Internat. J. Math. & Math. Sci.
Vol. 23, No. 8 (2000) 563–566
S016117120000257X

CAUCHY'S INTERLACE THEOREM AND LOWER BOUNDS FOR THE SPECTRAL RADIUS

A. MCD. MERCER and PETER R. MERCER

(Received 31 December 1998)

ABSTRACT. We present a short and simple proof of the well-known Cauchy interlace theorem. We use the theorem to improve some lower bound estimates for the spectral radius of a real symmetric matrix.

Keywords and phrases. Eigenvalues, symmetric matrix.

2000 Mathematics Subject Classification. Primary 15A42.

1. Cauchy's interlace theorem. We begin by presenting a short and simple proof of the Cauchy interlace theorem, which we believe to be new. See [1, 3, 4, 5], for example, for several other proofs. The theorem states that if a row-column pair is deleted from a real symmetric matrix, then the eigenvalues of the resulting matrix interlace those of the original one.

Let A be a real symmetric $n\times n$ matrix with eigenvalues (assumed distinct for now)

$$\lambda_1 < \lambda_2 < \cdots < \lambda_n \tag{1.1}$$

and normalized eigenvectors

$$\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n. \tag{1.2}$$

Let A_1 be the matrix obtained from A by deleting the first row and column. We list the eigenvalues of A_1 via $\mu_1 \le \mu_2 \le \cdots \le \mu_{n-1}$. Set

$$D(\lambda) := \det(A-\lambda I), \qquad D_1(\lambda) := \det(A_1 - \lambda I), \tag{1.3}$$

$$\mathbf{e} := [1,0,0,\dots,0]^T, \qquad \mathbf{x} := [x_1, x_2, \dots, x_n]^T. \tag{1.4}$$

Applying Cramer's rule to the set of equations $(A-\lambda I)\mathbf{x} = \mathbf{e}$ yields

$$x_1 = \frac{D_1(\lambda)}{D(\lambda)}. \tag{1.5}$$

If we write

$$\mathbf{e} = \sum c_k \mathbf{v}_k, \tag{1.6}$$

then the solution of the above set of equations reads

$$\mathbf{x} = \sum \frac{c_k}{\lambda_k - \lambda} \mathbf{v}_k. \tag{1.7}$$

On one hand,

$$\mathbf{x}\cdot\mathbf{e} = x_1, \tag{1.8}$$

while on the other hand,

$$\mathbf{x}\cdot\mathbf{e} = \sum \frac{c_k^2}{\lambda_k - \lambda}. \tag{1.9}$$

Therefore

$$\frac{D_1(\lambda)}{D(\lambda)} = x_1 = \mathbf{x}\cdot\mathbf{e} = \sum \frac{c_k^2}{\lambda_k - \lambda}. \tag{1.10}$$

Now if none of the c_k's is zero—i.e., if $\mathbf{e}$ is in *general position* with respect to $\{\mathbf{v_1},\mathbf{v_2},\dots,\mathbf{v_n}\}$—then it follows that the zeros of $D_1(\lambda)$ lie strictly between the zeros of $D(\lambda)$. That is, $\mu_k \in (\lambda_k,\lambda_{k+1})$ $(k = 1,2,\dots,n-1)$. If $\mathbf{e}$ is not in general position, then one may choose a sequence $\{\mathbf{u}_j\}$ of vectors which are in general position, and which tend to $\mathbf{e}$; passage to the limit yields $\mu_k \in [\lambda_k,\lambda_{k+1}]$. This is the Cauchy interlace theorem for the case in which A has distinct eigenvalues.

Little change in the proof is needed to deal with the case of multiple eigenvalues. We find, in particular, that if λ is an m-fold eigenvalue of A, then it is at least an $(m-1)$-fold eigenvalue of A_1 $(m \ge 2)$.

2. Lower bounds for the spectral radius. For any square matrix A we denote by $\rho(A)$ its spectral radius

$$\rho(A) = \max\ [|\lambda| : \lambda \text{ is an eigenvalue for } A]. \tag{2.1}$$

In [2], the following result is proved.

Theorem 2.1. *Let A be a real matrix with $m = \operatorname{rank}(A) \ge 2$.*

$$\text{If } \operatorname{tr}(A^2) \le (\operatorname{tr}(A))^2/m, \text{ then } \rho(A) \ge \sqrt{(\operatorname{tr}(A))^2 - \operatorname{tr}(A^2)/(m(m-1))}. \tag{2.2a}$$

$$\begin{gathered}\text{If } \operatorname{tr}(A^2) \ge (\operatorname{tr}(A))^2/m, \text{ then}\\ \rho(A) \ge |\operatorname{tr}(A)|/m + \sqrt{1/(m(m-1))[\operatorname{tr}(A^2) - (1/m)(\operatorname{tr}(A))^2]}.\end{gathered} \tag{2.2b}$$

Here we consider real symmetric matrices, in which case (2.2b) holds. We obtain a lower bound for $\rho(A)$ which is "usually" sharper than (2.2b), and which requires no knowledge of the rank. As in [2], we consider certain submatrices associated with A, but we employ Cauchy's interlace theorem instead of Lucas' theorem.

Theorem 2.2. *Let $A = [a_{jk}]$ be a real symmetric $n\times n$ matrix, with $n \ge 3$. Then*

$$\rho(A) \ge \frac{1}{2} \max_{1\le j<k\le n} \left[|a_{jj} + a_{kk}| + \sqrt{(a_{jj} - a_{kk})^2 + 4a_{jk}^2} \right]. \tag{2.3}$$

PROOF. Delete from A any $n-2$ row-column pairs, leaving a 2×2 submatrix B. It has characteristic polynomial, say, $p(\lambda) = \lambda^2 + b\lambda + c$, where $b = -\operatorname{tr}(B)$ and $2c = (\operatorname{tr}(B))^2 - \operatorname{tr}(B^2)$. As B is also symmetric it has real roots, the larger of their magnitudes being

$$\frac{1}{2}\left[|\operatorname{tr}(B)| + \sqrt{2\operatorname{tr}(B^2) - (\operatorname{tr}(B))^2}\right], \quad \text{where } B = \begin{bmatrix} a_{jj} & a_{jk} \\ a_{jk} & a_{kk} \end{bmatrix}. \tag{2.4}$$

By the Cauchy Interlace Theorem, each of the roots of p is no larger in magnitude than $\rho(A)$, and so a little manipulation gives us the desired result. □

REMARKS. (1) Deleting $n-1$ row-column pairs gives $\rho(A) \geq \max|a_{kk}|$. This result is already sharper than Theorem 2 of [2].

(2) We may delete (whenever possible) $n-3$ or $n-4$ row-column pairs to obtain characteristic polynomials of degree 3 or 4, then proceed as above to obtain increasingly sharper but less manageable estimates.

(3) Analogous results can be obtained for skew-symmetric matrices, which involve maximums of off-diagonal entries. We leave the interested reader to fill in the details.

(4) As was done in [2], we generated 1000 random (but symmetric) $n\times n$ matrices with integer entries in $[-10, 10]$, for $n = 4$, $n = 8$, and $n = 12$. We calculated the average ratios of each of the bounds obtained in Theorems 2.1 and 2.2 to the actual spectral radius. We used *Mathematica*, and our results are summarized in Table 2.1.

TABLE 2.1.

	Theorem 2.1	Theorem 2.2
$n = 4$	0.517802	0.802070
$n = 8$	0.285717	0.739505
$n = 12$	0.208946	0.694311

We add that our ratios also compare favorably with those arising from all of the results quoted in [2]—see Table 2.1.

(5) As the numerical evidence suggests, Theorem 2.2 is "usually" sharper than Theorem 2.1 (in the symmetric case). If A is $n\times n$, and $\operatorname{rank}(A) = n$, then Theorem 2.2 is at least as sharp as Theorem 2.1: the $\binom{n}{2}$ numbers whose maximum is taken in Theorem 2.2 are the roots of larger magnitude of $\binom{n}{2}$ quadratics, whose sum is the quadratic with the estimate in Theorem 2.1 as its root of larger magnitude. If $\operatorname{rank}(A) < n$, then there is no simple relationship: the matrices (each with eigenvalue $\lambda = 0$)

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} -1 & 0 & 2 \\ 0 & 1 & 2 \\ 2 & 2 & 0 \end{bmatrix}, \quad C = \begin{bmatrix} 0 & -1 & 0 & 0 \\ -1 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{2.5}$$

provide all three possibilities. For A, the estimates are equal. For B, Theorem 2.1 is sharper. For C, Theorem 2.2 is sharper.

References

[1] R. Bellman, *Introduction to Matrix Analysis*, 2nd ed., McGraw-Hill Book Co., New York, 1970. MR 41#3493. Zbl 216.06101.

[2] B. G. Horne, *Lower bounds for the spectral radius of a matrix*, Linear Algebra Appl. **263** (1997), 261–273. MR 98h:15034. Zbl 889.15014.

[3] Y. Ikebe, T. Inagaki, and S. Miyamoto, *The monotonicity theorem, Cauchy's interlace theorem, and the Courant-Fischer theorem*, Amer. Math. Monthly **94** (1987), no. 4, 352–354. MR 88a:15035. Zbl 623.15010.

[4] A. McD. Mercer, *On "deleting a row and column" from a differential operator*, Acta Math. Hungar. **74** (1997), no. 4, 321–331. MR 98c:47023. Zbl 990.64043.

[5] B. N. Parlett, *The Symmetric Eigenvalue Problem*, Prentice-Hall Inc., Englewood Cliffs, N.J., 1980. MR 81j:65063. Zbl 431.65017.

Mercer: Department of Mathematics and Statistics, University of Guelph, Ontario N1G-2W1, Canada

E-mail address: amercer@connect1.reach.net

Mercer: Department of Mathematics, SUNY College at Buffalo, NY 14222, USA

E-mail address: mercerpr@buffalostate.edu