Comparison Theorems for Symmetric Functions of Characteristic Roots

Marvin Marcus

(February 10, 1961)

Several theorems are proved that give necessary and sufficient conditions for $A-B$ to be positive semidefinite Hermitian. The conditions are in terms of the comparison of elementary symmetric functions of the characteristic roots of $A+X$ and $B+X$ as $X$ varies over positive definite Hermitian matrices.

1. Introduction

In a recent paper [4], M. Stone obtained the following result in reposing certain theorems of Ehrenfeld [1]: if $F$ and $G$ are positive definite $n$-square Hermitian matrices, then

$$d(I+AF) \geq d(I+AG)$$

for all positive definite $n$-square Hermitian matrices $A$ if and only if $F-G$ is positive semidefinite Hermitian. Here $d$ denotes determinant and henceforth $A>0$ \((\geq 0)\) will signify that $A$ is positive (positive semidefinite) Hermitian. \(F \geq G\) \((F>G)\) will mean $F-G \geq 0$ \((F-G>0)\). Note that since $A>0$ if and only if $A^{-1}>0$, the condition (1) is the same as saying

$$d(A+F) \geq d(A+G)$$

for all $A \geq 0$. It is in the form (2) that we investigate what happens when we replace $d$ by some other elementary symmetric function of the characteristic roots. To be specific, suppose $E_r(x_1, \ldots, x_n)$ denotes the $r^{th}$ elementary symmetric function of the indicated variables:

$$E_r(x_1, \ldots, x_n) = \sum_{1 \leq i_1 < \cdots < i_r \leq n} x_{i_1} \cdots x_{i_r}.$$ 

For a fixed $r$, $1 \leq r \leq n$, let $E_r(A)$ denote $E_r(\lambda_1(A), \ldots, \lambda_n(A))$ where $\lambda_i(A)$, $i=1, \ldots, n$, are the characteristic roots of $A$. If the $\lambda_i(A)$ are real we will choose our notation so that $\lambda_1(A) \geq \cdots \geq \lambda_n(A)$. Then the problem we pose is to find conditions on the characteristic roots of $F$ and $G$ such that

$$E_r(A+F) \geq E_r(A+G)$$

for all $A \geq 0$.

We have

Theorem 1. Assume $F \geq 0, G \geq 0$, and $1 \leq r \leq n$.

\begin{equation}
E_{r-p}(\lambda_1(F), \ldots, \lambda_{p+1}(F)) \geq E_{r-p}(\lambda_1(G), \ldots, \lambda_{n-p}(G)),
\end{equation}

$p=0, \ldots, r-1$

it follows that

$$E_r(A+F) \geq E_r(A+G) \text{ for all } A \geq 0.$$  \hspace{1cm} (4)

In case $G = I$, (3) becomes

$$E_{r-p}(\lambda_1(F), \ldots, \lambda_{p+1}(F)) \geq \left(\frac{n-p}{r-p}\right), \quad p=0, \ldots, r-1,$$

which is both necessary and sufficient for

$$E_r(A+F) \geq E_r(A+I) \quad \text{for all } A \geq 0.$$  \hspace{1cm} (6)

We remark that for $r=n$ the second part of theorem 1 simply becomes: $\lambda_n(F) \geq 1$ if and only if $d(A+F) \geq d(A+I)$ for all $A \geq 0$. Now in (1) multiply both sides by $d(A^{-1}R^*R)$ where $R$ is a non-singular matrix satisfying $R^*GR = I$, $R^*FR = K$ and we find that

$$d(R^*A^{-1}R+K) \geq d(R^*A^{-1}R+I).$$

Now as $A$ runs over all positive definite matrices so does $R^*A^{-1}R$ and from the above remark we conclude that $\lambda_n(K) \geq 1$. The characteristic roots of $K$ are just the characteristic roots of $G^{-1}F$ and thus

$$G^{-1/2}FG^{-1/2} \geq I, \quad F \geq G,$$

and the result in [4] follows. A somewhat stronger multiplicative analogue of theorem 1 is available.

Theorem 2. Assume $F \geq 0, G \geq 0$, and $1 \leq r \leq n$.

Then

$$E_r(AF) \geq E_r(AG) \text{ for all } A \geq 0.$$  \hspace{1cm} (7)
In order to prove these results we use some results concerning compound matrices and Grassmann products [2]. We give the coordinate definition of these items and list several of their properties. If \( x_1, \ldots, x_r \) are vectors in the unitary space of \( n \)-tuples, \( x_i = (x_{i1}, \ldots, x_{in}) \), \( 1 \leq r \leq n \), then \( x_1 \times \cdots \times x_r \) is the \( \binom{n}{r} \)-tuple whose coordinates are the \( r \)-square subdeterminants of the \( r \times n \) matrix \( (x_{ij}) \), \( i = 1, \ldots, r, j = 1, \ldots, n \) arranged in lexicographic order. If \( A \) is an \( n \times n \) matrix, then \( C_r(A) \) is the \( \binom{n}{r} \)-square matrix whose entries are the \( r \)-square subdeterminants of \( A \) arranged in doubly lexicographic order according to the row and column indices of \( A \). That is, if \( 1 \leq i_1 < \cdots < i_r \leq n \) and \( 1 \leq j_1 < \cdots < j_r \leq n \) are two increasing sets of \( r \) integers then the \( (i_1, \ldots, i_r) \) element of \( C_r(A) \) is the determinant of the matrix \( A[i_1, \ldots, i_r] \). The element of \( C_r(A) \) is the determinant of the matrix \( A[i_1, \ldots, i_r] \) whose \((s, t)\) entry is \( A_{i_s,j_t} \). If \( s = j_1, \ s = 1, \ldots, r \) we denote the corresponding principal submatrix by \( A[i_1, \ldots, i_r] \). If \( A \neq 0 \) then \( C_r(A) > 0 \); \( C_r(A) = 0 \) is a polynomial in (13) whose degree is \( r \). Since 1 > \( \cdots > \alpha \) we denote the ordered complementary set of \( \alpha \) by \( \alpha_1, \ldots, \alpha_r \). Setting \( \alpha = 1 \) in (13) we have

\[
\phi_1, \ldots, \phi_r = \sum' \left\{ d(K[\alpha_1, \ldots, \alpha_r]) \right\}, \tag{14}
\]

where the summation is taken only over those \( \alpha_1 < \cdots < \alpha \), containing \( \omega = (i_1 < \cdots < i_r) \) as a subset. Returning to (11) we set \( x_1 = \ldots = x_r = t = 0 \) and all other \( x_j \neq 0 \). Then \( \phi(0, \ldots, t, \ldots, t, 0) = t^r \phi_1, \ldots, \phi_r \). Conversely, if \( \phi_1, \ldots, \phi_r > 0 \) and \( \phi(0, \ldots, 0, 0) > 0 \), then \( \phi_1, \ldots, n_1, 0 \) \( n \) for all non-negative \( n_1, \ldots, n_r \). Setting \( \alpha = 1 \) in (14) we have

\[
\phi_1, \ldots, \phi_r = \sum' \left\{ d(K[\alpha_1, \ldots, \alpha_r]) \right\}. \tag{15}
\]
where $\sum' \geq$ indicates that the summation extends over precisely those increasing sequence $\sigma_i \leq \ldots \leq \sigma_{r-p}$ which are subsets of $j_1 \leq \ldots \leq j_{r-p}$. Also $u_j = U e_j$, $j = 1, \ldots, n$ is an orthonormal set of vectors (recall that $U$ is unitary). We use an extremal result in [3: theorem 1, p. 525] to conclude from (15) that

$$
\phi^0_{a_1 \ldots a_r} \geq E_{r-p}(\lambda_{p+1}(F), \ldots, \lambda_p(F))
$$

In case $G = I$ we have from (15) again that

$$
\phi^0_{a_1 \ldots a_r} = \sum'(C_{r-p}(F) u_{a_1} \ldots u_{a_{r-p}}) - (\Phi_{a_1 \ldots a_r})
$$

for all choices of sets of $n-p$ orthonormal vectors $u_1, \ldots, u_{n-p}$, and another application of the above cited extremal result completes the proof.

To proceed to the proof of theorem 2, choose a nonsingular $R$ such that $F = RR^*$, $G = RD^*$, $D = \text{diag} (\lambda_1(F^{-1} G), \ldots, \lambda_p(F^{-1} G))$. Let $P$ be an arbitrary nonsingular matrix and since any $A > 0$ is of the form $(PR^{-1})(PR^{-1})^*$ we have that (7) is equivalent to

$$
E_r((PR^{-1})^*(PR^{-1} RR^*) \geq E_r((PR^{-1})^*(PR^{-1} RD^*)^*)
$$

or

$$
E_r(P^*P) \geq E_r(P^*PD)
$$

for all $A > 0$. (16)

In (16) replace $A$ by $V^*X^*$, $V$ unitary, $X = \text{diag} (x_1, \ldots, x_n) \geq 0$ to obtain

$$
E_r(X) \geq E_r(XH), \quad H = V^*DV
$$

or

$$
tr[C_r(X)(I - C_r(H))]=0, \quad \quad (17)
$$

where $I$ is the $(n)$-square identity matrix. It is not difficult to check that (17) holds for all non-negative diagonal $X$ if and only if every diagonal element of $I - C_r(H)$ is non-negative. That is,

$$
1-C_r(V^*DV)e_{a_1} \ldots \cdot e_{a_r} = e_{a_1} \ldots \cdot e_{a_r} \geq 0
$$

must hold for all $V$ and all $1 \leq a_1 < \ldots < a_r \leq n$. But this is precisely equivalent to

$$
(C_r(D) u_{a_1} \ldots u_{a_r}) \leq 1
$$

for all orthonormal $u_{a_1}, \ldots, u_{a_r}$. As in the proof of theorem 1 we have finally that (7) holds if and only if (8) does.

To prove theorem 3 it will be convenient to let denote the relation of Hermitian congruence. Then if $A > 0$,

$$
C_r(I + AF) - C_r(I + AG) = C_r(A)
$$

$$
[C_r(A^{-1} + F) - C_r(A^{-1} + G)] = (C_r(A))^{1/2} (C_r(A) - 1/2)
$$

This last matrix has the same roots as the Hermitian matrix $C_r(A^{-1} + F) - C_r(A^{-1} + G)$ and hence (9) is equivalent to

$$
C_r(A + F) \geq C_r(A + G) \quad \text{for all } A > 0.
$$

Now

$$
C_r(A + F) - C_r(A + G) = C_r(A + F) - (C_r(A))^{1/2}
$$

$$
C_r(G^{-1/2} A G^{-1/2} + I) (C_r(A))^{1/2} \geq C_r(G^{-1/2} A G^{-1/2} + I) + G^{-1/2} F G^{-1/2} C_r(G^{-1/2} A G^{-1/2} + I).
$$

Thus (9) is equivalent to

$$
C_r(A + I) - C_r(A + I) \geq 0 \quad \text{for all } A > 0 \quad (18)
$$

where $H = G^{-1/2} F G^{-1/2}$.

By a unitary congruence we may assume $H = \text{diag} (h_1, \ldots, h_n)$ and by setting $A = \text{diag} (x_1, \ldots, x_n) \geq 0$ we see that

$$
\Pi_{i=1}^n (x_{i} + h_{i}) \geq \Pi_{i=1}^n (x{i} + 1)
$$

for any non-negative numbers $x_{i_1}, \ldots, x_{i_r}$. This clearly implies that each $h_{i} \geq 1$, $i = 1, \ldots, n$. Thus, $0 \geq n - I = G^{-1/2} F G^{-1/2} - I \leq F - G$.

Conversely suppose $F - G \geq 0$. Then $H \geq I$ and if we set $B = A + I > 0$ we would like to conclude that

$$
C_r(A + I) = C_r(B + H - I) \geq C_r(B) = C_r(A + I). \quad (19)
$$

But (19) is equivalent to

$$
C_r(I + K) \geq C_r(I), \quad K = B^{-1/2}(H - I)B^{-1/2}. \quad (20)
$$

After a unitary congruence (20) reduces to

$$
C_r(I + \text{diag} (k_1, \ldots, k_n)) \geq C_r(I),
$$

where $k_{a} \geq 0$ are the characteristic roots of $K$. The proof is complete.

2. References


