EXTREMAL PROPERTIES OF HERMITIAN MATRICES

M. MARCUS AND J. L. McGREGOR

1. Introduction. In **(1)** Fan showed that if A is a Hermitian matrix with eigenvalues $\lambda_1 \leqslant \ldots \leqslant \lambda_n$ then, for $k \leqslant n$,

$$\max \sum_{j=1}^{k} (Ax_j, x_j) = \sum_{j=1}^{k} \lambda_{n-j+1},$$

$$\min \sum_{j=1}^{k} (Ax_j, x_j) = \sum_{j=1}^{k} \lambda_j,$$

where $x_1, \ldots, x_k$ run over all sets of k orthonormal (o.n.) vectors in unitary n-space V.

It is the purpose of this paper to extend this result to the compound of a non-negative Hermitian (n.n.h.) matrix and investigate some of the consequences of this extension.

In the sequel $\mathrm{tr}(L)$ will denote the trace of the matrix L and the Euclidean norm of L will be designated by $||L|| = (\mathrm{tr}(L^*L))^{\frac{1}{2}}$ where L^* is the conjugate transpose of L. $F(L)$ is the convex image of the unit sphere $||x|| = 1$ in the complex plane under the mapping $x \to (Lx, x)$.

For $1 \leqslant r \leqslant n$ let $V^{(r)}$ denote the rth compound space of V. A vector $z \in V^{(r)}$ will be designated by

$$z = x_1 \wedge \ldots \wedge x_r, \qquad x_i \in V,$$

where the indicated product is the usual Grassmann notation for the exterior product **(2)**. The inner product in $V^{(r)}$ is defined by

$$(x_1 \wedge \ldots \wedge x_r,\ y_1 \wedge \ldots \wedge y_r) = \det\{(x_i, y_j)\}_{i,j=1,\ldots,r}.$$

If A is a linear transformation on V to V then the induced compound of A on $V^{(r)}$ to $V^{(r)}$ is denoted by $C_r(A)$ and is defined by

$$C_r(A)x_1 \wedge \ldots \wedge x_r = Ax_1 \wedge \ldots \wedge Ax_r.$$

We list some of the essential properties of $C_r(A)$ that will subsequently be used **(6)**.

(i) $C_r(AB) = C_r(A)\, C_r(B)$.

(ii) If A is non-singular, normal, Hermitian, unitary, non-negative, then $C_r(A)$ has the corresponding property.

Received July 25, 1955; in revised form April 14, 1956.

(iii) The eigenvalues of $C_r(A)$ are all possible $\binom{n}{r}$ products of r of the eigenvalues of A.

To state subsequent results more compactly we introduce some notation. The set of $\binom{k}{r}$ distinct choices of integers satisfying $1 \leqslant i_1 < i_2 < \ldots < i_r \leqslant k$ will be denoted by Q_{kr} and a typical sequence in Q_{kr} will be denoted by ω. If $x_1, \ldots, x_k$ is a choice of k vectors in V then a typical product

$$x_{i_1} \wedge \ldots \wedge x_{i_r} \in V^{(r)}$$

will be denoted by x_ω. $E_r(a_1, \ldots, a_k)$ will denote the rth elementary symmetric function of the numbers $a_1, \ldots, a_k$:

$$E_r(a_1, \ldots, a_k) = \sum_{\omega \in Q_{kr}} \prod_{j=1}^{r} a_{i_j}.$$

2. Results on Hermitian matrices. The basic result is contained in

THEOREM 1. *Let $1 \leqslant r \leqslant k \leqslant n$ and let A be an n-square positive definite Hermitian matrix with eigenvalues $0 < \alpha_1 \leqslant \alpha_2 \leqslant \ldots \leqslant \alpha_n$. Then*

$$\max \sum_{\omega \in Q_{kr}} (C_r(A)\, x_\omega, x_\omega) = E_r(\alpha_n, \ldots, \alpha_{n-k+1}),$$
$$\min \sum_{\omega \in Q_{kr}} (C_r(A)\, x_\omega, x_\omega) = E_r(\alpha_1, \ldots, \alpha_k)$$

where both max and min are taken over all sets of k o.n. vectors $x_1, \ldots, x_k$ in V.

Proof. Set

$$g(x_1, \ldots, x_k) = \sum_{\omega \in Q_{kr}} (C_r(A)\, x_\omega, x_\omega).$$

First it is clear that a set of maximizing (minimizing) o.n. vectors exist. This is easily seen using a standard continuity argument. If $k = n$ then

$$g(x_1, \ldots, x_n) = \operatorname{tr} C_r(A) = E_r(\alpha_1, \ldots, \alpha_n)$$

and the result thus follows trivially whenever the number of vectors is equal to the dimension of the space. Now for $k < n$ let $y_1, \ldots, y_k$ be a minimizing set for g. The following argument is the same if $y_1, \ldots, y_k$ is a maximizing set. Consider the linear subspace L of V spanned by $y_1, \ldots, y_k$. Let P be the orthogonal projection onto L. Consider the mapping PA on L to L. Clearly if x and y belong to L then

$$(PAx, y) = (Ax, Py) = (Ax, y) = (x, Ay)$$
$$= (Px, Ay) = (x, PAy),$$

so that PA is positive definite Hermitian on L to L. Let $u_1, \ldots, u_k$ be o.n. eigenvectors of PA in L. Then

https://doi.org/10.4153/CJM-1956-059-0 Published online by Cambridge University Press

$$
\begin{aligned}
g(y_1, \ldots, y_k) &= \sum_{\omega \in Q_{kr}} (C_r(A)\, y_\omega, y_\omega) \\
&= \sum \det\{(Ay_{i_s}, y_{i_t})\}_{s,\, t=1,\ldots,r} \\
&= \sum \det\{(PAy_{i_s}, y_{i_t})\} \\
&= \sum (C_r(PA)\, y_\omega, y_\omega) = \operatorname{tr} C_r\,(PA) \\
&= \sum (C_r(PA)\, u_\omega, u_\omega) = \sum (C_r(A)\, u_\omega, u_\omega) \\
&= g(u_1, \ldots, u_k).
\end{aligned}
$$

At this point we prove a lemma reducing this situation to the case $k = n$.

LEMMA 1. *L is an invariant subspace of A.*

Proof. If L is not invariant under A we lose no generality in assuming that $Au_1 \notin L$. Then there exists a unit vector v in the orthogonal complement of L such that

$$\rho = (Au_1, v) \neq 0.$$

We define

$$
\begin{aligned}
u_1' &= \frac{u_1 - t\,\rho\, v}{\sqrt{1 + t^2|\rho|^2}} \\
u_j' &= u_j, \qquad\qquad j = 2, \ldots, k,
\end{aligned}
$$

where t is a real number. It is easy to check that $u_1', \ldots, u_k'$ is an o.n. set. Since $g(u_1, \ldots, u_k)$ is a minimum for g it follows that

$$\frac{d}{dt}\, g(u_1', \ldots, u_k') = 0 \qquad \text{for } t = 0.$$

Using the multilinearity of the Grassmann product we compute that for $t = 0$

$$
\begin{aligned}
&\frac{d}{dt}\,\Big(C_r(A)\, \frac{u_1 - t\,\rho\, v}{\sqrt{(1 + t^2|\rho|^2)}} \wedge u_{i_2} \wedge \ldots \wedge u_{i_r}, \frac{u_1 - t\,\rho\, v}{\sqrt{(1 + t^2|\rho|^2)}} \wedge u_{i_2} \wedge \ldots \wedge u_{i_r}\Big) \\
&\quad = -\rho(C_r(A)\, v \wedge u_{i_2} \wedge \ldots \wedge u_{i_r}, u_1 \wedge u_{i_2} \wedge \ldots \wedge u_{i_r}) \\
&\qquad - \bar{\rho}(C_r(A)\, u_1 \wedge u_{i_2} \wedge \ldots \wedge u_{i_r}, v \wedge u_{i_2} \wedge \ldots \wedge u_{i_r}) \\
&\quad = -2|\rho|^2 \prod_{j=2}^{r} (Au_{i_j}, u_{i_j}).
\end{aligned}
$$

Here we have used the fact that if $s, t \geqslant 2$ and $s \neq t$ then

$$(Au_{i_s}, u_{i_t}) = (PA\, u_{i_s}, u_{i_t}) = 0,$$

since $u_1, \ldots, u_k$ is an o.n. set of eigenvectors of PA on L to L. Furthermore it is clear that

$$\prod_{j=2}^{r} (Au_{i_j}, u_{i_j}) > 0,$$

https://doi.org/10.4153/CJM-1956-059-0 Published online by Cambridge University Press

and hence at $t = 0$

$$\frac{d}{dt} g(u_1', \ldots, u_k') \neq 0$$

and the proof of Lemma 1 is complete.

The proof of Theorem 1 is now easily completed. Since L is invariant under A, let B be the restriction of A to L. Then B is a positive definite Hermitian transformation on a k dimensional subspace onto itself and the eigenvalues of B are some k of the eigenvalues of A, say $\alpha_{i_1}, \ldots, \alpha_{i_k}$. Thus

$$\begin{aligned} g(y_1, \ldots, y_k) &= \sum_{\omega \in Q_{kr}} (C_r(B)\, y_\omega, y_\omega) \\ &= \operatorname{tr} C_r(B) = E_r(\alpha_{i_1}, \ldots, \alpha_{i_k}) \\ &\geqslant E_r(\alpha_1, \ldots, \alpha_k). \end{aligned}$$

Thus

$$g(x_1, \ldots, x_k) \geqslant E_r(\alpha_1, \ldots, \alpha_k)$$

for any o.n. vectors $x_1, \ldots, x_k$ and equality is attained by choosing a set of o.n. eigenvectors of A corresponding to $\alpha_1, \ldots, \alpha_k$.

Remark. Theorem 1 is true for A simply n.n.h. and can be established by continuity from the case A positive definite. Actually Fan's Theorem for the sum can be proved in exactly the same way using only the condition that A is Hermitian. It is worth noting that Theorem 1 cannot be obtained directly by applying Fan's result to $C_r(A)$. The difficulty arises from the fact that the lexicographic ordering of the eigenvalues of $C_r(A)$ does not necessarily coincide with the ordering by magnitude. Throughout this section we will assume A is n.n.h. unless otherwise stated. A result of A. Ostrowski **(5)** now follows easily.

COROLLARY 1. *For* $1 \leqslant r \leqslant k \leqslant n$

$$\min E_r((Ax_1, x_1), \ldots, (Ax_k, x_k)) = E_r(\alpha_1, \ldots, \alpha_k).$$

where the min is taken over all sets of k o.n. vectors $x_1, \ldots, x_k$ in V.

Proof. It follows from the Hadamard determinant Theorem and Theorem 1 that

$$\begin{aligned} E_r(\alpha_1, \ldots, \alpha_k) &\leqslant g(x_1, \ldots, x_k) \\ &= \sum_{\omega \in Q_{kr}} (C_r(A)\, x_\omega, x_\omega) \\ &\leqslant \sum_{\omega \in Q_{kr}} \prod_{s=1}^{r} (Ax_{i_s}, x_{i_s}) \\ &= E_r((Ax_1, x_1), \ldots, (Ax_k, x_k)). \end{aligned}$$

As before, the minimum is taken on.

https://doi.org/10.4153/CJM-1956-059-0 Published online by Cambridge University Press

COROLLARY 2. *Under the same hypotheses as Corollary* 1,

$$\max E_r((Ax_1, x_1), \ldots, (Ax_k, x_k)) = \binom{k}{r}\left(\frac{1}{k}\sum_{j=1}^{k} \alpha_{n-j+1}\right)^r.$$

Proof. By Fan's result

$$\max \sum_{i=1}^{k} (Ax_i, x_i) = E_1(\alpha_n, \ldots, \alpha_{n-k+1}).$$

Then by (**3**; Theorem 52)

$$E_r((Ax_1, x_1), \ldots, (Ax_k, x_k)) \leqslant \binom{k}{r}\left(\frac{E_1((Ax_1, x_1), \ldots, (Ax_k, x_k))}{k}\right)^r$$
$$\leqslant \binom{k}{r}\left(\frac{1}{k}\sum_{j=1} \alpha_{n-j+1}\right)^r.$$

We must show that this value is actually taken on. This is accomplished by use of the following elementary lemma.

LEMMA 2. *If T is a linear transformation on V to V then there exists an o.n. set of vectors $v_j \in V$, $j = 1, \ldots, m$, $m \leqslant n$ such that*

$$(Tv_j, v_j) = n^{-1}\operatorname{tr}(T), \qquad j = 1, \ldots, m.$$

Proof. We use an induction argument to exhibit a unitary matrix R such that

$$(R^* TR)_{ii} = n^{-1}\operatorname{tr}(T), \qquad i = 1, \ldots, m.$$

For $m = 1$ it is clear since $n^{-1}\operatorname{tr}(T) \in F(T)$. Suppose there exists a unitary U such that

$$U^*TU = \begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix}$$

with T_{11}, T_{22}, r and $(n - r)$ square matrices respectively and $(T_{11})_{ii} = n^{-1}\operatorname{tr}(T)$. Then $\operatorname{tr}(T_{22}) = (n - r)r^{-1}\operatorname{tr}(T)$ and applying the case $m = 1$ to T_{22} we select a unitary $(n - r)$-square matrix S such that

$$(S^* T_{22}S)_{11} = r^{-1}\operatorname{tr}(T).$$

Define the n-square unitary matrix V by $V = \operatorname{diag}(I, S)$ and set $R = UV$. This completes the induction.

Actually, for the purposes of this proof, we need only know Lemma 2 for T Hermitian. In this case we can readily exhibit an o.n. set v_j satisfying Lemma 2; let $u_1, \ldots, u_n$ be an o.n. set of eigenvectors of T and let θ be a primitive nth root of unity. Then set

$$v_i = \sum_{j=1}^{n} \frac{\theta^{ij}}{n^{\frac{1}{2}}} u_j.$$

Returning to the proof of Corollary 2, we select $y_n, \ldots, y_{n-k+1}$ corresponding to the eigenvalues $\alpha_n, \ldots, \alpha_{n-k+1}$ respectively. These span a subspace invariant under A and by restricting A to this k-dimensional subspace and applying

Lemma 2 to the restricted transformation we select k o.n. vectors $x_1, \ldots, x_k$ such that

$$(Ax_j, x_j) = \frac{1}{k}\sum_{j=1}^{k} \alpha_{n-j+1}.$$

Clearly, for this choice of the x_i

$$E_r((Ax_1, x_1), \ldots, (Ax_k, x_k)) = \binom{k}{r}\left(\frac{1}{k}\sum_{j=1}^{k} \alpha_{n-j+1}\right)^r,$$

and the proof is complete.

Corollary 3. *For* $1 \leqslant i_1 < i_2 < \ldots < i_k \leqslant n$,

$$\prod_{j=1}^{k} \alpha_j \leqslant \prod_{j=1}^{k} A_{i_j i_j} \leqslant \left(\frac{1}{k}\sum_{j=1}^{k} \alpha_{n-j+1}\right)^k.$$

Proof. Let ϵ_j be the unit vector with 1 in the jth position and 0 elsewhere. Then

$$(A\epsilon_{i_j}, \epsilon_{i_j}) = A_{i_j i_j}$$

and the result follows from Corollaries 1 and 2. We remark that for $k = n$ we have the Hadamard determinant inequality. We also note that the lower inequality is contained in (5).

Corollary 4. *If A is an arbitrary matrix with row vectors $A_1, \ldots, A_n$ then for* $1 \leqslant i_1 < \ldots < i_k \leqslant n$

$$\prod_{j=1}^{k} \alpha_j \leqslant \prod_{j=1}^{k} ||A_{i_j}|| \leqslant \left(\frac{1}{k}\sum_{j=1}^{k} \alpha_{n-j+1}^2\right)^{\frac{1}{2}k}$$

where $\alpha_1 \leqslant \ldots \leqslant \alpha_n$ are the non-negative square roots of the eigenvalues of A^*A.

Proof. Apply Corollary 3 to A^*A.

Corollary 5. *Assume A satisfies the conditions of Theorem 1. Let $0 \leqslant \omega_1 \leqslant \ldots \leqslant \omega_k$ be k non-negative numbers $k \leqslant n$. Then*

$$\min \prod_{j=1}^{k} (Ax_j, x_j)^{\omega_j} = \prod_{j=1}^{k} \alpha_j^{\omega_{k-j+1}}.$$

Proof.

$$\begin{aligned}\prod_{j=1}^{k} (Ax_j, x_j)^{\omega_j} &= \prod_{j=1}^{k} (Ax_j, x_j)^{\omega_1} \prod_{j=2}^{k} (Ax_j, x_j)^{\omega_2-\omega_1} \ldots (Ax_k, x_k)^{\omega_k-\omega_{k-1}} \\ &\geqslant \prod_{j=1}^{k} \alpha_j^{\omega_1} \prod_{j=1}^{k-1} \alpha_j^{\omega_2-\omega_1} \ldots \alpha_1^{\omega_k-\omega_{k-1}} \\ &= \prod_{j=1}^{k} \alpha_j^{\omega_{k-i+1}}\end{aligned}$$

and the latter value is clearly assumed.

https://doi.org/10.4153/CJM-1956-059-0 Published online by Cambridge University Press

COROLLARY 6. *If A and B are arbitrary n-square complex matrices then*

$$||AB||^2 \geqslant \max\left\{||A||^2\left(\prod_{j=1}^{n} \beta_j^{\alpha_{n-j+1}}\right)^{||A||^{-2}}, ||B||^2\left(\prod_{j=1}^{n} \alpha_j^{\beta_{n-j+1}}\right)^{||B||^{-2}}\right\}$$

where $0 \leqslant \alpha_i \leqslant \alpha_{i+1}$ *and* $0 \leqslant \beta_i \leqslant \beta_{i+1}$ $(i = 1, \ldots, n-1)$ *are the eigenvalues of* $A^* A$ *and* $B^* B$ *respectively.*

Proof. $||AB||^2 = \operatorname{tr}\{ABB^*A^*\} = \operatorname{tr}\{(A^*A)^{\frac{1}{2}} (BB^*) (A^*A)^{\frac{1}{2}}\}$. Let $y_1, \ldots, y_n$ be an o.n. set of eigenvectors of $(A^* A)^{\frac{1}{2}}$ corresponding respectively to $\alpha_1^{\frac{1}{2}},\ldots,\alpha_n^{\frac{1}{2}}$. Then by Corollary 5

$$\begin{aligned} ||AB||^2 &= \sum_{j=1}^{n} \alpha_j(B^* By_j, y_j) \\ &\geqslant ||A||^2\left(\prod_{j=1}^{n} (B^* By_j, y_j)^{\alpha_j}\right)^{||A||^{-2}} \\ &\geqslant ||A||^2\left(\prod_{j=1}^{n} \beta_j^{\alpha_{n-j+1}}\right)^{||A||^{-2}} \end{aligned}$$

The argument is symmetric in A and B and the result follows.

THEOREM 2. *Let A and B be n.n.h. with eigenvalues* $\alpha_1 \leqslant \ldots \leqslant \alpha_n$ *and* $\beta_1 \leqslant \ldots \leqslant \beta_n$ *respectively. Let* $0 \leqslant \theta_1 \leqslant \ldots \leqslant \theta_n$ *denote the eigenvalues of* $A + B$. *Then for* $r \leqslant k \leqslant n$,

$$E_r(\theta_1, \ldots, \theta_k) \geqslant \max\left\{\binom{k}{r}\sum_{s=0}^{r}\prod_{j=1}^{r-s}\beta_j E_s(\alpha_1, \ldots, \alpha_r), \binom{k}{r}\sum_{s=0}^{r}\prod_{j=1}^{r-s}\alpha_j E_s(\beta_1, \ldots, \beta_r)\right\},$$

$$E_r(\theta_n, \ldots, \theta_{n-k+1}) \leqslant \min\left\{\binom{k}{r}\sum_{s=0}^{r}\binom{k}{s}\left(\frac{1}{r-s}\sum_{j=1}^{r-s}\beta_{n-j+1}\right)^{r-s}\left(\frac{1}{r}\sum_{j=1}^{r}\alpha_{n-j+1}\right)^{s}, \binom{k}{r}\sum_{s=0}^{r}\binom{k}{s}\left(\frac{1}{r-s}\sum_{j=1}^{r-s}\alpha_{n-j+1}\right)^{r-s}\left(\frac{1}{r}\sum_{j=1}^{r}\beta_{n-j+1}\right)^{s}\right\}$$

Proof. Let $x_1, \ldots, x_k$ be an o.n. set of eigenvectors of $C = A + B$ corresponding respectively to $\theta_1, \ldots, \theta_k$. Let $a_i = (Ax_i, x_i)$ and $b_i = (Bx_i, x_i)$. Then

$$\begin{aligned} E_r(\theta_1, \ldots, \theta_k) &= E_r(a_1 + b_1, \ldots, a_k + b_k) \\ &= \sum_{i_1<\ldots<i_r=\omega\in Q_{kr}} \sum_{s=0}^{r} \sum_{\mu=(i_1<\ldots<i_s)\subset\omega} \prod_{j=1}^{s} a_{i_j} \prod_{t\in\omega-\mu} b_t \\ &\geqslant \sum_{\omega\in Q_{kr}} \sum_{s=0}^{r} \prod_{j=1}^{r-s} \beta_j E_s(a_{i_1}, \ldots, a_{i_r}) \\ &\geqslant \sum_{\omega\in Q_{kr}} \sum_{s=0}^{r} \prod_{j=1}^{r-s} \beta_j E_s(\alpha_1, \ldots, \alpha_r) \\ &= \binom{k}{r} \sum_{s=0}^{r} \prod_{j=1}^{r-s} \beta_j E_s(\alpha_1, \ldots, \alpha_r). \end{aligned}$$

The result is symmetric in A and B and the first inequality follows. The second inequality is proved analogously.

References

1. Ky Fan, *On a theorem of Weyl concerning eigenvalues of linear transformations* I, Proc. N.A.S. (U.S.A.), *35* (1949), 652–655.
2. H. Flanders, *A note on the Sylvester-Franke theorem*, Amer. Math. Monthly *60* (1954), 8, 543.
3. G. H. Hardy, J. E. Littlewood, G. Polya, *Inequalities* (Cambridge, 2nd ed., 1952).
4. A. Horn, *On the singular values of a product of completely continuous operators*, Proc. N.A.S. (U.S.A.), *36* (1950), 374.
5. A. Ostrowski, *Sur quelques applications des fonctions convexes et concaves au sens de I. Schur*, J. Math. Pures Appl. (9) *31* (1952), 253–292.
6. J. H. M. Wedderburn, *Lectures on matrices*, Amer. Math. Soc. Colloq. Pub., *17* (1934).

United States Naval Ordnance Test Station,
Pasadena, California and
University of British Columbia.

California Institute
of Technology.