INEQUALITIES FOR SYMMETRIC FUNCTIONS AND HERMITIAN MATRICES

M. MARCUS AND L. LOPES

1. Introduction. The purpose of this paper is to present two concavity results for symmetric functions and apply these to obtain inequalities connecting the characteristic roots of the non-negative Hermitian (n.n.h.) matrices A, B and $A+B$.

H. F. Bohnenblust has recently communicated to one of the present authors a suggestion of H. Samelson indicating that Theorem 2 below can be derived from a concavity result on mixed volumes of convex bodies due to Fenchel **(2)**. Bohnenblust has also obtained a separate proof of Theorem 2 by different methods and we are indebted to him for several valuable conversations concerning his results. Theorem 1 in the sequel is believed to be new and we show that it actually provides a simple and direct algebraic proof of Theorem 2 without using the results on mixed volumes. We have not been able to find an explicit statement of Theorem 2 in the literature. Both the present proof and the unpublished proof of Bohnenblust were constructed before Samelson observed the connection with Fenchel's result.

In §2 we state and prove the two results on symmetric functions. In §3 these results are applied to obtain inequalities connecting the characteristic roots of n.n.h. matrices. Theorem 4 extends the classical Minkowski determinant inequality to all of the coefficients in the characteristic polynomial and Theorem 5 extends a recent concavity result of Fan **(1)** for determinants.

2. Results on Symmetric Functions. Let $(a) = (a_1, \ldots, a_n)$ be a set of real numbers and let $E_r(a)$ denote the rth elementary symmetric function of (a), $1 \leqslant r \leqslant n$:

$$E_r(a) = \sum_{1 \leqslant i_1 < \ldots < i_r \leqslant n} \prod_{j=1}^{r} a_{i_j},$$

where the summation extends over all $\binom{n}{r}$ choices of $i_1 < \ldots < i_r$ from $1, \ldots, n$. If $r = 0$ then $E_r(a) = 1$. If $(a) = (a_1, \ldots, a_n)$ and $(b) = (b_1, \ldots, b_n)$ then let $(a+b) = (a_1 + b_1, \ldots, a_n + b_n)$. We confine our attention here to sets (a) of non-negative numbers at least r of whose elements are assumed positive. If there exists a number λ such that

$$a_j = \lambda\, b_j, \qquad j = 1, \ldots, n,$$

we shall write $(a) \sim (b)$. If this is not the case we indicate it by $(a) \nsim (b)$.

Received May 14, 1956. Part of this research was completed under U.S. Air Force contract AF 18(603)–83.

https://doi.org/10.4153/CJM-1957-037-9 Published online by Cambridge University Press

The set $a_1, \ldots, a_{j-1}, a_{j+1}, \ldots, a_n$ will be denoted by (a'_j). We shall also use the abbreviation

$$\mathbf{E}_r(a) = \frac{E_r(a)}{E_{r-1}(a)}.$$

THEOREM 1. *If* $1 \leqslant r \leqslant n$ *then*

$$\mathbf{E}_r(a+b) > \mathbf{E}_r(a) + \mathbf{E}_r(b)$$

unless $(a) \sim (b)$ *or* $r = 1$ *in which cases we have equality.*

Proof. We use an induction on r. If $(a) \sim (b)$ we clearly have equality, hence assume $(a) \nsim (b)$. For $r = 2$ the result follows from

$$\mathbf{E}_2(a+b) - \mathbf{E}_2(a) - \mathbf{E}_2(b) = \frac{\sum_{i=1}^{n}\left(a_i \sum_{j=1}^{n} b_j - b_i \sum_{j=1}^{n} a_j\right)^2}{2\sum_{i=1}^{n}(a_i+b_i)\sum_{i=1}^{n} a_i \sum_{i=1}^{n} b_i}.$$

Hence assume $r > 2$. We first note that

(1) $$\sum_{i=1}^{n} a_i E_{r-1}(a'_i) = r E_r(a).$$

Now

$$E_r(a) = a_i E_{r-1}(a'_i) + E_r(a'_i),$$

and summing on i we have

$$n E_r(a) = \sum_{i=1}^{n} a_i E_{r-1}(a'_i) + \sum_{i=1}^{n} E_r(a'_i).$$

Using (1) it follows that

(2) $$\sum_{i=1}^{n} E_r(a'_i) = (n-r) E_r(a).$$

Now write

$$\begin{aligned} E_r(a) - E_r(a'_i) &= a_i E_{r-1}(a'_i) \\ &= a_i E_{r-1}(a) - a_i^2 E_{r-2}(a'_i) \end{aligned}$$

summing over i and using (2) we have

(3) $$\begin{aligned} r E_r(a) &= \sum_{i=1}^{n} a_i E_{r-1}(a) - \sum_{i=1}^{n} a_i^2 E_{r-2}(a'_i), \\ \frac{E_r(a)}{E_{r-1}(a)} &= \frac{1}{r}\left\{\sum_{i=1}^{n} a_i - \sum_{i=1}^{n} a_i^2 \frac{E_{r-2}(a'_i)}{E_{r-1}(a)}\right\} \\ &= \frac{1}{r}\left\{\sum_{i=1}^{n} a_i - \sum_{i=1}^{n} \frac{a_i^2}{a_i + \mathbf{E}_{r-1}(a'_i)}\right\}. \end{aligned}$$

https://doi.org/10.4153/CJM-1957-037-9 Published online by Cambridge University Press

From (3) we calculate that

$$\phi = \mathbf{E}_r(a+b) - \mathbf{E}_r(a) - \mathbf{E}_r(b)$$
$$= \frac{1}{r}\sum_{i=1}^{n}\left\{\frac{a_i^2}{a_i + \mathbf{E}_{r-1}(a'_i)} + \frac{b_i^2}{b_i + \mathbf{E}_{r-1}(b'_i)} - \frac{(a_i+b_i)^2}{a_i + b_i + \mathbf{E}_{r-1}(a'_i + b'_i)}\right\}.$$

We make the induction hypothesis that the inequality holds for $r - 1$. Since $r \geqslant 3$ we have

$$\mathbf{E}_{r-1}(a'_i + b'_i) > \mathbf{E}_{r-1}(a'_i) + \mathbf{E}_{r-1}(b'_i)$$

unless $(a'_i) \sim (b'_i)$ in which case we have equality. We first consider the case for which there exists an i such that $(a'_i) \nsim (b'_i)$; then it follows that

$$\phi > \frac{1}{r}\sum_{i=1}^{n}\left\{\frac{a_i^2}{a_i + \mathbf{E}_{r-1}(a'_i)} + \frac{b_i^2}{b_i + \mathbf{E}_{r-1}(b'_i)} - \frac{(a_i+b_i)^2}{a_i + b_i + \mathbf{E}_{r-1}(a'_i) + \mathbf{E}_{r-1}(b'_i)}\right\} \tag{4}$$
$$= \frac{1}{r}\sum_{i=1}^{n}\frac{(a_i\mathbf{E}_{r-1}(b'_i) - b_i\mathbf{E}_{r-1}(a'_i))^2}{(a_i + \mathbf{E}_{r-1}(a'_i))(b_i + \mathbf{E}_{r-1}(b'_i))(a_i + b_i + \mathbf{E}_{r-1}(a'_i) + \mathbf{E}_{r-1}(b'_i))}$$

Now in case $(a'_i) \sim (b'_i)$ for each i then the first inequality in (4) is equality. To resolve this case suppose

$$(a'_i) = \lambda_i (b'_i)$$

for each i. Then

$$(a_i\mathbf{E}_{r-1}(b'_i) - b_i\mathbf{E}_{r-1}(a'_i))^2 = (a_i - \lambda_i b_i)^2\omega_i^2$$

where

$$\omega_i = \mathbf{E}_{r-1}(b'_i) \neq 0$$

since at least r of the (b) are positive. But on the other hand $a_i = \lambda_i b_i$ implies $(a) \sim (b)$ contrary to assumption.

Hence in this case the second inequality in (4) is strict and the induction is complete.

THEOREM 2. *If (a) and (b) are sets of n positive variables then for* $1 \leqslant r \leqslant n$

$$E_r^{1/r}(a+b) > E_r^{1/r}(a) + E_r^{1/r}(b) \tag{5}$$

unless $(a) \sim (b)$ or $r = 1$, in which cases we have equality.

To prove (5) we first establish a preliminary

LEMMA. *If $(a'_n) \sim (b'_n)$ and $(a) \nsim (b)$ then the inequality* (5) *holds.*

Proof.[1] Set

$$ta_i = b_i, \qquad i = 1, \ldots, n-1$$
$$(t+\delta)a_n = b_n, \quad \delta \neq 0.$$

[1]The authors are indebted to K. Hoechsmann for simplifying the original proof of this Lemma.

https://doi.org/10.4153/CJM-1957-037-9 Published online by Cambridge University Press

From

$$E_r(b) = E_r(b'_n) + b_n\, E_{r-1}(b'_n)$$

we conclude that

$$\begin{aligned} E_r(b) &= t^r\, E_r(a'_n) + (t+\delta)t^{r-1}\, a_n\, E_{r-1}(a'_n) \\ &= t^r\, E_r(a) + \delta\, t^{r-1}\, a_n\, E_{r-1}(a'_n). \end{aligned}$$

Set

$$g(\delta) = E_r^{1/r}\,(a+b) - E_r^{1/r}\,(a) - E_r^{1/r}\,(b).$$

$$g(\delta) = ((t+1)^r\, E_r(a) + \delta(t+1)^{r-1}\, a_n\, E_{r-1}(a'_n))^{1/r}$$
$$- E_r^{1/r}\,(a) - (t^r\, E_r(a) + \delta t^{r-1}\, a_n\, E_{r-1}(a'_n))^{1/r}.$$

Since $g(0) = 0$ it will suffice to establish

$$\delta\, g'\,(\delta) > 0$$

for $\delta \neq 0$, for then g strictly increases from $\delta = 0$. First note that

$$g'(\delta) = \frac{a_n\, E_{r-1}(a_n)\ (t+1)^{r-1}\, E_r^{(r-1)/r}\,(b) - t^{r-1}\, E_r^{(r-1)/r}\,(a+b)}{r(E_r(b)\, E_r(a+b))^{(r-1)/r}}.$$

The sign of this expression is the same as the sign of

$$\begin{aligned} (t+1)^r\, E_r(b) - t^r\, E_r(a+b) &= (t+1)^r\, t^r\, E_r(a) + \delta t^{r-1}\, a_n E_{r-1}(a_n) \\ &\quad - t^r\, (t+1)^r\, E_r(a) + \delta(t+1)^{r-1} a_n E_{r-1}(a_n) \\ &= t^{r-1}(t+1)^{r-1}\, \delta\, a_n\, E_{r-1}\,(a_n). \end{aligned}$$

Hence $g'(\delta)$ has the same sign as δ.

The proof of (5) will be done by a double induction on r and n. The case $r = 2$ and n arbitrary can easily be verified directly. We thus make the induction hypotheses that (5) holds for

(i) $r - 1$ and all $n \geqslant r - 1$,
(ii) r and $n - 1$.

Since we know $r = 2$ and $r = n$ (Hölder's inequality) it will suffice to prove (5) for r, n using (i) and (ii).

By the Lemma we can assume $(a'_n) \not\sim (b'_n)$ and also $(a) \not\sim (b)$. Set

$$x^2 = a_n, \quad y^2 = b_n, \quad x = R\cos\phi, \quad y = R\sin\phi.$$

Let $f(x, y) = g(R, \phi)$ be the left side of (5) minus the right side:

$$\begin{aligned} (6)\qquad g(R,\phi) &= (E_{r-1}(a'_n + b'_n)R^2 + E_r\,(a'_n + b'_n))^{1/r} \\ &\quad - (E_{r-1}(a'_n)R^2\cos^2\phi + E_r\,(a'_n))^{1/r} \\ &\quad - (E_{r-1}(b'_n)R^2\sin^2\phi + E_r\,(b'_n))^{1/r}. \end{aligned}$$

We show first that there exists a sufficiently large R_0 such that

$$g(R, \phi) > 0$$

https://doi.org/10.4153/CJM-1957-037-9 Published online by Cambridge University Press

for $0 \leqslant \phi \leqslant 2\pi$ and $R \geqslant R_0$. Set

$$C_j = E_j\,(a'_n + b'_n),$$
$$A_j = E_j\,(a'_n),$$
$$B_j = E_j\,(b'_n).$$

By factoring an R^2 from each bracketed quantity in (6) it is clear that

$$g(R, \phi) > 0$$

if and only if

$$\left(C_{r-1} + \frac{C_r}{R^2}\right)^{1/r} - \left(A_{r-1}\cos^2\phi + \frac{A_r}{R^2}\right)^{1/r} - \left(B_{r-1}\sin^2\phi + \frac{B_r}{R^2}\right)^{1/r} > 0.$$

At $\phi = 0$ and $\phi = 2\pi$ this is clear. By setting

$$h'\,(\phi) = 0$$

we easily calculate that at a minimum for $h(\phi)$

$$\cos^2\phi = A_{r-1}{}^{1/(r-1)}(A_{r-1}{}^{1/(r-1)} + B_{r-1}{}^{1/(r-1)})^{-1},$$
$$\sin^2\phi = B_{r-1}{}^{1/(r-1)}(A_{r-1}{}^{1/(r-1)} + B_{r-1}{}^{1/(r-1)})^{-1}.$$

Replacing these values in $h(\phi)$ and using (i) we see that

$$h\,(\phi) > 0$$

and hence

$$f(x, y) = g(R, \phi) > 0$$

for $R \geqslant R_0$. Thus if $f(x, y)$ is to have a non-positive minimum it must occur in the interior of the circle $R = R_0$ where the partials satisfy

$$f_x = 0 \qquad f_y = 0. \tag{7}$$

The equations (7) become

$$\begin{aligned} x\{((x^2 + y^2)C_{r-1} + C_r)^{(1-r)/r}\, C_{r-1} - (x^2A_{r-1} + A_r)^{(1-r)/r}A_{r-1}\} &= 0 \\ y\{((x^2 + y^2)C_{r-1} + C_r)^{(1-r)/r}\, C_{r-1} - (y^2B_{r-1} + B_r)^{(1-r)/r}\, B_{r-1}\} &= 0. \end{aligned} \tag{8}$$

If $x = y = 0$ we have the result by (ii). If both the curly bracketed quantities in (8) vanish then the result follows easily from (i) by direct substitution. The less trivial case is, say, $y = 0$ and $x \neq 0$. We must then show

$$f(x, 0) > 0$$

for the value x given by

$$x^2 = \frac{C_{r-1}^{1/(r-1)}\dfrac{A_r}{A_{r-1}} - A_{r-1}^{1/(r-1)}\dfrac{C_r}{C_{r-1}}}{A_{r-1}^{1/(r-1)} - C_{r-1}^{1/(r-1)}}.$$

Replacing this in $f(x, 0)$ we have

$f(x, 0) =$

$$\left(\frac{C_r C_{r-1}^{1/(r-1)} - C_{r-1}^{r/(r-1)} \dfrac{A_r}{A_{r-1}}}{C_{r-1}^{1/(r-1)} - A_{r-1}^{1/(r-1)}}\right)^{1/r} - \left(\frac{A_{r-1}^{r/(r-1)} \dfrac{C_r}{C_{r-1}} - A_r A_{r-1}^{1/(r-1)}}{C_{r-1}^{1/(r-1)} - A_{r-1}^{1/(r-1)}}\right)^{1/r} - B_r^{1/r}$$

$$= C_{r-1}^{1/(r-1)}\left(\frac{\dfrac{C_r}{C_{r-1}} - \dfrac{A_r}{A_{r-1}}}{C_{r-1}^{1/(r-1)} - A_{r-1}^{1/(r-1)}}\right)^{1/r} - A_{r-1}^{1/(r-1)}\left(\frac{\dfrac{C_r}{C_{r-1}} - \dfrac{A_r}{A_{r-1}}}{C_{r-1}^{1/(r-1)} - A_{r-1}^{1/(r-1)}}\right) - B_r^{1/r}.$$

At this point we use the inequality of Theorem 1. Then

$$f(x, 0) > \left(\frac{B_r}{B_{r-1}}\right)^{1/r} (C_{r-1}^{1/(r-1)} - A_{r-1}^{1/(r-1)})^{(r-1)/r} - B_r^{1/r} > 0.$$

The last inequality follows from (i). This completes the induction.

To indicate briefly that the concavity of $E_r^{1/r}(a)$ can be obtained from Fenchel's result we set K_i equal to the segment from 0 to 1 on the ith axis $(i = 1, \ldots, n)$ and we let Q be the unit cube in n-space. For $a = (a_1, \ldots, a_n)$, $a_i \geqslant 0$, form the convex sum

$$K_a = \sum_{i=1}^{n} a_i K_i.$$

For λ, μ non-negative numbers form the volume

$$V(\lambda K_a + \mu Q) = \sum_{j=0}^{n} b_j \lambda^{n-j} \mu^j.$$

The coefficients b_j are the elementary symmetric functions of (a) to within a constant factor and are called the *mixed volumes*. It is easy to check that the concavity result of Fenchel[2] reduces in this case to the statement that $E_r^{1/r}(a)$ is concave.

3. Applications to Matrices. The following result is known **(3; 4)**: *Let*

$$f(x_1, \ldots, x_k) = E_r\,((Ax_1, x_1), \ldots, (Ax_k, x_k)),$$

where A is a n.n.h. *matrix with characteristic roots* $0 \leqslant \alpha_i \leqslant \alpha_{i+1}$ $(i = 1, \ldots, n-1)$. *Then*

$$\max f = \binom{k}{r}\left(\sum_{j=1}^{k} \frac{\alpha_{n-j+1}}{k}\right)^r$$

$$\min f = E_r(\alpha_1, \ldots, \alpha_k)$$

where both max and min are taken over all sets of k o.n. *vectors in the unitary n-space.*[3]

[2]A discussion of mixed volumes and a conjecture of this general concavity result may be found in the book by T. Bonnesen and W. Fenchel, *Theorie der konvexen Körper* (Chelsea, New York, 1948), p. 93.

[3]The paper of Ostrowski contains only the value of min f.

https://doi.org/10.4153/CJM-1957-037-9 Published online by Cambridge University Press

THEOREM 3. *Let* $A, B, A + B$ *be* n.n.h. *with characteristic roots* $0 \leqslant \alpha_i \leqslant \alpha_{i+1}$, $0 \leqslant \beta_i \leqslant \beta_{i+1}$, $0 \leqslant \lambda_i \leqslant \lambda_{i+1}$ *for* $i = 1, \ldots, n - 1$, *respectively. Then for* $1 \leqslant r \leqslant k \leqslant n$

$$E_r^{1/r}(\lambda_1, \ldots, \lambda_k) \geqslant E_r^{1/r}(\alpha_1, \ldots, \alpha_k) + E_r^{1/r}(\beta_1, \ldots, \beta_k).$$

Proof. Let $x_1, \ldots, x_k$ be an o.n. set of characteristic vectors of $A + B$ such that

$$c_j = ((A + B)\, x_j, x_j) = \lambda_j \qquad (j = 1, \ldots, k).$$

Set

$$(Ax_j, x_j) = a_j, \quad (Bx_j, x_j) = b_j.$$

Then

$$E_r^{1/r}(\lambda) = E_r^{1/r}(c) = E_r^{1/r}(a + b)$$
$$\geqslant E_r^{1/r}(a) + E_r^{1/r}(b) \geqslant E_r^{1/r}(\alpha) + E_r^{1/r}(\beta).$$

Let

$$x^n + \sum_{j=1}^{n} \rho_j(A)\, x^{n-j}$$

be the characteristic polynomial of A. By setting $k = n$ in Theorem 3 we obtain the following extension of the Minkowski inequality to all the coefficients in the characteristic polynomial:

THEOREM 4. *If* A *and* B *are* n.n.h. *then for* $1 \leqslant r \leqslant n$

$$|\rho_r\, (A + B)|^{1/r} \geqslant |\rho_r\, (A)|^{1/r} + |\rho_r\, (B)|^{1/r}.$$

The extension of Fan's result is contained in

THEOREM 5. *If* A *and* B *are* n.n.h. *and* $\omega + \sigma = 1$, $\omega \geqslant 0$, $\sigma \geqslant 0$, *then for* $1 \leqslant r \leqslant n$

$$|\rho_r(\omega A + \sigma B)| \geqslant |\rho_r(A)|^{\omega}\, |\rho_r(B)|^{\sigma}.$$

Proof. As in the proof of Theorem 3

$$E_r^{1/r}(\lambda) = E_r^{1/r}(\omega(a) + \sigma(b)) \geqslant \omega E_r^{1/r}(a) + \sigma E_r^{1/r}(b)$$
$$\geqslant \omega E_r^{1/r}(\alpha) + \sigma E_r^{1/r}(\beta) \geqslant E_r^{\omega/r}(\alpha)\, E_r^{\sigma/r}(\beta).$$

The result follows by taking rth powers.

THEOREM 6. *Under the same hypotheses as Theorem* 3,

$$k^{1-r}\binom{k}{r-1} \mathbf{E}_r(\lambda_1, \ldots, \lambda_k)$$
$$\geqslant E_r(\alpha_1, \ldots, \alpha_k)\left(\sum_{j=1}^{k} \alpha_{n-j+1}\right)^{1-r} \tag{9}$$
$$+ E_r(\beta_1, \ldots, \beta_k)\left(\sum_{j=1}^{k} \beta_{n-j+1}\right)^{1-r}.$$

Proof. We again choose $x_1, \ldots, x_k$ to be k o.n. characteristic vectors of $A + B$ corresponding respectively to the characteristic roots $\lambda_1, \ldots, \lambda_k$. Then

$$((A + B)\, x_j, x_j) = \lambda_j, \qquad j = 1, \ldots, k$$

and we have

$$\begin{aligned} \mathbf{E}_r(\lambda_1, \ldots, \lambda_k) &= \mathbf{E}_r(((A+B)x, x)) = \mathbf{E}_r((Ax, x) + (Bx, x)) \\ &\geqslant \mathbf{E}_r((Ax, x)) + \mathbf{E}_r((Bx, x)) \\ &\geqslant E_r(\alpha)\binom{k}{r-1}^{-1}\left(\sum_{j=1}^{k}\frac{\alpha_{n-j+1}}{k}\right)^{1-r} \\ &\quad + E_r(\beta)\binom{k}{r-1}^{-1}\left(\sum_{j=1}^{k}\frac{\beta_{n-j+1}}{k}\right)^{1-r}, \end{aligned}$$

and (9) follows.

COROLLARY 1. *Under the same hypotheses as Theorem* 4

$$n^{1-r}\binom{n}{r-1}\left|\frac{\rho_r(A+B)}{\rho_{r-1}(A+B)}\right| \geqslant \frac{|\rho_r(A)|}{(\text{tr}\, A)^{r-1}} + \frac{|\rho_r(B)|}{(\text{tr}\, B)^{r-1}}.$$

Proof. Set $k = n$ in (9). We immediately conclude from Corollary 1 that

$$\det\,(A + B) \geqslant \text{tr}\,(A + B)\left(n^{-\frac{1}{2}n(n-1)}\prod_{r=2}^{n}\binom{n}{r-1}^{-1}\sigma_r\right)$$

where

$$\sigma_r = \frac{|\rho_r(A)|}{(\text{tr}\, A)^{r-1}} + \frac{|\rho_r(B)|}{(\text{tr}\, B)^{r-1}}.$$

REFERENCES

1. K. Fan, *On a Theorem of Weyl concerning eigenvalues of linear transformations* II, Proc. Nat. Acad. Sci., *36* (1950), 31.
2. W. Fenchel, *Généralisation du théorème de Brunn-Minkowski concernant les corps convexes*, C.R. des Sci. de l'Acad. des Sci., *203* (1936), 764–766.
3. M. Marcus and J. L. McGregor, *Extremal properties of Hermitian matrices*, Can. J. Math., *8* (1956), 524–531.
4. A. Ostrowski, *Sur quelques applications des fonctions convexes et concaves au sens de I. Schur*, J. Math. pures et appl. (9), *31* (1952), 253–292.

University of British Columbia

United States Naval Ordnance Test Station Pasadena, Cal.