UNIVERSITY OF NORTH CAROLINA

Department of Statistics

Chapel Hill, N. C.

A NOTE ON THE INEQUALITY OF THE CHARACTERISTIC ROOTS OF THE PRODUCT AND THE SUM OF TWO SYMMETRIC MATRICES

by

C. G. Khatri

April 1964

Contract No. AF-AFOSR-84-63

This research was supported by the Mathematics Division of the Air Force Office of Scientific Research.

Institute of Statistics
Mimeo Series No. 387

A NOTE ON THE INEQUALITY OF THE CHARACTERISTIC ROOTS OF THE PRODUCT AND THE SUM OF TWO SYMMETRIC MATRICES[1]

by

C. G. Khatri
University of North Carolina and Gujarat University

= =

By a symmetric (sy.) matrix, we shall mean here a symmetric matrix with real elements only. We establish the bounds on the i-th maximum characteristic (ch.) root (denoted by ch_i) of (i) $\underset{\sim}{A}\underset{\sim}{B}$ where $\underset{\sim}{A}$ is any sy. matrix and $\underset{\sim}{B}$ is any sy. positive semi-definite (p.s.d) matrix, and (ii) of $\underset{\sim}{A} + \underset{\sim}{B}$ where $\underset{\sim}{A}$ and $\underset{\sim}{B}$ are both symmetric matrices. Anderson and Gupta [1], using somewhat different methods, established the bounds for the case (i) when $\underset{\sim}{A}$ was sy. p.s.d. and $\underset{\sim}{B}$ was sy. p.d. All the results established here are valid for hermitian matrices (with possibly complex elements) too.

We require the following three lemmas, which are given by Bellman [2].

<u>Lemma 1</u>: Let $\underset{\sim}{A}$ be any sy. matrix of order p and let $\underline{x}$, $\underline{\alpha}_j$ ($j = 1, 2, \ldots, i-1$) be column vectors with p elements, real and finite. Then

$$ch_i \underset{\sim}{A} = \underset{\substack{\underline{\alpha}_j \\ j = 1, 2, \ldots, i-1}}{\text{minimum}} \quad \underset{\underline{x}'\underline{\alpha}_j = 0}{\text{maximum}} \left(\frac{\underline{x}' \underset{\sim}{A} \underline{x}}{\underline{x}' \underline{x}} \right) .$$

Note that we replace the condition of $\underline{\alpha}_j' \underline{\alpha}_j = 1$ ($j = 1, 2, \ldots, i-1$) given by Bellman [2] for Courant-Fischer min-max. theorem by any $\underline{\alpha}_j$ ($j = 1, 2, \ldots, i-1$) having real and finite elements. It is easy to see that by doing this, we are not changing the value of $ch_i \underset{\sim}{A}$.

---

[1]This research was supported by the Mathematics Division of the Air Force Office of Scientific Research.

Lemma 2: Let $A_r = (a_{ij})$, $i, j = 1, 2, \ldots, r$ and $r = 1, 2, \ldots, p$ be a sequence of sy. matrices. Then

$$ch_{k+1}(A_{r+1}) \leq ch_k(A_r) \leq ch_k(A_{r+1}) .$$

Lemma 3: Let A and B be two matrices of respective order pxn and nxp with $n \geq p$. Then

$$ch_i\, AB = ch_i\, BA \qquad \text{for all positive ch. roots and}$$

$$ch_i\, AB = ch_{n-p+i}\, BA \qquad \text{for all negative ch. roots.}$$

Theorem 1: Let A be a sy. matrix of order p and B, a sy. matrix of order p, be p.s.d. of rank r ($\leq p$). Let a and b denote respectively the number of positive and the number of negative ch. roots of AB. Then for any $i, j = 1,2,\ldots, a$

$$ch_j B\; ch_{p+i-j} A \leq ch_i\, AB \leq ch_j\, B\; ch_{i-j+1}\, A$$

and for $i = p - b + 1, p - b + 2, \ldots, p$ and $j = 1, 2, \ldots, r$,

$$ch_j B\; ch_{r+i-j} A \leq ch_i\, AB \leq ch_j B\; ch_{r+i-j-(p-1)} A .$$

Proof: Since B is sy. p.s.d., we can find an orthogonal matrix Δ such that $B = \Delta D_1 \Delta'$, where D_1 = diagonal $(\omega_1, \ldots, \omega_r, 0, \ldots, 0)$, D_ω = diag. $(\omega_1, \ldots, \omega_r)$ and $\omega_1 \geq \omega_2 \geq \ldots \geq \omega_r > 0$ are the nonzero ch. roots of B. Let S be a sub-matrix of order r of $(\Delta A \Delta')$, obtained by deleting the last (p-r) rows and (p-r) columns of $\Delta A \Delta'$. Then by lemma 3, it is easy to see that

(1) $$ch_i AB = ch_i SD_\omega = ch_i (D_\omega^{\frac{1}{2}}\, S\, D_\omega^{\frac{1}{2}}) \qquad \text{for } i = 1, 2, \ldots, a \text{ and}$$

(2) $$ch_{p-r+i} AB = ch_i (SD_\omega) = ch_i (D_\omega^{\frac{1}{2}} S\, D_\omega^{\frac{1}{2}}) \qquad \text{for } i = r-b+1, r-b+2, \ldots, r ,$$

where $(D_\omega^{\frac{1}{2}})^2 = D_\omega$. Then by lemma 1,

(3) $$\operatorname{ch}_i(\underset{\sim}{S}\underset{\sim}{D}_\omega) = \underset{\substack{\underline{\alpha}_h \\ h=1,2,\ldots,i-1}}{\text{minimum}} \quad \underset{\underline{x}'\,\underline{\alpha}_h=0}{\text{maximum}} \left(\frac{\underline{x}'(\underset{\sim}{D}_{\sqrt{\omega}}\, \underset{\sim}{S}\, \underset{\sim}{D}_{\sqrt{\omega}})\, \underline{x}}{\underline{x}'\underline{x}} \right)$$

where $\underline{x}$ and $\underline{\alpha}_h$ $(h = 1,2,\ldots,i-1)$ are column vectors with r elements, real and finite. Now minimize over a sub-space given by

$\underline{\alpha}_{i-j+h} = (0,\ldots,0,\ c_h, 0,\ldots,0)$, c_h is nonzero finite and real at the h-th place, for $h = 1,2,\ldots,j-1$, and $\underline{\alpha}_h = (0,\ldots,0,\ \underline{\alpha}_{j,h},\ldots,\ \underline{\alpha}_{r,h})$ for $h = 1,2,\ldots,i-j$. Then $\underline{x}'\ \underline{\alpha}_{i-j+h} = 0$ for $h = 1,2,\ldots,j-1$ give $x_1 = x_2 = \ldots = x_{j-1} = 0$. Now let us write $\underline{y}' = (\sqrt{\omega_j}\ x_j,\ldots,\ \sqrt{\omega_r}\ x_r)$, $\underline{\delta}_h = (\omega_j^{-\frac{1}{2}}\alpha_{j,h},\ldots,\ \omega_r^{-\frac{1}{2}}\alpha_{r,h})$ for $h = 1, 2, \ldots, i-j$, $\underset{\sim}{S}_1$ be a sub-matrix of S, obtained by deleting the first $(j-1)$ rows and $(j-1)$ columns of $\underset{\sim}{S}$, and $\underset{\sim}{D}_2 = \text{diag.}\ (\omega_j,\ \omega_{j+1},\ldots,\ \omega_r)$. Note that the elements of $\underline{y}$ and $\underline{\delta}_h$ $(h = 1, 2, \ldots, i-j)$ are finite and real, and the minimum over a sub-set must always be greater than or equal to the minimum over the larger set. Hence (3) gives

(4) $$\operatorname{ch}_i(\underset{\sim}{S}\underset{\sim}{D}_\omega) \leq \underset{\substack{\underline{\delta}_h \\ h=1,2,\ldots,i-j}}{\text{minimum}} \quad \underset{\underline{y}'\underline{\delta}_h=0}{\text{maximum}} \left(\frac{\underline{y}'\ \underset{\sim}{S}_1\ \underline{y}}{\underline{y}'\ \underset{\sim}{D}_2^{-1}\ \underline{y}} \right).$$

Since $\underline{y}'\ \underset{\sim}{D}_2^{-1}\ \underline{y} \geq \omega_j^{-1}\ \underline{y}'\underline{y}$ or $(\underline{y}'\ \underset{\sim}{D}_2^{-1}\ \underline{y})^{-1} \leq \omega_j(\underline{y}'\underline{y})^{-1}$, (4) gives

(5) $$\operatorname{ch}_i(\underset{\sim}{S}\underset{\sim}{D}_\omega) \leq \underset{\substack{\underline{\delta}_h \\ h=1,2,\ldots,i-j}}{\text{minimum}} \quad \underset{\underline{y}'\ \underline{\delta}_h=0}{\text{maximum}} \left(\frac{\underline{y}'\ \underset{\sim}{S}_1\ \underline{y}}{\underline{y}'\ \underline{y}} \right)\omega_j = (\operatorname{ch}_{i-j+1}\underset{\sim}{S}_1)\omega_j .$$

Using the lemma 2, and $\omega_j = \operatorname{ch}_j \underset{\sim}{D}_\omega$, (5) gives

(6) $$\operatorname{ch}_i(\underset{\sim}{S}\underset{\sim}{D}_\omega) \leq (\operatorname{ch}_{i-j+1}\underset{\sim}{S})\ (\operatorname{ch}_j \underset{\sim}{D}_\omega) .$$

Now consider

$$\text{ch}_{r+i-j}\underset{\sim}{S} = \text{ch}_{r+i-j}[(\underset{\sim}{S}\underset{\sim}{D}_\omega)(\underset{\sim}{D}_\omega^{-1})] \leq (\text{ch}_i\ \underset{\sim}{S}\underset{\sim}{D}_\omega)(\text{ch}_{r-j+1}\ \underset{\sim}{D}_\omega^{-1})$$

and $\text{ch}_{r-j+1}\ \underset{\sim}{D}_\omega^{-1} = (\text{ch}_j\ \underset{\sim}{D}_\omega)^{-1}$. Hence, we get

(7) $$\text{ch}_i\ \underset{\sim}{S}\underset{\sim}{D}_\omega \geq (\text{ch}_j\ \underset{\sim}{D}_\omega)\ (\text{ch}_{r+i-j}\underset{\sim}{S})$$

Combining (6) and (7), and using lemma 2 for $\underset{\sim}{S}$ matrix, we get

(8) $$(\text{ch}_j\underset{\sim}{B})\ (\text{ch}_{p+i-j}\ \underset{\sim}{A}) \leq \text{ch}_i\ \underset{\sim}{S}\underset{\sim}{D}_\omega \leq (\text{ch}_j\underset{\sim}{B})\ (\text{ch}_{i-j+1}\underset{\sim}{A})$$

for $i, j = 1, 2, \ldots, r$.

Now, the use of (1) and (2) in (8) proves the theorem 1.

<u>Corollary 1:</u> Let $\underset{\sim}{A}$ be a sy. matrix, $\underset{\sim}{B}$ be sy. p.s.d. of rank r and $\underset{\sim}{C}$ be sy. p.d. Then for any $i, j = 1, 2, \ldots, a$ (= number of positive ch. roots of $\underset{\sim}{A}\underset{\sim}{B}$) ,

$$(\text{ch}_j\ \underset{\sim}{B}\underset{\sim}{C})(\text{ch}_{p+i-j}\ \underset{\sim}{A}\ \underset{\sim}{C}^{-1}) \leq \text{ch}_i\ \underset{\sim}{A}\underset{\sim}{B} \leq (\text{ch}_j\ \underset{\sim}{B}\underset{\sim}{C})(\text{ch}_{i-j+1}\ \underset{\sim}{A}\underset{\sim}{C}^{-1})$$

and for $j = 1, 2, \ldots, r$ and $i = p - b + 1, p - b + 2, \ldots, p$, (b(= number of negative ch. roots of $\underset{\sim}{A}\underset{\sim}{B}$) ,

$$(\text{ch}_j\ \underset{\sim}{B}\underset{\sim}{C})\ (\text{ch}_{r+i-j}\underset{\sim}{A}\underset{\sim}{C}^{-1}) \leq \text{ch}_i\ \underset{\sim}{A}\underset{\sim}{B} \leq (\text{ch}_j\ \underset{\sim}{B}\underset{\sim}{C})(\text{ch}_{r+i-j-(p-1)}\underset{\sim}{A}\underset{\sim}{C}^{-1}) .$$

<u>Proof:</u> Since $\underset{\sim}{C}$ is, sy. p.d., we have $\underset{\sim}{C} = \underset{\sim}{T}\underset{\sim}{T}'$ where $\underset{\sim}{T}$ is a non-singular matrix. Then by lemma 3,

$\text{ch}_i\underset{\sim}{A}\underset{\sim}{B} = \text{ch}_i(\underset{\sim}{A}\underset{\sim}{T}'^{-1}\ \underset{\sim}{T}'\ \underset{\sim}{B}\underset{\sim}{T}\underset{\sim}{T}^{-1}) = \text{ch}_i[(\underset{\sim}{T}^{-1}\underset{\sim}{A}\ \underset{\sim}{T}'^{-1})(\underset{\sim}{T}'\underset{\sim}{B}\underset{\sim}{T})]$ and then by theorem 1, we get the corollary 1.

<u>Corollary 2:</u> Let $r = p$ in corollary 1. Then, we get for any $i, j = 1, 2, \ldots, p$,

$$(\text{ch}_j\underset{\sim}{B}\underset{\sim}{C})\ (\text{ch}_{p+i-j}\underset{\sim}{A}\underset{\sim}{C}^{-1}) \leq \text{ch}_i\ \underset{\sim}{A}\underset{\sim}{B} \leq (\text{ch}_j\ \underset{\sim}{B}\underset{\sim}{C})\ (\text{ch}_{i-j+1}\ \underset{\sim}{A}\underset{\sim}{C}^{-1}) .$$

The above result was established by Anderson and Gupta [1] by using somewhat different methods under the condition of $\underset{\sim}{A}$ being p.s.d., but we have proved it for $\underset{\sim}{A}$ any sy. matrix.

Corollary 3: As a special case of corollary 2, we get for any $i = 1, 2, \ldots, p$

$$\max.[(ch_p \underset{\sim}{B}\underset{\sim}{C})(ch_i \underset{\sim}{A}\underset{\sim}{C}^{-1}),\ (ch_i \underset{\sim}{B}\underset{\sim}{C})(ch_p \underset{\sim}{A}\underset{\sim}{C}^{-1})] \leq ch_i(\underset{\sim}{A}\underset{\sim}{B})$$
$$\leq \min\ [(ch_1 \underset{\sim}{B}\underset{\sim}{C})(ch_i \underset{\sim}{A}\underset{\sim}{C}^{-1}),\ (ch_i \underset{\sim}{B}\underset{\sim}{C})(ch_1 \underset{\sim}{A}\underset{\sim}{C}^{-1})].$$

Theorem 2: Let A and B be any two sy. matrices of order p. Then for any $i, j = 1, 2, \ldots, p$,

$$ch_j \underset{\sim}{B} + ch_{p+i-j} \underset{\sim}{A} \leq ch_i\ (\underset{\sim}{A} + \underset{\sim}{B}) \leq ch_j \underset{\sim}{B} + ch_{i-j+1} \underset{\sim}{A} \ .$$

Proof: Since $\underset{\sim}{B}$ is a sy. matrix of order p, there exists an orthogonal matrix $\underset{\sim}{\Delta}$ such that $\underset{\sim}{B} = \underset{\sim}{\Delta}\, \underset{\sim}{D}_\omega\, \underset{\sim}{\Delta}'$ where $\underset{\sim}{D}_\omega = \text{diag.}\ (\omega_1, \omega_2, \ldots, \omega_p)$ and $\omega_1 \geq \omega_2 \geq \cdots \geq \omega_p$ are the ch. roots of $\underset{\sim}{B}$. Moreover, by lemma 1, we have

$$ch_i(\underset{\sim}{A}+\underset{\sim}{B}) = \underset{\substack{\underset{\sim}{\alpha}_h \\ h = 1,2,\ldots,i-1}}{\text{minimum}} \quad \underset{\underline{x}'\underset{\sim}{\alpha}_h = 0}{\text{maximum}} \left[\left(\frac{\underline{x}'(A+B)\underline{x}}{\underline{x}'\underline{x}} \right) = \frac{\underline{x}'A\ \underline{x}}{\underline{x}'\underline{x}} + \frac{\underline{x}'B\ \underline{x}}{\underline{x}'\underline{x}} \right].$$

Let us write $\underset{\sim}{\Delta}'\underline{x} = \underline{y}$ and $\underset{\sim}{\Delta}'\underset{\sim}{\alpha}_h = \underset{\sim}{\beta}_h$ $(h = 1, 2, \ldots, i-1)$. Then

$$(9) \qquad ch_i(\underset{\sim}{A}+\underset{\sim}{B}) = \underset{\substack{\underset{\sim}{\beta}_h \\ h = 1,2,\ldots,i-1}}{\text{minimum}} \quad \underset{\underset{\sim}{\alpha}'\underset{\sim}{\beta}_h = 0}{\text{maximum}} \left[\frac{\underline{y}'(\underset{\sim}{\Delta}'\underset{\sim}{A}\underset{\sim}{\Delta})\underline{y}}{\underline{y}'\underline{y}} + \frac{\underline{y}'\ \underset{\sim}{D}_\omega\ \underline{y}}{\underline{y}'\underline{y}} \right].$$

Now minimize over the sub-region given by

$\underset{\sim}{\beta}'_{h+i-j} = (0,\ldots,0,\ c_h,\ 0,\ldots,0)$, c_h is nonzero finite and real, for $r = 1, 2, \ldots, j-1$, and $\underset{\sim}{\beta}'_h = (0,\ldots,0,\ \beta_{j,h},\ \ldots,\ \beta_{r,h})$ for $h = 1,2,\ldots,i-j$. Then $\underline{y}'\ \underset{\sim}{\beta}_{h+i-j} = 0$ $(h = 1,2,\ldots,j-1)$ give $y_1 = y_2 = \cdots = y_{j-1} = 0$. Now let us write $\underline{Z} = (y_j, \ldots, y_r)$, $\underline{\delta}_h = (\beta_{j,h}, \ldots, \beta_{r,h})(h = 1,2,\ldots,i-j)$, $\underset{\sim}{D}_3$ = diag. $(\omega_j, \omega_{j+1}, \ldots, \omega_p)$ and $\underset{\sim}{A}_1$ be the sub-matrix of $\underset{\sim}{A}$, obtained by deleting first (j-1) rows and (j-1) columns of $(\underset{\sim}{\Delta}'\underset{\sim}{A}\underset{\sim}{\Delta})$. With t his, it is clear that

$$\mathrm{ch}_i(\underset{\sim}{A}+\underset{\sim}{B}) \leq \underset{\substack{\underline{\delta}_h \\ h = 1,2,\ldots,i-j}}{\text{minimum}} \quad \underset{\underline{Z}'\underline{\delta}_h=0}{\text{maximum}} \left[\frac{\underline{Z}' \underset{\sim}{A}_1 \underline{Z}}{\underline{Z}'\underline{Z}} + \frac{\underline{Z}' \underset{\sim}{D}_3 \underline{Z}}{\underset{\sim}{Z}'\underset{\sim}{Z}} \right]$$

$$\leq \underset{\substack{\underline{\delta}_h \\ h = 1,2,\ldots,i-j}}{\text{minimum}} \quad \underset{\underline{Z}'\underline{\delta}_h=0}{\text{maximum}} \left[\frac{\underline{Z}' \underset{\sim}{A}_1 \underline{Z}}{\underline{Z}'\underline{Z}} + \omega_j \right] = (\mathrm{ch}_{i-j+1}\, \underset{\sim}{A}_1) + \omega_j$$

That is, using lemma 2,

$$\mathrm{ch}_i(\underset{\sim}{A}+\underset{\sim}{B}) \leq \mathrm{ch}_j\underset{\sim}{B} + \mathrm{ch}_{i-j+1}\underset{\sim}{A} .$$

For the other part, $\mathrm{ch}_j\underset{\sim}{B} = \mathrm{ch}_j[(\underset{\sim}{A}+\underset{\sim}{B}) + (-\underset{\sim}{A})] \leq \mathrm{ch}_i(\underset{\sim}{A}+\underset{\sim}{B}) + \mathrm{ch}_{j-i+1}(-\underset{\sim}{A})$.

But $\mathrm{ch}_{j-i+1}(-\underset{\sim}{A}) = -\mathrm{ch}_{p+i-j}\underset{\sim}{A}$. Hence we get

$$\mathrm{ch}_i(\underset{\sim}{A}+\underset{\sim}{B}) \geq \mathrm{ch}_j\underset{\sim}{B} + \mathrm{ch}_{p+i-j}\underset{\sim}{A} .$$

Thus, the theorem 2 is established.

<u>Corollary 4</u>: As a special case of theorem 2, we have for any $i = 1,2,\ldots,p$

$$\text{max.}\ [\mathrm{ch}_i\underset{\sim}{B} + \mathrm{ch}_p\underset{\sim}{A} ,\ \mathrm{ch}_p\underset{\sim}{B} + \mathrm{ch}_i\underset{\sim}{A}] \leq \mathrm{ch}_i(\underset{\sim}{A}+\underset{\sim}{B}) \leq \text{min.}\ [\mathrm{ch}_1\underset{\sim}{B}+\mathrm{ch}_i\underset{\sim}{A}, \mathrm{ch}_i\underset{\sim}{B}+\mathrm{ch}_1\underset{\sim}{A}].$$

The author thanks Professor S. N. Roy for his kind help.

REFERENCES

[1] Anderson, T. W. and S. Das Gupta (1963), Some inequalities on characteristic roots of matrices, <u>Biometrika</u>, Vol. 50, pp. 522-524.

[2] Bellman, (1960), <u>Introduction to Matrix Analysis</u>, McGraw Hill Book Co., New York.