AN EXTENSION OF THE MINKOWSKI DETERMINANT THEOREM

by MARVIN MARCUS AND WILLIAM R. GORDON†

(Received 21st September 1970)

Minkowski proved the following (for a proof see (**4**)): if A and B are $n \times n$ positive semi-definite hermitian matrices then

$$(\det (A+B))^{1/n} \geqq (\det A)^{1/n} + (\det B)^{1/n}. \tag{1}$$

It is known (**4**) that if both A and B are non-singular, then the equality holds in (1) if and only if $B = cA$ where c is a positive number.

In this note we shall investigate the cases of equality in an extension of the result (1).

Theorem 1. *For each $n \times n$ matrix X and each integer r, $1 \leqq r \leqq n$, let $d_r(X)$ denote the sum of all r-square principal subdeterminants of X. If A and B are n-square positive semi-definite hermitian matrices and $0 < q \leqq 1$, then*

$$d_r^{1/r}((A+B)^q) \geqq 2^{q-1} d_r^{1/r}(A^q) + 2^{q-1} d_r^{1/r}(B^q). \tag{2}$$

If A and B both have rank at least r and if $q < 1$, then equality holds in (2) *if and only if $A = B$. If $q = 1$ and $r > 1$, then equality holds in* (2) *if and only if $B = cA$ for some $c > 0$.*

We shall deduce Theorem 1 from Theorem 2 below. In order to state Theorem 2 we introduce some notation and definitions.

By C_r we denote the set of all k-tuples of non-negative reals with at least r positive coordinates. If f is a real valued function defined on k-tuples of reals then we say that

(i) f is *strictly C_r-concave* if f is concave on C_r and for x and y in C_r and for $0 < \theta < 1$, the equality $f(\theta x + (1-\theta)y) = \theta f(x) + (1-\theta) f(y)$ implies that x is a positive multiple of y, $x \sim y$;

(ii) f is *C_r-positive* means that $f(x) > 0$ if and only if $x \in C_r$;

(iii) f is *strictly C_r-monotone* if $f(x+u) > f(x)$ for x in C_r and for u any non-zero k-tuple of non-negative reals.

Theorem 2. *Let A and B be n-square positive semi-definite hermitian matrices with eigenvalues $0 \leqq \lambda_1 \leqq \ldots \leqq \lambda_n$ and $0 \leqq \mu_1 \leqq \ldots \leqq \mu_n$ respectively and let $0 \leqq \sigma_1 \leqq \ldots \leqq \sigma_n$ denote the eigenvalues of $A+B$. Let $1 \leqq k \leqq n$ and assume*

† The research of both authors was supported by the U.S. Air Force Office of Scientific Research under Grant AFOSR 698-67.

that $f(x) = f(x_1, \ldots, x_k)$ is symmetric concave and non-decreasing in each variable. Then

$$2f(\sigma_1, \ldots, \sigma_k) \geqq f(2\lambda_1, \ldots, 2\lambda_k)+f(2\mu_1, \ldots, 2\mu_k). \tag{3}$$

Assume that $1 \leqq r \leqq k$, and that f is strictly C_r-monotone, strictly C_r-concave and C_r-positive. Moreover assume that A and B both have rank at least $n-k+r$. Then equality can hold in (3) if and only if there exists a unitary matrix X such that

$$X^*(A+B)X = \operatorname{diag}(\sigma_1, \ldots, \sigma_n),$$

$$X^*AX = \operatorname{diag}(\lambda_1, \ldots, \lambda_k) \dotplus A_{n-k},$$

$$X^*BX = c \operatorname{diag}(\lambda_1, \ldots, \lambda_k) \dotplus B_{n-k},\ c>0,$$

$\mu_i = c\lambda_i$, $i = 1, \ldots, k$, (A_{n-k} and B_{n-k} are $(n-k)$-square matrices) and c satisfies

$$2f((1+c)\lambda_1, \ldots, (1+c)\lambda_k) = f(2\lambda_1, \ldots, 2\lambda_k)+f(2c\lambda_1, \ldots, 2c\lambda_k). \tag{4}$$

Proof. Let $x_1, \ldots, x_n$ be orthonormal eigenvectors of $A+B$ corresponding respectively to $\sigma_1 \leqq \ldots \leqq \sigma_n$. Then

$$\begin{aligned} f(\sigma_1, \ldots, \sigma_k) &= f(((A+B)x_1, x_1), \ldots, ((A+B)x_k, x_k)) \\ &= f((Ax_1, x_1)+(Bx_1, x_1), \ldots, (Ax_k, x_k)+(Bx_k, x_k)) \\ &= f\left(\frac{(2Ax_1, x_1)+(2Bx_1, x_1)}{2}, \ldots, \frac{(2Ax_k, x_k)+(2Bx_k, x_k)}{2}\right) \\ &\geqq \tfrac{1}{2}[f((2Ax_1, x_1), \ldots, (2Ax_k, x_k))+f((2Bx_1, x_1), \ldots, (2Bx_k, x_k))]. \end{aligned}$$

Suppose that $u_1, \ldots, u_n$ are orthonormal eigenvectors of A corresponding to $\lambda_1, \ldots, \lambda_n$ respectively. Then

$$(Ax_i, x_i) = \sum_{j=1}^{n} |(x_i, u_j)|^2 \lambda_j.$$

Since the vectors $x_1, \ldots, x_n$ are also orthonormal it follows that the matrix S whose (i,j) element is $|(x_i, u_j)|^2$ is doubly stochastic. Thus by a theorem of Birkhoff (**1**; **2**, p. 97) S is a convex combination of permutation matrices

$$S = \sum_{\sigma \in G} c_\sigma P_\sigma,$$

where G is a subset of S_n, the permutation group of degree n. Let λ denote the n-tuple $(\lambda_1, \ldots, \lambda_n)$. For each permutation σ let λ^σ denote $(\lambda_{\sigma(1)}, \ldots, \lambda_{\sigma(n)})$ and for each n-tuple $x = (x_1, \ldots, x_n)$ let $x[k] = (x_1, \ldots, x_k)$. Then the concavity of f implies that

$$\begin{aligned} f((Ax_1, x_1), \ldots, (Ax_k, x_k)) &= f\Big(\sum_{\sigma \in G} c_\sigma \lambda^\sigma[k]\Big) \\ &\geqq \sum_{\sigma \in G} c_\sigma f(\lambda^\sigma[k]). \end{aligned}$$

https://doi.org/10.1017/S0013091500009597 Published online by Cambridge University Press

Since f is symmetric and non-decreasing and $\lambda_1 \leqq \ldots \leqq \lambda_n$,

$$f(\lambda^\sigma[k]) \geqq f(\lambda_1, \ldots, \lambda_k).$$

Thus $f((Ax_1, x_1), \ldots, (Ax_k, x_k)) \geqq f(\lambda_1, \ldots, \lambda_k)$. Similarly it follows that

$$f((Bx_1, x_1), \ldots, (Bx_k, x_k)) \geqq f(\mu_1, \ldots, \mu_k).$$

Hence

$$f(\sigma_1, \ldots, \sigma_k) \geqq \tfrac{1}{2}[f(2\lambda_1, \ldots, 2\lambda_k) + f(2\mu_1, \ldots, 2\mu_k)].$$

This proves the inequality. Suppose that equality holds in (3), f satisfies the given conditions, and A and B have rank at least $n-k+r$. Then at least r of the inner products (Ax_i, x_i), $i = 1, \ldots, k$, must be positive and similarly at least r of the inner products (Bx_i, x_i), $i = 1, \ldots, k$, must be positive. Now in (3) we proved the following result: let $H = (h_{ij})$ be an n-square positive semi-definite hermitian matrix with eigenvalues $0 \leqq \gamma_1 \leqq \ldots \leqq \gamma_n$; let $1 \leqq r \leqq k \leqq n$ and suppose that f is a real valued function defined on the set of k-tuples of non-negative reals which is symmetric, concave and non-decreasing in each variable. Then for any set of k orthonormal vectors $x_1, \ldots, x_k$

$$f((Hx_1, x_1), \ldots, (Hx_k, x_k)) \geqq f(\gamma_1, \ldots, \gamma_k);$$

if f is strictly C_r-monotone, strictly C_r-concave and C_r-positive and if at least r of the inner products (Hx_j, x_j), $j = 1, \ldots, k$ are positive then the preceding inequality is equality if and only if

$$Hx_j = \gamma_{\phi(j)} x_j, \quad j = 1, \ldots, k$$

for some permutation ϕ on $\{1, \ldots, k\}$, i.e., $x_1, \ldots, x_k$ is an orthonormal set of eigenvectors corresponding to $\gamma_1, \ldots, \gamma_k$ in some order. In view of this result and the strict C_r-concavity of f we can conclude that

$$Ax_j = \lambda_{\phi(j)} x_j, \quad j = 1, \ldots, k, \quad \phi \in S_k,$$

$$Bx_j = \mu_{\theta(j)} x_j, \quad j = 1, \ldots, k, \quad \theta \in S_k,$$

and

$$c\lambda[k]^\phi = \mu[k]^\theta, \quad c > 0, \tag{5}$$

where $\lambda = (\lambda_1, \ldots, \lambda_n)$, $\mu = (\mu_1, \ldots, \mu_n)$. Let $\tau = \theta\phi^{-1}$ so that (5) becomes

$$\mu_{\tau(i)} = c\lambda_i, \quad i = 1, \ldots, k. \tag{6}$$

Since $\lambda_1 \leqq \ldots \leqq \lambda_k$ and $\mu_1 \leqq \ldots \leqq \mu_k$ we conclude from (6) that $\mu_{\tau(i)} = \mu_i$, $i = 1, \ldots, k$, i.e., $c\lambda_i = \mu_i$, $i = 1, \ldots, k$. But then

$$\begin{aligned}\mu_{\phi(i)} &= c\lambda_{\phi(i)} \\ &= \mu_{\theta(i)}\end{aligned}$$

or

$$\mu_{\phi(i)} = \mu_{\theta(i)}, \quad i = 1, \ldots, k.$$

However

$$\begin{aligned}\sigma_i &= ((A+B)x_i, x_i) = (Ax_i, x_i) + (Bx_i, x_i) = \lambda_{\phi(i)} + \mu_{\theta(i)} \\ &= \lambda_{\phi(i)} + \mu_{\phi(i)} = \lambda_{\phi(i)} + c\lambda_{\phi(i)} = (1+c)\lambda_{\phi(i)}, \quad i = 1, \ldots, k.\end{aligned}$$

From $\sigma_1 \leqq \ldots \leqq \sigma_k$, it follows that $\lambda_{\phi(i)} = \lambda_i$, $i = 1, \ldots, k$, and hence

$$\mu_{\theta(i)} = \mu_i = c\lambda_i, \quad i = 1, \ldots, k.$$

Thus $\sigma_i = (1+c)\lambda_i$, $i = 1, \ldots, k$, and equality holds if and only if (4) holds.

Corollary. *Let A, B, and f satisfy the conditions of Theorem 2 and let $k = n$. If f is homogeneous of degree $q \neq 0$ then*

$$f(\sigma_1, \ldots, \sigma_n) \geqq 2^{q-1}f(\lambda_1, \ldots, \lambda_n) + 2^{q-1}f(\mu_1, \ldots, \mu_n). \tag{7}$$

If A and B both have rank at least r then (7) is equality if and only if $B = cA$, $c > 0$. If $q \neq 1$ then equality can hold if and only if $B = A$.

Proof. According to Theorem 2 there exists a unitary X such that

$$X^*AX = \text{diag}\,(\lambda_1, \ldots, \lambda_n),$$

$$X^*BX = c\,\text{diag}\,(\lambda_1, \ldots, \lambda_n), \quad c > 0,$$

and (from (4))

$$(1+c)^q = 2^{q-1}(1+c^q). \tag{8}$$

It is easy to check that for $q \neq 1$, $c = 1$ is the only positive solution to (8). This completes the proof of the Corollary.

In (3) we show that if $E_r(y_1, \ldots, y_n)$ denotes the rth elementary symmetric function of $y_1, \ldots, y_n$ and if $f(x_1, \ldots, x_n) = E_r^{1/r}(x_1^q, \ldots, x_n^q)$ with $0 < q \leqq 1$, then for $r > 1$, or $r = 1$ and $q < 1$, f is strictly C_r-concave. With this choice of f, Theorem 1 now follows from the Corollary.

REFERENCES

(1) G. Birkhoff, Tres observaciones sobre el algebra lineal, *Univ. Nac. Tucuman Rev.*, Ser. A **5** (1946), 147-150.

(2) M. Marcus and H. Minc, *A Survey of Matrix Theory and Matrix Inequalities* (Prindle, Weber and Schmidt, Boston, Mass., 1964).

(3) W. R. Gordon and M. Marcus, An analysis of equality in certain matrix inequalities I, *Pacific J. Math.* **34** (1970), 407-413.

(4) L. Mirsky, *An Introduction to Linear Algebra* (Oxford 1955), 419.

University of California, Santa Barbara
University of Victoria, Victoria B.C.