a finite $\mu > 0$ and a submatrix N of M of dimensions say $u \times v$, such that μM is doubly superstochastic, and $s(N) = (u + v - m)/\mu > 0$.*

Proof. There are finitely many $u \times v$ submatrices of M satisfying $u + v > m$. Each such submatrix S must contain at least one positive entry since $\mathrm{per}(M) > 0$ by assumption, so there is some scalar $\lambda_S > 0$ for which $s(\lambda_S S) = u + v - m$. By Lemma 2.2, we take $\mu = \lambda_N$ where N is the choice of S that maximises λ_S. □

Next we show that the permanent of a ϕ-maximising matrix is asymptotic to $\mathrm{per}(J_n)$ as $n \to \infty$. In fact:

Theorem 2.4. *Suppose $A \in K_n$ satisfies $\phi(A) \geqslant \phi(J_n)$. Then*

$$\left(1-\sqrt{8\delta n^3}\right)\mathrm{per}(J_n) < \mathrm{per}(A) \leqslant \mathrm{per}(J_n), \tag{5}$$

where $\delta = \mathrm{per}(J_n) - \mathrm{per}(A) < 8n^3(\mathrm{per}(J_n))^2 = \mathcal{O}(n^4 e^{-2n})$.

Proof. The upper bound on $\mathrm{per}(A)$ is trivial given that $\phi(A) \leqslant 2 - \mathrm{per}(A)$ and $\phi(J_n) = 2 - \mathrm{per}(J_n)$. Let $\delta = \mathrm{per}(J_n) - \mathrm{per}(A)$. To prove the lower bound it suffices to show that $(1-\sqrt{8\delta n})^{-1}A$ must be doubly superstochastic, since it will then follow that

$$\mathrm{per}(J_n) \leqslant \mathrm{per}\left(\left(1-\sqrt{8\delta n}\right)^{-1} A\right) = \left(1-\sqrt{8\delta n}\right)^{-n} \mathrm{per}(A).$$

This proves (5) since $\left(1-\sqrt{8\delta n}\right)^n > 1 - n\sqrt{8\delta n}$. Note that we may assume that n is large enough that $8\delta n^3 < 1$ since otherwise the result is trivial.

So suppose, by way of contradiction, that $\left(1-\sqrt{8\delta n}\right)^{-1} A$ is not doubly superstochastic. By Lemma 2.2 there must be a $u \times v$ submatrix X of A such that $s(X) < \left(1-\sqrt{8\delta n}\right)(u+v-n)$. Let Y denote the (possibly vacuous) $u \times (n-v)$ submatrix of A which shares the same rows as X but is disjoint from X. Then by Theorem 2.1 we have $s(Y) \leqslant n - v + \sqrt{2\delta(n-v)}$ and $s(X) + s(Y) \geqslant u - \sqrt{2\delta u}$. But this implies the contradiction that

$$s(X) \geqslant u - \sqrt{2\delta u} - n + v - \sqrt{2\delta(n-v)} > u + v - n - 2\sqrt{2\delta n} \geqslant \left(1-\sqrt{8\delta n}\right)(u+v-n)$$

since $u + v - n \geqslant 1$.

Having shown (5), we find that $\delta = \mathrm{per}(J_n) - \mathrm{per}(A) < \sqrt{8\delta n^3}\mathrm{per}(J_n)$ from which it follows that $\delta < 8n^3(\mathrm{per}(J_n))^2 = \mathcal{O}(n^4e^{-2n})$ as required. □

Theorem 2.4 shows that if the Dittert conjecture fails then it only barely does, in the sense that $\phi(A) < \phi(J_n) + \mathcal{O}(n^4e^{-2n})$ for all $A \in K_n$. Also note that Theorem 2.4 implies that the bounds in (1) and (2) are $\mathcal{O}\left(\sqrt{k}n^2e^{-n}\right)$, meaning that any ϕ-maximising matrix has to be very close to doubly stochastic.

3. Zero patterns

We say that two matrices A and B with the same dimensions have the same *zero pattern* if an entry in A is zero if and only if the corresponding entry in B is zero. The basic aim of this section is to prove restrictions on the zero pattern of any possible counterexample to the Dittert conjecture.

For our first result we will need a technical lemma.

Lemma 3.1. *For a fixed integer t, let s be an integer satisfying $1 \leqslant s < \frac{t}{2}$ and define $f(s) = \frac{s!(t-s)!}{s^s(t-s)^{t-s}}$. Then $f(s) \leqslant f(s+1)$.*

Proof. Treating f as a function of a real variable, we find

$$\begin{aligned}\frac{d}{ds}\log(f(s)) &= (\log(t-s) - \log(s)) - \sum_{i=s+1}^{t-s}\frac{1}{i}\\ &= \int_s^{t-s}\frac{1}{x}\,dx - \sum_{i=s+1}^{t-s}\frac{1}{i},\end{aligned}$$

which is positive because the sum is a lower Riemann sum approximating the integral. Hence f is an increasing function in the required range. □

Theorem 3.2. *If $A \in K_n$ is partly decomposable then $\phi(A) < \phi(J_n)$.*

Proof. Let $A \in K_n$ be partly decomposable. Then there exist permutation matrices P and Q such that

$$PAQ = \begin{bmatrix} O & X \\ Y & Z \end{bmatrix} \tag{6}$$

where O is a zero submatrix and X and Y are square matrices of order w and $n-w$ respectively, where $1 \leqslant w \leqslant n-1$. Let

$$\lambda = \left(1 - \frac{1}{n}\right)^{\frac{1}{n}-1} = 1 + \frac{1}{n} - O(n^{-3}).$$

(*Case* 1) Suppose that both λX and λY are doubly superstochastic matrices. Then

$$\frac{w!}{w^w} \leqslant \mathrm{per}(\lambda X) = \lambda^w\mathrm{per}(X),$$

and similarly we have

$$\frac{(n-w)!}{(n-w)^{n-w}} \leqslant \lambda^{n-w}\mathrm{per}(Y).$$

Applying Lemma 3.1 gives

$$\begin{aligned}\mathrm{per}(A) &= \mathrm{per}(X)\mathrm{per}(Y)\\ &\geqslant \lambda^{-n}\frac{w!(n-w)!}{w^w(n-w)^{n-w}} = \left(1 - \frac{1}{n}\right)^{n-1}\frac{w!(n-w)!}{w^w(n-w)^{n-w}}\end{aligned}$$

$$\geqslant \left(1-\frac{1}{n}\right)^{n-1}\frac{(n-1)!}{(n-1)^{n-1}}$$
$$= \frac{n!}{n^n} = \mathrm{per}(J_n).$$

Thus we have

$$\phi(A) \leqslant 2 - \mathrm{per}(A) \leqslant 2 - \mathrm{per}(J_n) = \phi(J_n). \tag{7}$$

Moreover, for equality to hold in the first inequality in (7), A must be doubly stochastic. However, in that case the second inequality in (7) is strict, by Theorem 1.1, since $A \neq J_n$ because J_n is fully indecomposable. So we have proved the strict inequality required.

(*Case* 2) Without loss of generality, we may assume that λX is not doubly superstochastic and hence, by Lemma 2.2, that X has the form

$$X = \begin{bmatrix} X_1 & X_2 \\ X_3 & X_4 \end{bmatrix} \tag{8}$$

where X_1 is a $u \times v$ matrix and $\lambda s(X_1) < u + v - w$. Applying Theorem 2.1 shows that

$$s(X_1) + s(X_2) > u - \sqrt{2\delta u} > u - \sqrt{2\delta n}, \quad \text{and}$$
$$s(X_2) < w - v + \sqrt{2\delta(w-v)} < w - v + \sqrt{2\delta n}.$$

Hence $\frac{1}{\lambda}(u+v-w) > s(X_1) > u+v-w-2\sqrt{2\delta n}$. Since $u+v-w \geqslant 1$ and $\lambda > 1$ this means that $2\sqrt{2\delta n} > \left(1-\frac{1}{\lambda}\right)(u+v-w) > \left(1-\frac{1}{\lambda}\right)$, which is a contradiction for all $n \geqslant 12$, given that $\delta \leqslant n!/n^n$.

Below we describe a computation using interval arithmetic that we used to dispose of the cases $4 \leqslant n \leqslant 11$. This completes the proof of the present theorem, since Hwang [9] showed that the Dittert conjecture holds for $n \leqslant 3$. □

The basic idea of our computation to prove Theorem 3.2 for $4 \leqslant n \leqslant 11$ was to assume that there is a partly decomposable matrix $A \in K_n$ satisfying $\phi(A) \geqslant \phi(J_n)$. For each of a number of variables the computer maintained an interval in which the variable was assumed to lie. Successive iterations then used various inequalities to narrow these intervals until such time as one of the intervals became empty, at which point we had arrived at a contradiction to our assumption of the existence of A.

We could assume without loss of generality that A had the structure given by (6) where X is a $w \times w$ submatrix for some $w \leqslant \frac{1}{2}n$ and $\mathrm{per}(A) = \mathrm{per}(X)\mathrm{per}(Y) > 0$.

From Lemma 2.3 we know there is $\mu > 0$ such that μX is doubly superstochastic and X has the form (8) where X_1 is a $u_1 \times v_1$ submatrix of X such that $u_1 + v_1 > w$ and $s(X_1) = (u_1 + v_1 - w)/\mu$. Similarly, we may assume that νY is doubly superstochastic and Y has the form

$$Y = \begin{bmatrix} Y_1 & Y_2 \\ Y_3 & Y_4 \end{bmatrix} \tag{9}$$

where Y_1 is a $u_2 \times v_2$ submatrix of Y such that $u_2 + v_2 > n - w$ and $s(Y_1) = (u_2 + v_2 - n + w)/\nu$. The block structures (8) and (9) induce a block structure on Z so that we have

$$PAQ = \begin{bmatrix} O & O & X_1 & X_2 \\ O & O & X_3 & X_4 \\ Y_1 & Y_2 & Z_1 & Z_2 \\ Y_3 & Y_4 & Z_3 & Z_4 \end{bmatrix}. \tag{10}$$

Our program considered, in turn, the possibility of A having the structure (10) for each possible choice of positive integers n, w, u_1, v_1, u_2, v_2 satisfying $u_1 \leqslant w < u_1 + v_1, u_2 \leqslant n - w < u_2 + v_2, v_1 \leqslant w \leqslant \frac{1}{2}n$, $v_2 \leqslant n - w$ and $4 \leqslant n \leqslant 11$. In some cases, such as when $u_1 = w$ or $v_2 = n - w$, some of the blocks in (10) were vacuous, in which case the following description should be adjusted accordingly. For the

sake of simplicity, though, we describe the case when $u_1 < w, v_1 < w, u_2 < n - w$ and $v_2 < n - w$ so that no block is vacuous.

Our program maintained intervals for the following variables:

- $R = \prod_i r_i, C = \prod_i c_i, P = \mathrm{per}(A)$.
- $s(X_i), s(Y_i), s(Z_i)$ for $i = 1, 2, 3, 4$.
- $\rho_1 = \sum_{i=1}^{u_1} r_i, \rho_2 = \sum_{i=u_1+1}^{w} r_i, \rho_3 = \sum_{i=w+1}^{w+u_2} r_i, \rho_4 = \sum_{i=w+u_2+1}^{n} r_i$.
- $\gamma_1 = \sum_{i=1}^{v_2} c_i, \gamma_2 = \sum_{i=v_2+1}^{n-w} c_i, \gamma_3 = \sum_{i=n-w+1}^{n-w+v_1} c_i, \gamma_4 = \sum_{i=n-w+v_1+1}^{n} c_i$.

The main procedure took intervals for $s(X_1)$ and $s(Y_1)$ as its parameters, say $s(X_1) \in [\alpha_X, \beta_X]$ and $s(Y_1) \in [\alpha_Y, \beta_Y]$. If it was unsuccessful in deriving a contradiction it would call itself four more times with respective parameters

$$(a) \qquad s(X_1) \in \left[\alpha_X, \frac{1}{2}(\alpha_X + \beta_X)\right], \quad s(Y_1) \in \left[\alpha_Y, \frac{1}{2}(\alpha_Y + \beta_Y)\right],$$

$$(b) \qquad s(X_1) \in \left[\frac{1}{2}(\alpha_X + \beta_X), \beta_X\right], \quad s(Y_1) \in \left[\alpha_Y, \frac{1}{2}(\alpha_Y + \beta_Y)\right],$$

$$(c) \qquad s(X_1) \in \left[\alpha_X, \frac{1}{2}(\alpha_X + \beta_X)\right], \quad s(Y_1) \in \left[\frac{1}{2}(\alpha_Y + \beta_Y), \beta_Y\right],$$

$$(d) \qquad s(X_1) \in \left[\frac{1}{2}(\alpha_X + \beta_X), \beta_X\right], \quad s(Y_1) \in \left[\frac{1}{2}(\alpha_Y + \beta_Y), \beta_Y\right].$$

This recursion was continued until the intervals were small enough that a contradiction could be derived. The starting point was to specify $s(X_1) \in [0, n]$ and $s(Y_1) \in [0, n]$.

Inside the main procedure, the intervals for variables were updated on the basis of the following simple information.

$$\begin{aligned}
&R + C - P = \phi(A) \geqslant \phi(J_n) = 2 - n!/n^n \\
&\rho_1 = s(X_1) + s(X_2) \qquad (11)\\
&\rho_2 = s(X_3) + s(X_4) \\
&\rho_3 = s(Y_1) + s(Y_2) + s(Z_1) + s(Z_2) \\
&\rho_4 = s(Y_3) + s(Y_4) + s(Z_3) + s(Z_4) \\
&\gamma_1 = s(Y_1) + s(Y_3) \\
&\gamma_2 = s(Y_2) + s(Y_4) \\
&\gamma_3 = s(X_1) + s(X_3) + s(Z_1) + s(Z_3) \\
&\gamma_4 = s(X_2) + s(X_4) + s(Z_2) + s(Z_4) \\
&\rho_1 + \rho_2 + \rho_3 + \rho_4 = n = \gamma_1 + \gamma_2 + \gamma_3 + \gamma_4. \qquad (12)
\end{aligned}$$

For example, any lower bound for ρ_1 and upper bound for $s(X_1)$ provides a lower bound for $s(X_2)$, by (11).

Crucially, we also used three more advanced deductions. First of all we knew that $\mu \leqslant (u_1 + v_1 - w)/\alpha_X$ and $\nu \leqslant (u_2 + v_2 - n + w)/\alpha_Y$ and thus

$$P = \mathrm{per}(A) \geqslant w!\left(\frac{\alpha_X}{(u_1 + v_1 - w)w}\right)^w (n - w)!\left(\frac{\alpha_Y}{(u_2 + v_2 - n + w)(n - w)}\right)^{n-w}.$$

Secondly, we used Theorem 2.1 to bound the ρ_i and γ_i. Thirdly, we found an upper bound for R using the following consequence of the fact that if $x + y$ is fixed then xy is maximised by choosing x and y as equal as possible.

Lemma 3.3. *Suppose that $\rho_i \in [\xi_i, \xi_i']$ for each i. Define*

$$a_1 = \rho_1/u_1, \quad a_2 = \rho_2/(w - u_1), \quad a_3 = \rho_3/u_2, \quad a_4 = \rho_4/(n - w - u_2).$$

Then there is value x such that R is maximised at a point where, for each i, one of these statements holds:

- $a_i = x$,
- $a_i < x$ *and* $\rho_i = \xi_i'$ *or*
- $a_i > x$ *and* $\rho_i = \xi_i$.

Moreover, $R \leqslant a_1^{u_1} a_2^{w-u_1} a_3^{u_2} a_4^{n-w-u_2}$.

Given (12), it is easy to find the value of x in Lemma 3.3 and thereby an upper bound on R. Of course, a corresponding bound holds for C, based on the γ_i.

All of the above information was used to update the interval for each variable. This process was then iterated 50 times, since narrowing the interval for one variable can allow stronger deductions for other variables. If a contradiction was discovered at any point the procedure call was successful, otherwise it would call itself recursively as described above. In this manner, the theorem was proved for $5 \leqslant n \leqslant 11$.

For $n = 4$ the above technique had to be refined by assigning an interval for every cell of the matrix, and splitting one of these whenever a contradiction could not be found. This required a separate program to be written, though the principle was the same. In this way the proof of Theorem 3.2 was completed.

For any matrix A with row sum vector $(r_1, \ldots, r_n)$ and column sum vector $(c_1, \ldots, c_n)$ we define

$$\phi_{ij}(A) = \prod_{k \neq i} r_k + \prod_{l \neq j} c_l - \operatorname{per}(A(i|j)).$$

Here $A(i|j)$ denotes the $(n-1) \times (n-1)$ matrix obtained from A by deleting the ith row and the jth column.

Our next theorem needs three results due to Hwang (the first two are from [8] and the third was proved in [9]).

Lemma 3.4. *Let Z be the set of all matrices in K_n which have a particular zero pattern. Then there exists $Z_0 \in Z$ such that*

(*i*) $\phi(Z_0) \geqslant \phi(Z_1)$ *for all* $Z_1 \in Z$;
(*ii*) *Any two rows in Z_0 which have the same zero pattern are equal*;
(*iii*) *Any two columns in Z_0 which have the same zero pattern are equal*.

Lemma 3.5. *Let A be a ϕ-maximising matrix on K_n. If A is row-stochastic then $A = J_n$.*

Lemma 3.6. *Let A be a ϕ-maximising matrix on K_n. Then $\phi(A) \geqslant \phi_{ij}(A)$, with equality holding if $a_{ij} > 0$.*

With the aid of these lemmas, we can show:

Theorem 3.7. *If the zeros of $A \in K_n$ form a single block up to permutations, such as*

$$A = \begin{bmatrix} O & X \\ Y & Z \end{bmatrix}$$

where the zero block O is $p \times q$ and $X, Y, Z > 0$, then A is not a ϕ-maximising matrix.

Proof. Without loss of generality we assume that $p \geqslant q$ and $n \geqslant 4$. By Theorem 3.2 we can assume that $p + q \leqslant n - 1$. Also by Lemma 3.4 we may assume that $X = xJ_{p,n-q}$, $Y = yJ_{n-p,q}$, and $Z = zJ_{n-p,n-q}$ where $J_{s,t}$ is the $s \times t$ matrix all of whose entries are 1.

Let $R = \prod_i r_i$, $C = \prod_j c_j$ and $P = \mathrm{per}(A)$ and note that without loss of generality we may assume that

$$1 - \frac{n!}{n^n} < R - P < R < 1, \quad 1 - \frac{n!}{n^n} < C - P < C < 1, \quad 0 < P < \frac{n!}{n^n}. \tag{13}$$

Also let $a = (n-q)x$ so that $r_i = a$ for $i \leqslant p$ and $r_i = (n-pa)/(n-p)$ for $i > p$. Similarly, let $b = (n-p)y$. From straightforward calculations it then follows that

$$\phi_{1n}(A) = \frac{1}{a}R + \frac{n-q}{n-qb}C - \frac{1}{a}P,$$

$$\phi_{n1}(A) = \frac{n-p}{n-pa}R + \frac{1}{b}C - \frac{1}{b}P,$$

$$\phi_{nn}(A) = \frac{n-p}{n-pa}R + \frac{n-q}{n-qb}C - \frac{n-p-q}{n-pa-qb}P.$$

However, Lemma 3.6 says the above three quantities are equal. So

$$0 = \phi_{1n}(A) - \phi_{nn}(A) = \frac{n(1-a)}{a(n-pa)}R - \frac{n(1-a) - q(b-a)}{a(n-pa-qb)}P$$

from which we deduce that

$$b = \frac{n-pa}{q}\left(1 - \frac{Pa(n-p-q)}{(n-pa)P - n(1-a)R}\right).$$

Substituting this value for b into the equation $0 = \phi_{n1}(A) - \phi_{nn}(A)$ reveals that

$$0 = Pa(1-a)(n-pa)(n-p-q)\left(a - \frac{n(R-P)((n-q)(C-P)+pR)}{q(R+C-P)(nR-(n-q)P)}\right).$$

Now P, a and $n-p-q$ are all positive by assumption. Also $n - pa > 0$, since $n \geqslant pa + qb$ and $qb > 0$. The root $a = 1$ corresponds to the case when A is row-stochastic, which is handled by Lemma 3.5. The only case remaining is when

$$\begin{aligned} a &= \frac{n(R-P)\,((n-q)(C-P)+pR)}{q(R+C-P)(nR-(n-q)P)} \\ &> (R-P)\frac{(n-q)(C-P)+p(R-P)}{q(R+C-P)} \\ &> \frac{n-q+p}{2p}\left(1 - \frac{n!}{n^n}\right)^2 \\ &> \frac{2p+1}{2p}\left(1 - 2\frac{n!}{n^n}\right), \end{aligned}$$

using (13) and $n - q > p \geqslant q$. But this means that

$$pa - p > \frac{1}{2} - (2p+1)\frac{n!}{n^n} > \frac{1}{2} - \frac{2n!}{n^{n-1}}$$

whereas Theorem 2.1 says that

$$pa - p < \sqrt{2pn!/n^n} < \sqrt{2n!/n^{n-1}}. \tag{14}$$

These last two results are incompatible for all $n \geqslant 7$ since in that case $n!/n^{n-1} < 1/20$.

For $4 \leqslant n \leqslant 6$ we proceed as follows:

$$a = \frac{n(R-P)\left((n-q)(R+C-P)-(n-p-q)R\right)}{q(R+C-P)(n(R-P)+qP)}$$
$$> \frac{n(R-P)\left((n-q)(2-n^{-n}n!)-(n-p-q)\right)}{q(R-P+1)(n(R-P)+qn^{-n}n!)}.$$

Elementary calculus shows this bound is minimised when $R-P=1$, so

$$a > \frac{n\left((n-q)(2-n^{-n}n!)-(n-p-q)\right)}{2q(n+qn^{-n}n!)}. \tag{15}$$

An exhaustive case analysis shows that (15) contradicts (14) for all integers n, p, q where $4 \leqslant n \leqslant 6$ and $n-q > p \geqslant q \geqslant 1$. $\square$

References

[1] G.-S. Cheon, S.-G. Hwang, Maximization of a matrix function related to the Dittert conjecture, Linear Algebra Appl. 165 (1992) 153–165.
[2] G.-S. Cheon, H.-W. Yoon, A note on the Dittert conjecture for permanents, Int. Math. Forum 1 (2006) 1943–1949.
[3] G.-S. Cheon, I.M. Wanless, An update on Minc's survey of open problems involving permanents, Linear Algebra Appl. 403 (2005) 314–342.
[4] G.-S. Cheon, I.M. Wanless, An interpretation of the Dittert conjecture in terms of semi-matchings, Discrete Math. 307 (2007) 2501–2507.
[5] G.P. Egoryĉev, Solution of the van der Waerden problem for permanents, Soviet Math. Dokl. 23 (1981) 619–622.
[6] D.F. Falikman, A proof of van der Waerden's conjecture on the permanent of a stochastic matrix, Mat. Zametki 29 (1981) 931–938.
[7] B. Hajek, A conjectured generalized permanent inequality and a multi-access problem, in: T. Cover, B. Gopinath (Eds.), Open Problems in Communication and Computation, Springer-Verlag, New York, 1987.
[8] S.-G. Hwang, A note on a conjecture on permanents, Linear Algebra Appl. 76 (1986) 31–44.
[9] S.-G. Hwang, On a conjecture of E. Dittert, Linear Algebra Appl. 95 (1987) 161–169.
[10] J. Körner, K. Marton, Random access communication and graph entropy, IEEE Trans. Inform. Theory 34 (1988) 312–314.
[11] C.-K. Li, On certain convex matrix sets, Discrete Math. 79 (1990) 323–326.
[12] H. Minc, Theory of permanents 1982–1985, Linear Multilinear Algebra 21 (1987) 109–148.
[13] R. Sinkhorn, A problem related to the van der Waerden permanent theorem, Linear Multilinear Algebra 16 (1984) 167–173.
[14] B.L. van der Waerden, Aufgabe 45, Jahresber Deutch Math.-Verein. 35 (1926) 117.