Two Matrix Eigenvalue Inequalities

Seymour Haber

(Feburary 15, 1962)

A lower bound is given for the quantity \( \lambda_1/\lambda_n \), and an upper bound for the quantity \( \lambda_i - \lambda_n \), where \( \lambda_1 \) and \( \lambda_n \) are respectively the greatest and least characteristic roots of a matrix with positive roots. The bounds involve the first and second coefficients of the characteristic equation of the matrix.

Suppose \( A = (a_{ij}) \) is a nonsingular \( n \times n \) matrix with characteristic roots \( \lambda_1, \lambda_2, \ldots, \lambda_n \), so ordered that \( |\lambda_1| \geq |\lambda_2| \geq \ldots \geq |\lambda_n| \). The quantity \( |\lambda_1|/|\lambda_n| \) provides a rough measure of the probable error in the computation of the inverse of \( A \); it has been called by J. Todd [1, 2, 3] the \( P \)-condition number of \( A \) and may be denoted \( P(A) \). Von Neumann and Goldstine [4] have shown that if \( A \) is symmetric and positive definite (in which case the \( \lambda_i \) are all positive), then the error in the inverse of \( A \) computed by a certain elimination method, can be bounded by a quantity proportional to \( P(A) \); if \( A \) is not symmetric positive definite, the error can be bounded by a quantity proportional to \( P(\bar{A} \bar{A}') \).

We shall restrict our consideration to matrices whose roots are all positive. For these, P. J. Davis, E. V. Haynsworth, and M. Marcus [5] obtained bounds on \( P \) involving \( \det A \) and one other symmetric function of the roots of \( A \). If the characteristic polynomial of \( A \) is \( p(x) = x^n - C_1 x^{n-1} + C_2 x^{n-2} + \ldots + (-1)^n C_n \) and we set \( D_i = (n^i C_n)/(C_i) \), they showed that

\[
\frac{1}{1} \leq P \leq 1 + \sqrt{1 - D_1} / 1 - \sqrt{1 - D_1},
\]

and they found similar inequalities involving \( C_n (= \det A) \) and any other one of the \( C_i \).

In many cases, however, \( C_n \) is not known and it is in general difficult to calculate. \( C_i (= \text{trace} A) \) is easy to calculate, and \( C_2 \) can in general be calculated more easily than can \( C_n \) since it is the sum of \( [n(n-1)]/2 \) determinants of order two. In this paper we present (in Theorems 1 and 1') a lower bound for \( P \) in terms of \( C_1 \) and \( C_2 \); an attempt to obtain a corresponding upper bound fails, but leads to an inequality (Theorem 2) on \( \lambda_1 - \lambda_n \), the “spread” of the roots of \( A \). Finally we apply the method of proof of Theorem 1 to obtain an improvement of the lower bound in (1).

If we set \( C_K = (K)S_K(\lambda_1, \lambda_2, \ldots, \lambda_n) \), then \( S_K \) is the \( K \)th symmetric mean of \( \lambda_1, \ldots, \lambda_n \). The \( S_K \) satisfy [6] the inequalities

\[
S_1 \geq S_2^{1/2} \geq S_3^{1/3} \geq \ldots \geq S_n^{1/n}. \tag{2}
\]

Setting \( \mu_K = \lambda_K/\lambda_n \) we have \( P = \mu_1 \geq \mu_2 \geq \mu_3 \geq \ldots \geq \mu_n = 1 \), and

\[
S^n_1(\mu_1, \mu_2, \ldots, \mu_n) \geq S_2(\lambda_1, \ldots, \lambda_n) \tag{2}
\]

Let

\[
R(x_2, x_3, \ldots, x_{n-1}) = \frac{S_1^n(P, x_2, x_3, \ldots, x_{n-1})}{S_2^n(P, x_2, x_3, \ldots, x_{n-1})}
\]

and

\[
f_1(P) = \max_{1 \leq x_i \leq P} R(x_2, x_3, \ldots, x_{n-1})
\]

and

\[
f_2(P) = \min_{1 \leq x_i \leq P} R(x_2, x_3, \ldots, x_{n-1}).
\]

Then \( f_1 \) and \( f_2 \) can be seen to be increasing in \( P \), and

\[
f_2(P) \leq \frac{S_2^n}{S_2} \leq f_1(P). \tag{3}
\]

Thus the right half of (3) should provide a lower bound for \( P \), while the left half should provide an upper bound. We first calculate \( f_2(P) \).

By direct calculation we can show that \( (\partial R)/\partial x_i \) is nonnegative at all points for each \( i \). Therefore \( R \) attains its maximum at a point where each \( x_i \) is either 1 or \( P \). Letting \( R_K(P) \) denote the value of \( R \) when \( K-1 \) of \( x_2, x_3, \ldots, x_{n-1} \) are equal to \( P \) and the remainder are equal to one, we find that \( R_K \) is equal to:

\[
n-1 \frac{(KP+n-K)^2}{n(KP+n-K)^2-(KP^2+n-K)}. \tag{4}
\]
This rational function of $K$ attains its maximum at $K = \frac{n}{P+1}$ and so we obtain

$$\frac{S_1}{S_2} \leq \frac{n-1}{n} - \frac{1}{4Pn}$$

(5)

which is equivalent to:

Theorem 1:

$$\frac{1+\sqrt{1-A}}{1-\sqrt{1-A}} \leq P, \text{where } A = \frac{1}{n - \frac{S_1}{S_2} (n-1)}.$$

The upper bound (5) can be sharpened, since in fact we need only consider integer values of $K$ in (4). It can be shown that $R_1 = \max R_K$ if $P^2 \geq \frac{1}{[(n-1)(n-2)]/2}$. Thus if $P^2 \geq \frac{1}{[(n-1)(n-2)]/2}$, $S_1^2/S_2 \leq R_1(P)$; and so, setting $\rho = S_1^2/S_2$, we have:

$$[(\rho-1) + \sqrt{\rho(\rho-1)}]n+1 \leq P.$$

(6)

Now if $P^2 \geq \frac{1}{[(n-1)(n-2)]/2}$, then $\rho \leq \max R_K(P) = R_K^\ast(P)$, say; and since each $R_K(P)$ is a strictly increasing function of $P$, $\rho \leq R_K^\ast([\frac{1}{(n-1)(n-2)]/2}) \leq R_1([\frac{1}{(n-1)(n-2)]/2})$ and so $\sqrt{\frac{(n-1)(n-2)}{2}} > R_1^{-1}(\rho)$. However it can easily be seen that if $\rho \geq \frac{(\sqrt{2} + \frac{3}{2})}{(\sqrt{2}+1)}$, then $R_1^{-1}(\rho) \geq \frac{(n-1)(n-2)}{2}$; and so we conclude that:

Theorem 1': If $\rho \geq \frac{(\sqrt{2} + \frac{3}{2})}{(\sqrt{2}+1)}$, then $[(\rho-1) + \sqrt{\rho(\rho-1)}]n+1 \leq P$.

This lower bound is better than the previous one, whenever it applies.

The attempt to derive an upper bound for $P$ from the left half of (3) fails, because $f_2(P)$ approaches 1 uniformly in $P$ as $n$ increases, and so the inequality $f_2(P) \leq S_1^2/S_2$ will in most cases hold for all values of $P$. We can, however, by considering the minimum of $S_1^2/(\lambda_1, y_2, y_3, \ldots, y_{n-1}, \lambda_n) - S_2^2/(\lambda_1, y_2, y_3, \ldots, y_{n-1}, \lambda_n)$ subject to the condition $\lambda_1 \geq y_2 \geq y_3 \geq \ldots \geq y_{n-1} \geq \lambda_n$, obtain an upper bound for the spread of the roots of $A$. Calling the above function of $y_2, \ldots, y_{n-1}, D$, we evaluate its minimum directly by observing that $|D/\partial y_i| = -2[(n(n-1))/(S_1-y_i)]$ and $|D/\partial y_j| \geq -2[(n^2(n-1))/(n-1)]$ if $i \neq j; i, j = 2, 3, \ldots, n-1$. The $(n-2) \times (n-2)$ matrix $(d_{ij})$, with $d_{ij} = \partial D/\partial y_i \partial y_j$ is symmetric, and by Gerschgorin's theorem each of its eigenvalues lies in the circle $|Z-2/n^2| \leq 2/n^2$ and so is positive. Thus $(d_{ij})$ is positive definite, and, setting each $\partial D/\partial y_i$ equal to zero, we find that $D$ is at a minimum when each $y_i$ is equal to $(\lambda_1 + \lambda_n)/2$, and that the minimal value is $|(\lambda_1 - \lambda_n)/2|/2n(n-1)]$. We may then conclude:

Theorem 2: $\lambda_1 - \lambda_n \leq \sqrt{2n(n-1)(S_1^2 - S_2^2)}$.

By the method of Theorem 1, it is possible to sharpen the left side of inequality (1). Following the notation of [5] we write

$$D_1(x_1, x_2, \ldots, x_n) = \frac{S_1(x_1, \ldots, x_n)}{S_2(x_1, \ldots, x_n)} = \frac{x_1 \times x_2 \times \ldots \times x_n}{(x_1 + \ldots + x_n)^n/n},$$

and seek an upper bound for $D_1$ subject to the condition $P = x_1 \geq x_2 \geq \ldots \geq x_n = 1$. As in the proof of Theorem 1, we show that $\partial D_1/\partial x_i \geq 0$ for $i = 2, 3, \ldots, n-1$ and finally obtain the relation

$$D_1 \leq \left[ \frac{P - 1}{\log P} \right]^{1/\log P} \frac{1}{2} \leq \log P \cdot P.$$

(7)

which leads to the inequality:

Theorem 3:

$$\frac{1}{2} \leq \log P \cdot P.$$

This inequality yields a lower bound for $P$ which is always higher than that given by (1) (as may be seen by comparing the proof of Theorem 3 with the proof of (1) in [5]). However it is cumbersome. It can be simplified (and somewhat weakened) as follows: Since $x^e = a$, $a \geq e$, implies that $x \geq a - e + 1$, and $e/D_1 \geq e$ by (2), we may conclude that $(P-1)/\log P \geq e/D_1^{1/\log P} + 1$. Denoting this last quantity by $A$, we have $(P-1)/\log P \geq A$, which has as an immediate consequence:

Theorem 3': $P \geq A \log A - 1; A = e/D_1^{1/\log P} + 1$.

This lower bound is most often, though not always, better than that given in (1).

References