SYMMETRIC MEANS AND MATRIX INEQUALITIES

PETER BULLEN AND MARVIN MARCUS

1. Introduction. The purpose of this note is to present several new inequalities for elementary symmetric functions and exploit these to investigate the inequalities that exist among the eigenvalues of submatrices and sums of positive definite hermitian (p.d.h.) matrices.

We describe our notation: \((a) = (a_1, \cdots, a_n)\) will denote an \(n\)-tuple of positive numbers, \((a_i)\) is the \((n-1)\)-tuple obtained by deleting \(a_i\) from \((a)\). \(E_r(a)\) denotes the \(r\)th elementary symmetric function of the \((a)\):

\[
E_r(a) = \sum_{1 \leq i_1 < \cdots < i_r \leq n} \prod_{j=1}^{r} a_{i_j}, \quad E_0 = 1.
\]

\(\bar{p}_r(a)\) is the mean of \(E_r(a)\),

\[
\bar{p}_r = \bar{p}_r(a) = \binom{n}{r}^{-1} E_r(a).
\]

If \((a) = (a_1, \cdots, a_n)\) then \((a) = (a_1, \cdots, a_n, a_{n+1})\) and \(p_r(a_1, \cdots, a_{n+1}) = p_r.\) If two sets \((a)\) and \((b)\) are related by \(a_i = \lambda b_i, \lambda > 0, i = 1, \cdots, n\) then we say \((a)\) is proportional to \((b)\). We note that if \(A\) is an \(n\)-square complex matrix with eigenvalues \(\alpha_1, \cdots, \alpha_n\) then \((-1)^r E_r(\alpha)\) is the coefficient of \(x^{n-r}\) in the characteristic polynomial \(\det(xI - A)\).

2. Results.

Theorem 1. If \(1 \leq r \leq k \leq n\) then

\[ \frac{p_r}{p_k} \leq \frac{p_k+1}{p_k+1}, \]

with equality if and only if \(a_1 = \cdots = a_{n+1}\).

We give two proofs.

Proof 1. The following results are known [2, p. 52]

\[
\bar{p}_s = \frac{n + 1 - s}{n + 1} p_s + \frac{s}{n + 1} a_{n+1 p_{s-1}},
\]

(3) if \(s < t\) then \(p_s^{1/s} \geq p_t^{1/t}\) with equality if and only if \(a_1 = \cdots = a_n\),

(4) if \(a_1 = \cdots = a_n = \alpha\) then

Presented to the Society, January 24, 1961; received by the editors May 24, 1960.

1 The work of this author was supported by U. S. Air Force Contract No. AF 49(638)-776.
If \( r = k \) then (1) is implied by (3) so we assume \( r < k \). Using (2) with both \( s = p \) and \( s = k+1 \) (1) is seen to be equivalent to

\[
L = \left\{ \frac{n - k}{n + 1} \left( \frac{p_{k+1}}{p_k} \right) + \frac{k + 1}{n + 1} \right\}^r p_r \leq \left\{ \frac{n + 1 - r}{n + 1} p_r + \frac{r}{n + 1} a_{n+1} p_{r-1} \right\}^{k+1} = R.
\]

By (3)

\[
R \geq \left\{ \frac{n + 1 - r}{n + 1} p_r + \frac{r}{n + 1} a_{n+1} p_{r-1} \right\}^{k+1}.
\]

By (4) with \( \alpha = p_r^{-1/r} \) the right-hand side of (5) is the mean of the \( r \)th elementary symmetric function of the numbers \( a_1 = a_2 = \cdots = a_n = p_r^{-1/r} \) and \( a_{n+1} \). Hence by (3) with \( t = r \) and \( s = k+1 \) \( R \) is not less than

\[
\left\{ \frac{n - k}{n + 1} \left( \frac{p_{k+1}}{p_k} \right) + \frac{k + 1}{n + 1} \right\}^r p_r \leq L.
\]

This proves the inequality. The inequality is strict in (3) and hence in our arguments unless \( a_1 = \cdots = a_n = p_r^{-1/r} = a_{n+1} \), completing the proof.

**Proof. 2.** In this argument we show that the right-hand side of (1) is a convex function of \( a_{n+1} \) with at most one critical value in

\( a_{n+1} > 0 \).

Let \( x = a_{n+1} \) and set

\[
\phi(x) = \frac{p_r^{k+1}}{p_{k+1}^k}.
\]

Since \( \phi(x) \) is of the form \((x+a)^{k+1}/(x+b)^r \) to within a positive constant multiple and \( k > r \) \((k=r \) is again an application of (3)) we can check that \( \phi''(x) > 0 \) for \( x \geq 0 \) and hence \( \phi(x) \) is convex. Thus it suffices to show that

\[
\phi(0) \geq \frac{p_r^k}{p_k^r},
\]

and
\[(7) \quad \phi(x_0) \geq \frac{p^k}{p^r_k} \quad \text{when} \quad \phi'(x_0) = 0, \quad x_0 > 0.\]

Now
\[
\phi(0) = \frac{(n + 1 - r)^{k+1}}{(n + 1)^{k+1-r}(n - k)^r} \frac{p^k_r}{p^k_{r+1}}
\]
and by (3) we need only check that
\[(8) \quad (n + 1 - r)^{k+1} > (n - k)^r(n + 1)^{k+1-r}\]
whenever \(1 \leq r \leq k \leq n\). If \(k = n\) (8) is clear, otherwise let \(K = k + 1\), \(N = n + 1\) and (8) becomes
\[(9) \quad \left(1 - \frac{r}{N}\right)^K > \left(1 - \frac{K}{N}\right)^r, \quad \frac{r}{N} < \frac{K}{N} < 1.\]

Taking natural logarithms of both sides and expanding, (9) is immediately implied by \(r < K\). We next see that \(\phi'(x_0) = 0\) and \(x_0 \geq 0\) implies
\[
x_0 = \frac{(n + 1 - r)p_r p_k - (n - k)p_{k+1}p_{r-1}}{(k + 1 - r)p_k p_{r-1}},
\]
and hence
\[
\phi(x_0) = (\frac{p^k_r}{p^r_k})(\frac{p^r_{r-1}}{p^r_r}) \cdot \left[\frac{(k + 1)(n + 1 - r) - r(n - k)(p_{r-1}/p_r)(p_{k+1}/p_k)}{(n + 1)(k + 1 - r)}\right]^{k+1-r}.
\]

By (3) it suffices to show that the quantity in the square brackets is greater than 1. This is equivalent to
\[(10) \quad p_{r-1} p_{k+1} \leq p_r p_k\]
which is established by an easy induction on \(k - r\) from the case \(k = r [2, p. 52]\). Now, the only time that equality can hold in (3) and (10) is for \(a_1 = \cdots = a_n = a\). But then we check that \(x_0 = a\) also. The proof is complete.

**Corollary.** If \(A_n\) and \(G_n\) are the arithmetic and geometric means of \((a)\) then
\[(11) \quad (A_n/G_n)^n \leq (A_{n+1}/G_{n+1})^{n+1}\]
with equality if and only if \(a_1 = \cdots = a_{n+1}\).

This result is an analogue of an inequality due to R. Rado [2, p. 61]:

License or copyright restrictions may apply to redistribution; see http://www.ams.org/journal-terms-of-use
The inequality (12) can be derived in the same way as (11) was from (3) with \( s = n \) and \( t = 1 \).

Theorem 2. Let \( H \) be a p.d.h. \((n+1)\)-square matrix and let \( K \) be an \( n \)-square principal submatrix of \( H \). If \( \mu_1 \geq \cdots \geq \mu_{n+1} > 0 \) and \( \lambda_1 \geq \cdots \geq \lambda_n > 0 \) are the eigenvalues of \( H \) and \( K \) respectively then

\[
\frac{p_k(\lambda_1, \ldots, \lambda_n)}{p_k(\lambda_1, \ldots, \lambda_n)} \leq \frac{p_{k+1}(\mu_1, \ldots, \mu_{n+1})}{p_{k+1}(\mu_1, \ldots, \mu_{n+1})}. 
\]

Proof. Let \( K \) be obtained from \( H \) by deleting row and column \( i \) of \( H \). Let \( x_1, \ldots, x_n \) be an orthonormal (o.n.) set of \( n \)-dimensional eigenvectors of \( K \). Let \( y_1, \ldots, y_n+1 \) be the \((n+1)\)-vectors given by

\[
y_i = (x_{i1}, \ldots, x_{i,i-1}, 0, x_{i,i}, \ldots, x_{in}),
\]

where \( x_j = (x_{j1}, \ldots, x_{j,i-1}, x_{ji}, \ldots, x_{jn}) \) and let \( y_{n+1} \) be the \((n+1)\)-vector with 1 in position \( i \), zero elsewhere. Then we check easily that \( y_1, \ldots, y_{n+1} \) is an o.n. set of \((n+1)\)-vectors and moreover

\[
(Hy_j, y_j) = (Kx_j, x_j), \quad j = 1, \ldots, n.
\]

Then

\[
\frac{p_k(\lambda_1, \ldots, \lambda_n)}{p_k(\lambda_1, \ldots, \lambda_n)} = \frac{\left( \sum_{j=1}^{n} (Kx_j, x_j)/n \right)^k}{p_k((Kx_1, x_1), \ldots, (Kx_n, x_n))} \leq \frac{\left( \sum_{j=1}^{n+1} (Hy_j, y_j)/(n+1) \right)^{k+1}}{p_{k+1}((Hy_1, y_1), \ldots, (Hy_{n+1}, y_{n+1}))} \leq \frac{p_{k+1}(\mu_1, \ldots, \mu_{n+1})}{p_{k+1}(\mu_1, \ldots, \mu_{n+1})}. 
\]

The inequalities above follow from:

(i) the sum of the quadratic forms over a complete o.n. set is the trace; (ii) \( p_{k+1}((Hy_1, y_1), \ldots, (Hy_{n+1}, y_{n+1})) \geq p_{k+1}(\mu_1, \ldots, \mu_{n+1}) \) for any o.n. set \( y_1, \ldots, y_{n+1} \) [4].

Corollary. If \( \tau_k \) and \( d_k \) be the trace and determinant of \( H_k \), the \( k \)-square principal submatrix of \( H \) whose elements lie in the first \( k \) rows and columns of \( H \), then
It might be conjectured that Theorem 2 follows from the Cauchy inequalities [1, p. 75]. However a proof via this route is made difficult by the fact that \( p_k^p / p_k^k \) is not an increasing function.

In [5] it was proved that \( E_r / E_{r-1} \) and \( E_{r/p} \) are both concave functions for positive variables. We extend these results as follows.

**Theorem 3.** If \( 1 \leq p \leq r \leq n \) and \( Fr,p = (E_r / E_{r-p})^{1/p} \) then

\[
Fr,p(a + b) \geq Fr,p(a) + Fr,p(b)
\]

with equality if and only if \( r = 1, p = 1 \) or \( (a) \) is proportional to \( (b) \).

Moreover, \( Fr,p \) is a nondecreasing function of each \( a_j \).

**Proof.**

\[
Fr,p(a + b) = \left( \prod_{j=1}^{p} \frac{E_{r-j+1}(a + b)}{E_{r-j}(a + b)} \right)^{1/p}
\]

\[
\geq \left\{ \prod_{j=1}^{p} \left( \frac{E_{r-j+1}(a)}{E_{r-j}(a)} \right) + (E_{r-j+1}(b)/E_{r-j}(b)) \right\}^{1/p}
\]

\[
\geq \left\{ \prod_{j=1}^{p} E_{r-j+1}(a)/E_{r-j}(a) \right\}^{1/p} + \left\{ \prod_{j=1}^{p} E_{r-j+1}(b)/E_{r-j}(b) \right\}^{1/p}
\]

\[
= Fr,p(a) + Fr,p(b).
\]

The first inequality above is the result in [5] and inequality can hold if and only if \( (a) \) is proportional to \( (b) \). The second inequality is the Hölder inequality. Conversely if \( (a) \) is proportional to \( (b) \) the equality is easily checked.

We next compute

\[
\frac{\partial}{\partial a_1} \left\{ E_r(a)/E_{r-p}(a) \right\} = \left\{ E_{r-p}(a)E_{r-1}(a') - E_r(a)E_{r-p-1}(a') \right\}/E_{r-p}(a).
\]

Now by (2) the numerator in this last expression becomes

\[
E_{r-1}(a')(E_{r-p}(a) - E_r(a)E_{r-p-1}(a')) \geq 0.
\]

This last inequality is found in [2, p. 52].

We apply Theorem 3 to obtain

**Theorem 4.** If \( A \) and \( B \) are \( n \)-square p.d.h. matrices with eigenvalues \( 0 < \alpha_1 \leq \cdots \leq \alpha_n \), \( 0 < \beta_1 \leq \cdots \leq \beta_n \) respectively and \( C = A + B \) has eigenvalues \( 0 < \delta_1 \leq \cdots \leq \delta_n \), then for \( 1 \leq p \leq r \leq k \leq n \),
\[(15) \quad F_{r,p}(\delta_1, \ldots, \delta_k) \geq F_{r,p}(\alpha_1, \ldots, \alpha_k) + F_{r,p}(\beta_1, \ldots, \beta_k).\]

**Proof.** By Theorem 3 and the fact that \(F_{r,p}\) is homogeneous of degree 1 we conclude that \(F_{r,p}\) is concave for positive variables. By a general extremal result in [3] it follows that if \(x_1, \ldots, x_k\) are an o.n. set of vectors then

\[(16) \quad F_{r,p}(Ax_1, x_1), \ldots, Ax_p, x_p) \geq F_{r,p}(\alpha_i_1, \ldots, \alpha_i_k)\]

for some index set \(1 \leq i_1 < \cdots < i_k \leq n\). But since \(F_{r,p}\) is nondecreasing we know that

\[(17) \quad F_{r,p}(\alpha_1, \ldots, \alpha_k) \geq F_{r,p}(\alpha_1, \ldots, \alpha_k).\]

Now select \(x_1, \ldots, x_k\) to be an o.n. set of eigenvectors of \(C\) corresponding to \(\delta_1, \ldots, \delta_k\) respectively. Then

\[
F_{r,p}(\delta_1, \ldots, \delta_k) = F_{r,p}((Cx_1, x_1), \ldots, (Cx_k, x_k))
\]

\[
= F_{r,p}(Ax_1, x_1) + (Bx_1, x_1), \ldots, (Ax_k, x_k) + (Bx_k, x_k))
\]

\[
\geq F_{r,p}(Ax_1, x_1), \ldots, (Ax_k, x_k)) + F_{r,p}(Bx_1, x_1), \ldots, (Bx_k, x_k))
\]

\[
\geq F_{r,p}(\alpha_1, \ldots, \alpha_k) + F_{r,p}(\beta_1, \ldots, \beta_k)
\]

in which we have used Theorem 3 and (16) and (17) in succession.

**Corollary.** If \(x^n + \sum_{j=1}^{n-1} (-1)^j c_j(A)x^{n-j}\) is the characteristic polynomial of \(A\) then

\[
(c_r(C)/c_{r-p}(C))^{1/p} \geq (c_r(A)/c_{r-p}(A))^{1/p} + (c_r(B)/c_{r-p}(B))^{1/p}.
\]

**Proof.** Take \(k = n\) in Theorem 4.

**References**


THE UNIVERSITY OF BRITISH COLUMBIA, VANCOUVER, CANADA