VAN DER WAERDEN'S CONJECTURE ON PERMANENTS AND ITS RESOLUTION

Thomas J. Laffey

Let $A = (a_{ij})$ be an nxn matrix. The *permanent*, per A, of A is given by the formula

$$\text{per } A = \sum_{\sigma \in S_n} a_{1,\sigma(1)} a_{2,\sigma(2)} \cdots a_{n,\sigma(n)}$$

where the sum is over the symmetric group S_n. Thus per A is obtained from det A by formally replacing the factors sign(σ) in the expansion of det A by +1. Let a_i be the i^{th} column of A. Then it is clear that per A = per$(a_1,\ldots,a_n)$ is multilinear. Also if A(i,j) is the (n-1)x(n-1) submatrix obtained from A by deleting row i and column j, we have the Laplace-like expansions

$$\text{per } A = \sum_{i=1}^{n} a_{ij} \text{per } A(i,j) \qquad (j=1,2,..n)$$
$$= \sum_{j=1}^{n} a_{ij} \text{per } A(i,j) \qquad (i=1,2,..n).$$

However, per A does not have the alternating properties of det A and it is not in general multiplicative, so it is not a similarity invariant. However, it is clear that per P^TAP = per A for all permutation matrices P,Q. This last property enables one to replace A by a matrix equivalent to A by permutation matrices in carrying out calculations and it is used many times without explicit mention in this article.

A real nxn matrix is called *doubly-stochastic* if its entries are non-negative and

$$\sum_{j=1}^{n} a_{ij} = 1 \qquad (i=1,2,\ldots,n)$$
$$\sum_{i=1}^{n} a_{ij} = 1 \qquad (i=1,2,\ldots,n)$$

Let DS(n) be the set of nxn doubly-stochastic matrices. Then DS(n) is a compact subset of $\mathbb{R}^{n^2}$. Let

$$f(n) = \inf\{\text{per } A \mid A \in DS(n)\}.$$

By compactness, there exist elements $A \in DS(n)$ with per A = f(n). Such a matrix A is called a ***minimizing matrix***. Thus A is a minimizing matrix if A is an (nxn) doubly-stochastic matrix such that its permanent achieves the ***absolute minimum*** of the permanent on the set of all doubly-stochastic matrices.

The famous van der Waerden conjecture (1926) states

Van der Waerden Conjecture

(1) $f(n) = n!/n^n$

(2) there is exactly one minimizing matrix, namely the matrix J_n that has all its entries equal to 1/n.

This conjecture was resolved in the affirmative by G.P. Egorychev of Krasnoyarsk in the U.S.S.R. in 1980. Independently D.I. Falikman, also from the U.S.S.R., proved part (1) of the conjecture in a paper submitted in 1979. Various special cases of the conjecture had been resolved earlier by various authors. Of particular beauty was the verification of the conjecture for the class of positive semi-definite symmetric doubly-stochastic matrices by Marcus and Newman (1962), later improved by Minc (1963), and the work of Friedland in the 1970s who showed in particular that per A > 1/n! Of particular relevance to subsequent interest in the problem as well as to its solution was the verification by Marcus and Newman (1959) of the conjecture for matrices that have all their entries positive. While the verification of the van der Waerden conjecture for n=2 is an elementary exercise, the problem quickly increases in difficulty as n increases and it was not until 1968 that Eberlein and Mudholkar settled the case n=4 and 1969 that Eberlein settled the case n=5.

In this expository article we present an account of Egorychev's work and describe the necessary background results. As well as Egorychev's own account [2] which appeared in English in Advances in Mathematics, an account of his work has been published by van Lint [10] and a detailed account with the background filled in has been given by Knuth in the American Mathematical Monthly [5]. The presentation here has been greatly influenced by the accounts of van Lint and Knuth. In the final section we describe a few more recent results.

A full and authoritative account of the properties and importance of permanents has been given by Minc in his enjoyable book [7]. The problem of computing permanents is described by Nijenhuis and Wilf in Chapter 23 of [8]. Permanents arise in many combinatorial problems and the "permanental polynomial" per (xI-A) is sometimes referred to as one of the isomorphism invariants of a graph with incidence matrix A.

1. Preliminaries

Let $A = (a_{ij})$ be an nxn matrix. The (directed) graph G(A) is the graph with vertices 1,2,3,..,n and such that for $i \neq j$, ij is a (directed) edge of G(A) if and only if $a_{ij} \neq 0$. G(A) is connected if for all $i \neq j$, there exists $s \geq 1$ and a sequence $i_0 = i, i_2, \ldots, i_s = j$ such that $i_0 i_1, i_1 i_2, \ldots, i_{s-1} i_s$ are edges of G(A). Equivalently, A is irreducible under permutation similarity, i.e. there is no permutation matrix P such that

$$P^T AP = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}$$

where A_{11} is an rxr, A_{22} an (n-r)x(n-r) matrix, some $1 \leq r < n$. (A special case of) the Perron-Frobenius theorem states that if A is a permutation irreducible non-negative real matrix, then A has a real eigenvalue r with $r \geq |\lambda|$ for all eigenvalues λ of A and r is a simple eigenvalue.

A theorem of Birkhoff states that the set DS(n) of doubly-stochastic matrices is precisely the set of convex combinations of the permutation matrices, i.e. $A \in DS(n)$ if and only if there exist non-negative real numbers $a(\sigma)$ with $\sum a(\sigma) = 1$ such that

$$A = \sum_{\sigma \in S_n} a(\sigma)P(\sigma)$$

where $P(\sigma)$ is the permutation matrix corresponding to σ. Note that this result in particular implies that $f(n) > 0$.

2. Minimizing Matrices

Throughout this section $A = (a_{ij}) \in DS(n)$ is such that per A = f(n).

Lemma 2.1 A is irreducible under permutation similarity.

Proof Suppose not. Since per $(P^T AP)$ = per A for P a permutation matrix, we may assume

$$A = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}$$

where A_{11} is rxr, A_{22} (n-r)x(n-r), some $1 \leq r < n$. Since $A \in DS(n)$, looking at the sum of all the entries in A_{11}, we see that $A_{11} \in DS(r)$, that $A_{12} = 0$ and thus that $A_{22} \in DS(n-r)$. Note that per A = per A_{11} per A_{22}.

Now a simple induction yields that per $A_{11} > 0$ and that per $A_{22} > 0$. We may assume that $a_{ii} > 0$, (i=1,2,...,n). Let A(e) be the matrix obtained from A by replacing a_{11} by $a_{11}-e$, $a_{1,r+1}$ by $a_{1,r+1}+e$, $a_{r+1,1}$ by $a_{r+1,1}+e$, $a_{r+1,r+1}$ by $a_{r+1,r+1}-e$. Then for sufficiently small $e > 0$, $A(e) \in DS(n)$. But a simple calculation yields per A(e) < per A for all sufficiently small $e > 0$. This is a contradiction.

The next result, due to Marcus and Newman, is crucial to the discussion.

Theorem 2.2 For all i,j for which $a_{ij} > 0$, we have $\operatorname{per} A(i,j) = \operatorname{per} A$.

Proof Let $Z = \{B \in DS(n) \mid b_{ij} = 0 \text{ if } a_{ij} = 0\}$.

Using Lemma 2.1, we see that A is an interior element of Z and hence it must satisfy the analytic criteria for a local minimum. A matrix $X = (x_{ij}) \in Z$ if the following conditions hold

$$x_{ij} \geqq 0 \quad (\text{all } i,j)$$

$$x_{ij} = 0 \quad \text{if } a_{ij} = 0$$

$$\sum_{j=1}^{n} x_{ij} - 1 = 0 \quad (i=1,..,n)$$

$$\sum_{i=1}^{n} x_{ij} - 1 = 0 \quad (j=1,..,n)$$

Introducing Lagrange multipliers, we consider the function

$$F(X) = \operatorname{per}(X) - \sum_{i=1}^{n} \lambda_i \Big(\sum_{j=1}^{n} x_{ij} - 1\Big) - \sum_{j=1}^{n} \mu_j \Big(\sum_{i=1}^{n} x_{ij} - 1\Big).$$

If $x_{ij} \neq 0$, the partial derivative

$$\frac{\partial F}{\partial x_{ij}} = \operatorname{per} X(i,j) - \lambda_i - \mu_j$$

so

$$(*) \qquad \operatorname{per} A(i,j) = \lambda_i + \mu_j \quad \text{if } a_{ij} \neq 0.$$

Now the expansions

$$\operatorname{per} A = \sum_{i=1}^{n} a_{ij} \operatorname{per} A(i,j)$$

$$= \sum_{j=1}^{n} a_{ij} \operatorname{per} A(i,j)$$

yield

$$(1) \qquad \operatorname{per} A = \lambda_i + \sum_{j=1}^{n} a_{ij}\mu_j$$

$$(2) \qquad \operatorname{per} A = \sum_{i=1}^{n} a_{ij}\lambda_i + \mu_j.$$

Let $e = (1,\ldots,1)^T$, $\lambda = (\lambda_1,\ldots,\lambda_n)^T$, $\mu = (\mu_1,\ldots,\mu_n)^T$. The equations become

$$(\operatorname{per} A)e = \lambda + A\mu$$

$$(\operatorname{per} A)e = A^T\lambda + \mu.$$

Since $A \in DS(n)$, $Ae = A^Te = e$. Thus we obtain

$$A^T\lambda + A^TA\mu = A^T\lambda + \mu$$

$$\lambda + A\mu = AA^T\lambda + A\mu.$$

Thus $A^TA\mu = \mu$, $AA^T\mu = \lambda$. But A and therefore AA^T, A^TA are irreducible with maximum eigenvalue 1 and corresponding eigenvector e. By the Perron-Frobenius theorem, 1 is a simple eigenvalue, so

$$\lambda_1 = \ldots = \lambda_n = a, \text{ say}$$

$$\mu_1 = \ldots = \mu_n = b, \text{ say}.$$

But then $\operatorname{per} A = \sum_{i=1}^{n} a_{ij} \operatorname{per} A(i,j) = a + b$ and the result follows.

Remark We note that Knuth [5] gives a purely combinatorial argument to establish (*).

We note also that Marcus and Newman were able to obtain a proof that if $a_{ij} > 0$ for all i,j, then $A = J_n$ easily from (2.2). This is not used in Egorychev's work, so we omit it. Details are given in Minc [7], page 79.

The following partial extension of (2.2) to the case where $a_{ij} = 0$ is due to London (1971) ([7], page 85).

Theorem 2.3 For all i,j, $\operatorname{per} A(i,j) \geqq \operatorname{per} A$.

Proof In proving the result for a pair i,j we may assume $a_{ij} = 0$. Using the remark on permutation equivalence in the introduction and (2.1) we may assume i=1, j=1 and further that $a_{kk} \neq 0$ (k=2,...,n).

Note that for sufficiently small $e > 0$, $(1-e)A + eI \in DS(n)$ and using the fact that for $C = (c_{ij})$, $D = (d_{ij})$,

$$p(C + eD) = \text{per } C + e \sum_{i,j=1}^{n} d_{ij} \text{ per } C(i,j) + O(e^2)$$

and (2.2) we obtain

$$\text{per}((1-e)A + eI) = \text{per } A + e(\text{per } A(1,1) - \text{per } A) + O(e^2).$$

Since A is minimizing, we obtain per $A(1,1) \geq$ per A, as required.

3. Aleksandrov's Inequality

The next ingredient in Egorychev's solution is (a special case of) an inequality of Aleksandrov (1938) [1]. This arose in the context of computing the volumes of convex sets.

Suppose $a_1,\ldots a_{n-2}$ are (column)-vectors in $\mathbb{R}^n$. We can define an inner product by

$$x.y = \text{per }(a_1,\ldots a_{n-2},x,y)$$

for $x,y \in \mathbb{R}^n$. (Of course this is not a positive definite inner product.) We may write $x.y = x^TQy$ for a symmetric matrix Q.

The result of Aleksandrov we require is

Theorem 3.1 Let $a_1,\ldots a_{n-1}$ be elements of $\mathbb{R}^n$ with all their entries positive. Then (using the notation above) for $x \in \mathbb{R}^n$

(*) $$(x.a_{n-1})^2 \geq (x.x)(a_{n-1}.a_{n-1})$$

with equality if and only if $x = ba_{n-1}$ for some real b. (Note

that (*) is the reverse of the Cauchy-Schwarz inequality valid for positive definite inner products.)

We show that Theorem (3.1) follows from

Theorem 3.2 Let $a_1,\ldots a_{n-2} \in \mathbb{R}^n$ have positive entries. Then (in the notation above) Q is non-singular and has exactly one positive eigenvalue.

For suppose (3.2) holds. Suppose x and a_{n-1} are independent. Then on the two dimensional space span(x,a_{n-1}), there exists an element $x + ha_{n-1}$ such that $(x + ha_{n-1}).(x + ha_{n-1}) < 0$. Thus

$$h^2a_{n-1}.a_{n-1} + 2hx.a_{n-1} + x.x = 0.$$

Since $a_{n-1}.a_{n-1} > 0$ (as all the a_i have positive entries) the discriminant of the polynomial

$$\lambda^2a_{n-1}.a_{n-1} + 2\lambda x.a_{n-1} + x.x$$

is positive, proving (*).

We now prove (3.2) by induction on n.

If n=2, $Q = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ and the result is trivial. Suppose $n > 2$ and that the result holds for n-1. We first show Q is non-singular. For suppose $Qx = 0$. Since $Q = (q_{ij})$ where

$$q_{ij} = \text{per}(a_1,\ldots,a_{n-2},e_i,e_j)$$

(where $e_1,\ldots e_n$ is the standard basis of $\mathbb{R}^n$) we see that

(†) $$\text{per}(a_1,\ldots a_{n-2},x,e_j) = 0 \quad (j=1,2,..,n).$$

This equation is the same as

$$\text{per}((a_1,\ldots,a_{n-2},x,e_j)(i,n)) = 0.$$

Applying the induction hypothesis and hence (*) to the (n-1)x(n-1) matrix

$$(a_1,\ldots,a_{n-2},x,e_j)(j,n)$$

and the fact that $(a_1,\ldots,a_{n-2},a_{n-2},e_j)(j,n)$ has a positive permanent, we obtain

$$\text{per}((a_1,\ldots,a_{n-3},x,x,e_j)(j,n)) \leq 0$$

with equality if and only if $x - ca_{n-2}$ is zero at all positions except possibly the j^{th} for some real c. But in the case of equality we must have $c = 0$ since a_{n-2} has positive entries.

Hence

(††) $$\text{per}(a_1,\ldots a_{n-3},x,x,e_j) \leq 0$$

with equality if and only if x has all its entries except possibly the j^{th} zero.

But by (†)

$$\text{per}(a_1,\ldots,a_{n-2},x,x) = 0$$

and since a_{n-2} has positive entries, this with (††) gives

$$\text{per}(a_1,\ldots,a_{n-3},e_j,x,x) = 0$$

for all j and hence x has all its entries zero.

Thus Q is non-singular. Let $Q(\lambda)$ be defined by replacing a_i by $\lambda e + (1-\lambda)a_i$ where $e = (1,1,\ldots,1)^T$. Applying the above argument to $Q(\lambda)$ we conclude that $Q(\lambda)$ is non-singular for $0 \leq \lambda \leq 1$. Hence by continuity, the number of positive eigen-values of $Q(0) = Q$ is the same as that of $Q(1)$. But $Q(1) = (n-1)!\,(E-I)$ where E is the nxn matrix that has all its entries 1 so the eigen values of $Q(1)$ are

$$(n-1)!\,(n-1),\ -(n-1)!,\ldots,\ -(n-1)!$$

So (3.2) holds.

By continuity we obtain from (3.1)

Corollary 3.3 If $a_1,\ldots a_{n-1} \in \mathbb{R}^n$ have non-negative entries, then for $x \in \mathbb{R}^n$

$$(x.a_{n-1})^2 \geq (x.x)(a_{n-1}.a_{n-1}).$$

4. Egorychev's Resolution

Suppose $A \in DS(n)$, $n \geq 3$ with per $A = f(n)$.

We first show that for all i,j

(†) $$\text{per } A(i,j) = \text{per } A.$$

This is true by (2.2) if $a_{ij} > 0$ and by (2.3), per $A(i,j) \geq$ per A if $a_{ij} = 0$. Suppose that for some i,j, per $A(i,j) >$ per A. Now for some t, $a_{it} > 0$. By Corollary 3.3,

$$\text{per}(a_1,\ldots,a_j,\ldots,a_t,\ldots,a_n)^2 \geq$$

$$\text{per}(a_1,\ldots,a_j,\ldots,a_j,\ldots a_n)\text{per}(a_1,\ldots,a_t,\ldots,a_t,\ldots,a_n).$$

Using the fact that per $A(u,v)$ = per A for $a_{uv} \neq 0$ and a_{it} per $A(i,j) > a_{it}$ per A, we see, by expanding the terms along the t^{th} column, that the right-hand side is greater than (per A)(per A). This is a contradiction.

Next, note that using (†)

$$\text{per}(a_1,a_2,\ldots a_n) = \text{per}(\tfrac{1}{2}(a_1+a_2),\ \tfrac{1}{2}(a_1+a_2),a_3,\ldots a_n)$$

and since the matrix on the right is also in DS(n) and hence minimizing, we may repeat this process to find a minimizing matrix

$$(b_1,b_2,\ldots,b_{n-1},a_n)$$

in which $b_1,b_2,\ldots b_{n-1}$ have positive entries. But now using (†) again

$$\text{per}(b_1,\ldots,b_{n-1},a_n)^2 =$$

$$\text{per}(b_1,\ldots b_{n-1},b_{n-1})\text{per}(b_1,\ldots,b_{n-2},a_n,a_n)$$

so by Aleksandrov's result (3.1), $b_{n-1} = c_{n-1}a_n$ for some real c_{n-1}. Expanding

$$\text{per}(b_1,\ldots b_{n-1},a_n)$$

by its $(n-1)^{st}$ and n^{th} columns and using (†) gives $c_{n-1} = 1$. Thus $b_{n-1} = a_n$. Similarly $b_{n-2} = a_n,\ldots,b_1 = a_n$. Hence since $A \in DS(n)$, $a_n = e/n$ where $e = (1,1,\ldots,1)^T$. Similarly $a_1 = a_2 = \ldots = a_{n-1} = e/n$. Thus $A = J_n$ and the conjecture is proved.

5. More Recent Developments

With the solution of the van der Waerden conjecture, the interest in permanents has increased rather than waned. Many conjectures related to the van der Waerden conjecture had been formulated and while several were special cases of the conjecture, some were more general. A detailed account is given in Minc [7] Chapter 8. We refer briefly to some recent work on a few of these conjectures.

Let $A \in DS(n)$ and let $\sigma_k(A)$ be the sum of the permanents of all the kxk submatrices of A. (Thus for example $\sigma_1(A) = n$, $\sigma_n(A) = \text{per } A$.)

The *Tverberg conjecture* (1963) states that if $A \in DS(n)$ and $2 \leq k \leq n$, then

$$\sigma_k(A) \geq \sigma_k(J_n)$$

with equality only if $A = J_n$. (The case k=n is the van der Waerden conjecture.) In a beautiful paper [4], Friedland has proved this conjecture. He first expresses $\sigma_k(A)$ as a permanent of a $(2n-k)x(2n-k)$ doubly-stochastic matrix having an $(n-k)x(n-k)$ block of zeros. Modifying Egorychev's methods and using many ingenious arguments, he then solves the more general problem of finding min(per A) taken over all $B \in DS(m)$ having a given rxr block of zeros.

Another conjecture more general than the van der Waerden conjecture is due to Djokovic (1967). In the notation of the last paragraph, Djokovic conjectures that for $k = 2,\ldots,n$, $A \in DS(n)$

$$\sigma_k((1 - \theta)J_n + \theta A)$$

is strictly increasing for $0 \leq \theta \leq 1$. Many special cases of this have been settled. Friedland and Minc [7] proved it for $k = n$, $A = J_n$ or $(nJ_n - I_n)/(n-1)$. London [6] has proved it for $k = n$, $A = \alpha I_n + \beta P_n$, $\alpha \geq 0, \beta \geq 0, \alpha + \beta = 1$, where P_n denotes the permutation matrix corresponding to the n-cycle $(1\ 2\ 3\ \ldots\ n)$ and for $A = (nJ_n - I_n - P_n)/(n-2)$ $(n > 2)$.

An important advance on this problem has been reported by Egorychev in his review of London's paper (MR 83g 15005). He asserts that if $f_0, f_1, f_{m+1},\ldots,f_n$ are column n-vectors with positive entries and $f_\lambda = \lambda f_0 + (1-\lambda)f_1$, $0 \leq \lambda \leq 1$, then the function $\text{per}^{1/m}(B)$ where

$$B = (f_\lambda, f_\lambda,\ldots,f_\lambda, f_{m+1},\ldots f_n)$$

is concave (convex upwards).

Finally we describe a conjecture of Schrijver and Valiant [9] which in his review of their paper, Minc (MR 82a 15004) suggests is a worthy successor to the van der Waerden conjecture.

Let Λ_n^k be the set of all nxn matrices with non-negative *integer* entries such that each row sum and each column sum equals k. Let

$$\lambda(n) = \min\{\text{per}(a) \mid A \in \Lambda_n^k\}$$

$$\theta_k = \lim_{n\to\infty} \lambda_k(n)^{1/n}$$

In their paper Schrijver and Valiant show that

$$(1) \quad \lambda_k(n) \leq k^{2n}/\binom{nk}{n}$$

$$(2) \quad \theta_k \leq (k-1)^{k-1}/k^{k-2}$$

and their conjecture states that (2) is an equality for all k. (The positive solution of the van der Waerden conjecture yields $\theta_k \geq k/e$ here.)

References

1. A.D. Aleksandrov, Math. Sbornik, 3 (1938), 227-251 (Russian with German summary).
2. G.P. Egorychev, Inst. Fiziki im L.V. Kirenskogo USSR Acad. Sc. Sibirsk, Preprint IFSO-13M Krasnoyarsk (1980) (Russian). Published in English in Advances in Math., 42, (1981), 299-305.
3. D.I. Falikman, Mat. Zametki, 29 (1981), 931-938 (Russian).
4. S. Friedland, Lin. & Multilin. Alg., 11 (1982), 107-120.
5. D.E. Knuth, Amer. Math. Monthly, 88 (1981), 731-740.
6. D. London, Lin. Alg. & Appl., 37 (1981), 235-249.
7. H. Minc, *Permanents* Encyclopedia Math. & Appl., Vol. 6, Addison-Wesley, Reading, Mass., 1978.
8. A. Nijenhuis and H.S. Wilf, *Combinatorial Algorithms*, Academic Press, New York, 1968.
9. A. Schrijver and W. Valiant, Nederl. Akad. Wetensch. Indag. Math., 42 (1980), 425-427.
10. J.H. van Lint, Lin. Alg. & Appl., 39 (1981), 1-8.

Mathematics Department,
University College,
Dublin.

SOME APPLICATIONS OF THE CLASSIFICATION OF FINITE SIMPLE GROUPS TO PERMUTATION GROUP THEORY [1]

Martin W. Liebeck

The classification of finite simple groups has made it possible to prove many new and striking results in the theory of finite permutation groups. We survey some of these results and describe some of the methods used in proving them. We also present a theorem on maximal subgroups of finite classical groups which is of use in extending the techniques.

(A) The Classification Theorem This states that any finite simple group is isomorphic to one of the following groups:

cyclic	Z_p		
alternating	A_n $(n \geq 5)$		
groups of Lie type	classical:	$PSL(n,q)$, $PSp(2m,q)$, $PSU(n,q)$, $P\Omega^{\pm}(n,q)$	
groups of Lie type	Chevalley:	$G_2(q)$, $F_4(q)$, $E_6(q)$, $E_7(q)$, $E_8(q)$	exceptional
	twisted:	${}^2B_2(q)$, ${}^2G_2(q)$, ${}^2F_4(q)$, ${}^3D_4(q)$, ${}^2E_6(q)$	
26 sporadic groups			

See [5] for descriptions of the groups of Lie type.

(B) Some Recent Applications to Permutation Groups. As explained, for example, in Sections 2 and 3 of [2], at the heart of

[1] This is the text of a talk given by the author to the Group Theory Conference held in Galway, 13-14 May, 1983, under the auspices of the I.M.S.