THE HADAMARD THEOREM FOR PERMANENTS

MARVIN MARCUS

1. Statement of the result. Let $A=(a_{ij})$ be an $n$-square non-negative hermitian matrix. A classical inequality of Hadamard states that

$$\det A \leq \prod_{i=1}^{n} a_{ii}$$

with equality if and only if $A$ has a zero row (and column) or $A$ is a diagonal matrix: $A = \text{diag} \left( a_{11}, \cdots, a_{nn} \right)$. Several generalizations are known (e.g. [1]).

Let $\text{per} \ A$ denote the permanent of $A$,

$$\text{per} \ A = \sum_{\sigma} \prod_{i=1}^{n} a_{i\sigma(i)},$$

where the summation extends over the whole symmetric group of degree $n$. It was conjectured in [4] that in analogy with (1),

$$\text{per} \ A \geq \prod_{i=1}^{n} a_{ii}$$

with the conditions of equality precisely those in the Hadamard determinant theorem. L. Mirsky recently listed this conjecture among several other problems concerning the permanent function [5]. This conjecture was suggested by an inequality of I. Schur [6] (see also [4]):

$$\text{per} \ A \geq \det A.$$

In an unsuccessful attempt [3] to prove (3) H. Minc and the present author obtained an inequality of the form

$$\text{per} \ A \geq c_{n} \prod_{i=1}^{n} a_{ii}$$

in which the constant $c_{n}$ depends only on $n$ and not on $A$.

It is the purpose of this paper to present the proof of an inequality that substantially generalizes (3) and to discuss the somewhat delicate cases of equality. Let $A(i)$ denote the principal submatrix of $A$ obtained by deleting row and column $i$ of $A$. The main result is contained in the following

Received by the editors June 26, 1963.
Theorem 1. Let \( A = (a_{ij}) \) be an \((r+1)\)-square non-negative hermitian matrix. Then
\[
(r + 1)a_{ii}\ \text{per} \ A(i) \geq \ \text{per} \ A \geq a_{ii}\ \text{per} \ A(i), \quad 1 \leq i \leq n.
\]
If \( A \) has a zero row then \( (4) \) is equality throughout. If \( A \) has no zero row then the lower equality holds if and only if \( a_{ii} \) is the only nonzero entry in row and column \( i \) of \( A \); the upper equality holds if and only if the rank of \( A \) is \( 1 \).

The permanent is unaltered by pre- and post-multiplication by permutation matrices so that we can take \( i = 1 \) in proving \( (4) \).

Once \( (4) \) is established it is clear that an obvious induction on \( r \) will yield

Theorem 2. If \( A = (a_{ij}) \) is an \( n \)-square non-negative hermitian matrix then
\[
\text{per} \ A \geq \prod_{i=1}^{n} a_{ii}
\]
with equality if and only if \( A \) has a zero row or \( A \) is a diagonal matrix.

2. Preliminaries. Let \( U \) be an \( n \)-dimensional unitary space with inner product \( \langle x, y \rangle \). For \( 1 \leq r \leq n \) define \( U^{(r)} \) to be the space of \( r \)-tensors on \( U \) [2, Chapter 7]; that is, \( U^{(r)} \) is the dual space of the space \( M_r(U) \) of all multilinear functionals of \( r \)-tuples of vectors from \( U \). If \( x_1, \ldots, x_r \) are in \( U \) then their tensor product \( x_1 \otimes \cdots \otimes x_r \in U^{(r)} \) is defined by
\[
x_1 \otimes \cdots \otimes x_r(\phi) = \phi(x_1, \ldots, x_r), \quad \phi \in M_r(U).
\]
An inner product in \( U^{(r)} \) is given by
\[
(x_1 \otimes \cdots \otimes x_r, y_1 \otimes \cdots \otimes y_r) = \prod_{i=1}^{r} \langle x_i, y_i \rangle.
\]
Define the completely symmetric operator \( S^{(r)} : U^{(r)} \to U^{(r)} \) by
\[
S^{(r)}(x_1 \otimes \cdots \otimes x_r) = \frac{1}{r!} \sum_{\sigma \in S_r} x_{\sigma(1)} \otimes \cdots \otimes x_{\sigma(r)},
\]
where \( S_r \) is the symmetric group of degree \( r \). The symmetric product of \( x_1, \ldots, x_r \) is then defined by
\[
x_1 \cdots x_r = S^{(r)}(x_1 \otimes \cdots \otimes x_r).
\]
The range space of \( S^{(r)} \) is the symmetry class of completely symmetric
tensors on $U$ denoted by $U(\omega)$ and since it is a subspace of $U^{(r)}$ it is unitary. By combining (6), (7), (8) we compute that the inner product of two symmetric products is

$$(9) \quad (x_1 \cdots x_r, y_1 \cdots y_r)_r = \frac{1}{r!} \text{per} \left( (x_i, y_i) \right), \quad 1 \leq i, j \leq r.$$  

(The formula (9) is an immediate consequence of the fact that $S^{(r)}$ is idempotent and hermitian.) Next let $G_{r,n}$ denote the totality of non-decreasing sequences of length $r$ chosen from $1, \cdots, n$. If $\omega \in G_{r,n}$, let $\mu(\omega)$ be the product of the factorials of the multiplicities of the distinct integers in $\omega$; e.g., $\mu(3,3,7,7,7,9) = 2!3!$. If $e_1, \cdots, e_n$ is an orthonormal (o.n.) basis of $U$ then the

$$\binom{n+r-1}{r}$$

symmetric products $\sqrt{(r!/\mu(\omega))}e_{a_1} \cdots e_{a_r} (\omega_1, \cdots, \omega_r) = \omega \in G_{r,n}$, constitute an o.n. basis of $U^{(r)}$ in the inner product $(\cdot, \cdot)_r$. We let $e_a = e_{a_1} \cdots e_{a_r}$.

3. Proofs. Assume that $A$ is $(r+1)$-square and has no zero row and let $D = \text{diag}(a_{11}^{-1/2}, 1, \cdots, 1)$. Then the $1, 1$ entry of $B = DAD$ is 1, per $B = a_{11}^{-1}$ per $A$, $A(1) = B(1)$, and $B$ is also non-negative hermitian. If we prove Theorem 1 for $B$ we will clearly have the result we want for $A$. Since $B$ is non-negative hermitian it follows that $B$ is a Gram matrix based on some set $e_1, v_1, \cdots, v_r$ where $e_1$ is a unit vector:

$$b_{11} = 1 = (e_1, e_1),$$

$$b_{1j+1} = b_{j+1,1} = (e_1, v_j), \quad j = 1, \cdots, r,$$

$$b_{ij+1} = (v_i, v_j), \quad i, j = 1, \cdots, r.$$ 

Let $e_2, \cdots, e_n$ be a completion of $e_1$ to an o.n. basis of $U$. Define a map $T: U(\omega) \to U^{(r+1)}$ by $T(e_a) = e_1 \cdot e_{a_1} \cdots e_{a_r}$, all $\omega = (\omega_1, \cdots, \omega_r) \in G_{r,n}$, and extend linearly. It is an easy consequence of the symmetry and linearity of the symmetric product in its factors that

$$T(x_1 \cdots x_r) = e_1 \cdot x_1 \cdots x_r$$

for all $x_1, \cdots, x_r$ in $U$. Let $R \subseteq U^{(r+1)}$ denote the range of $T$ and let $T^*$ denote the conjugate dual map of $T$; $T^*: R \to U^{(r)}$. That is, $T^*$ satisfies

$$(Th, g)_{r+1} = (h, T^*g), \quad \text{for all} \ h \in U^{(r)}, g \in R.$$
Next let \( f(x_1, \ldots, x_r) \) denote the Rayleigh quotient
\[
\frac{(T^* T x_1, \ldots, x_r, x_1, \ldots, x_r)_{r+1}}{(x_1, \ldots, x_r, x_1, \ldots, x_r)_r}
\]
(10)
\[
= \frac{(T^* x_1, \ldots, x_r, x_1, \ldots, x_r)_r}{(x_1, \ldots, x_r, x_1, \ldots, x_r)_r}.
\]

Now let \( H \) denote the non-negative hermitian transformation \( T^* T : U(r) \to U(\omega) \). It is easy to verify [4, Theorem 3] that \( x_1 \cdots x_r = 0 \) if and only if some \( x_i = 0 \). Thus \( f(x_1, \ldots, x_r) \) is defined for all sets of \( r \) nonzero vectors \( x_1, \ldots, x_r \) in \( U \) and it is known that such values of \( f \) lie in the interval between the largest and smallest eigenvalues of \( H \).

We compute the eigenvalues of \( H \) by obtaining a matrix representation of \( H \). The basis \( \sqrt{(r!/\mu(\omega))} e_{\omega}, \omega \in G_{r,n} \), ordered lexicographically in \( \omega \), is o.n. for \( U(\omega) \) and the \((r, \omega)\) entry of the matrix representation of \( H \) in this ordered basis is
\[
\frac{r!}{\sqrt{(\mu(\omega)\mu(\tau))}} (He_{\omega}, e_{\tau})_r.
\]
(11)

Now
\[
(He_{\omega}, e_{\tau})_r = (T^* T e_{\omega}, e_{\tau})_r = (Te_{\omega}, Te_{\tau})_{r+1}
\]
(12)
\[
= (e_{\omega_1}, \ldots, e_{\omega_r}, e_1, \ldots, e_{r+1})
\]
\[
= (e_{(1,\omega)}, e_{(1,\tau)})_{r+1}
\]
where \((1, \omega)\) is the sequence \((1, \omega_1, \ldots, \omega_r) \in G_{r+1,n}\) and similarly for \((1, \tau)\). The vectors \( \sqrt{(r+1)!/\mu(\omega)}) e_{\omega}, \omega \in G_{r+1,n} \), are an o.n. basis for \( U_{(r+1)} \) in the inner product \( (,)_r \). Moreover \((1, \omega) = (1, \tau)\) if and only if \( \omega = \tau \). Hence from (12) we have
\[
(He_{\omega}, e_{\tau})_r = \frac{\mu((1, \omega))}{(r+1)!} \delta_{\omega,\tau}.
\]

Thus the matrix representation of \( H \) in the basis \( \sqrt{(r!/\mu(\omega))} e_{\omega} \) is diagonal and the eigenvalues of \( H \) are seen from (11) to be
\[
\lambda(\omega) = \frac{r!}{\mu(\omega)} (He_{\omega}, e_{\omega})_r
\]
\[
= \frac{r!}{\mu(\omega)} \frac{\mu((1, \omega))}{(r+1)!}
\]
\[
= \frac{1}{r+1} \frac{\mu((1, \omega))}{\mu(\omega)}.
\]
Clearly \( \mu((1, \omega)) \geq \mu(\omega) \) with equality if and only if \( \omega_1 > 1 \). Thus the minimum eigenvalue of \( H \) is \( 1/(r+1) \) and the symmetric tensors \( e_\omega, \omega_1 < 1 \), constitute the totality of corresponding eigenvectors. Suppose the multiplicities greater than 1 of the distinct integers in \( \omega \) are \( m_1, \cdots, m_p \). Then the multiplicities for \((1, \omega)\) are either \( m_1, \cdots, m_p \) or \( m_1+1, m_2, \cdots, m_p \). In the first instance \( \omega_1 > 1 \) and \( \mu((1, \omega))/\mu(\omega) = 1 \); in the second \( \omega_1 = 1 \) and \( \mu((1, \omega))/\mu(\omega) = m_1+1 \). This latter expression is maximal only when \( m_1 = r \), i.e., for the sequence \( \omega = (1, \cdots, 1) \). Thus we conclude that

\[
1 \geq f(x_1, \cdots, x_r) \geq \frac{1}{r+1}.
\]

The lower equality holds if and only if \( x_1 \cdots x_r \) lies in the space spanned by the tensors \( e_\omega, \omega_1 > 1, \omega \in G_{r,n} \). The upper equality holds if and only if \( x_1 \cdots x_r \) is a multiple of \( e_1 \cdots e_1 \). Now by (10) we have

\[
f(v_1, \cdots, v_r) = \frac{(T_{v_1} \cdots v_r, T_{v_1} \cdots v_r)_{r+1}}{(v_1 \cdots v_r, v_1 \cdots v_r)_r} = \frac{(e_1 \cdot v_1 \cdots v_r, e_1 \cdot v_1 \cdots v_r)_{r+1}}{(v_1 \cdots v_r, v_1 \cdots v_r)_r} = \frac{1}{(r+1)!} \text{per } B = \frac{1}{r!} \frac{1}{\text{per } B(1)} = \frac{1}{r+1} \frac{\text{per } B}{\text{per } B(1)}\]

and it follows from (13) that

\[
(r+1) \text{ per } B(1) \geq \text{per } B \geq \text{per } B(1).
\]

As we indicated earlier (4) follows from (14).

We noted following (13) that the lower equality can hold if and only if

\[
v_1 \cdots v_r = \sum \epsilon_\omega e_\omega
\]

where the summation extends over all \( \omega \in G_{r,n} \) for which \( \omega_1 > 1 \). We prove that (15) implies that \( (v_i, e_i) = 0, i = 1, \cdots, r \). Let \( h \) denote the tensor on the right side of (15). Then
\[(v_1 \cdots v_r, e_1 \cdots e_i)_r = \prod_{k=1}^{r} (v_k, e_1)\]

and

\[(h, e_1 \cdots e_i)_r = \sum_{a_1 > 1} c_a(e_a, e_1 \cdots e_i)_r = 0.\]

Hence some \((v_j, e_i) = 0\) and from the symmetry of the symmetric product we can assume \(j = 1\). Suppose we have proved that \((v_1, e_i) = \cdots = (v_k, e_1) = 0\). Since no \(v_i = 0, i = 1, \cdots, r\), we know [4, Theorem 3] that \(0 \neq v_1 \cdots v_k \in U(k)\). The tensors \(e_a, a \in G_{k,n}\), constitute a basis for \(U(k)\) and because \(v_1 \cdots v_k \neq 0\) there exists an \(a \in G_{k,n}\) for which \((v_1 \cdots v_k, e_a) \neq 0\). Let \(\beta = (1, \cdots, 1, \alpha_1, \cdots, \alpha_k) \in G_{r,n}\) where the notation means that \(\alpha\) has been preceded with \(r - k\) 's to make a nondecreasing sequence of length \(r\). Now once again \((h, e_\beta) = 0\) and hence

\[0 = (v_1 \cdots v_r, e_\beta)_r = (v_1 \cdots v_k, v_{k+1} \cdots v_r, e_1 \cdots e_1 e_a \cdots e_a)_r\]

\[= \frac{1}{r!} \per \begin{bmatrix}
(v_1, e_1) & \cdots & (v_1, e_i) & (v_1, e_a) & \cdots & (v_1, e_a) \\
\vdots & & \vdots & \vdots & & \vdots \\
(v_k, e_1) & \cdots & (v_k, e_1) & (v_k, e_a) & \cdots & (v_k, e_a) \\
(v_{k+1}, e_1) & \cdots & (v_{k+1}, e_1) & (v_{k+1}, e_a) & \cdots & (v_{k+1}, e_a) \\
\vdots & & \vdots & \vdots & & \vdots \\
(v_r, e_1) & \cdots & (v_r, e_1) & (v_r, e_a) & \cdots & (v_r, e_a)
\end{bmatrix}.
\]

The upper left block in this last matrix is \(k \times (r - k)\), the upper right block is \(k \times k\), the lower left is \((r - k) \times (r - k)\) and the lower right is \((r - k) \times k\). The upper left block consists of zeros and hence by the Laplace expansion theorem for permanents we have

\[0 = \per \begin{bmatrix}
(v_1, e_a) & \cdots & (v_1, e_a) \\
\vdots & & \vdots \\
(v_k, e_a) & \cdots & (v_k, e_a) \\
(v_{k+1}, e_1) & \cdots & (v_{k+1}, e_1) \\
\vdots & & \vdots \\
(v_r, e_1) & \cdots & (v_r, e_1)
\end{bmatrix} = k! (v_1 \cdots v_k, e_a)_k (r - k)! \prod_{j=k+1}^{r} (v_j, e_1).
\]

Now \((v_1 \cdots v_k, e_a)_k \neq 0\) and hence some \((v_j, e_i) = 0, j = k+1, \cdots, r\). We can assume \((v_{i+1}, e_i) = 0\). Thus we have proved \((v_i, e_i) = 0, i = 1, \cdots, r\). In terms of the matrix \(B\) this implies that \(b_{ij} = b_{ji} = 0, j = 2, \cdots, r+1\), and hence \(B = 1 + B(1)\).
If the upper equality holds then \( v_1 \cdots v_r \) was seen to be a multiple of \( e_1 \cdots e_r \). Since we are assuming that \( A \) has no zero row it follows from [4, Theorem 3] that \( v_i = d_i e_1, \ i = 1, \ldots, r \), and hence \( B \) is the Gram matrix based on the set \( e_1, d_1 e_1, \cdots, d_r e_1 \). This means that \( B \) and hence \( A \) has rank 1, say \( A = (x, y), i, j = 1, \cdots, r + 1 \). Of course, when \( A \) has this form then

\[
\text{per} A = (r + 1)^2 \prod_{j=1}^{r+1} |z_j|^2, \quad \text{per} A(1) = r^1 \prod_{j=2}^{r+1} |z_j|^2
\]

and clearly \( \text{per} A = (r + 1)a_{11} \) per \( A(1) \). This completes the proof of Theorem 1.

References


University of California, Santa Barbara