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)

\[
\lambda_1 < \lambda_2 < \cdots < \lambda_n
\]

(1.1)

and normalized eigenvectors

\[
v_1, v_2, \ldots, v_n.
\]

(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 \( \mu_1 \leq \mu_2 \leq \cdots \leq \mu_{n-1} \). Set

\[
D(\lambda) := \det(A - \lambda I), \quad D_1(\lambda) := \det(A_1 - \lambda I),
\]

(1.3)

\[
e := [1, 0, 0, \ldots, 0]^T, \quad x := [x_1, x_2, \ldots, x_n]^T.
\]

(1.4)

Applying Cramer’s rule to the set of equations \((A - \lambda I)x = e\) yields

\[
x_1 = \frac{D_1(\lambda)}{D(\lambda)}.
\]

(1.5)

If we write

\[
e = \sum c_k v_k,
\]

(1.6)

then the solution of the above set of equations reads

\[
x = \sum \frac{c_k}{\lambda_k - \lambda} v_k.
\]

(1.7)
On one hand,
\[ x \cdot e = x_1, \quad (1.8) \]
while on the other hand,
\[ x \cdot e = \sum c_k^2 \frac{\lambda_k - \lambda}{\lambda_k - \lambda}. \quad (1.9) \]

Therefore
\[ \frac{D_1(\lambda)}{D(\lambda)} = x_1 = x \cdot e = \sum c_k^2 \frac{\lambda_k - \lambda}{\lambda_k - \lambda}. \quad (1.10) \]

Now if none of the \( c_k \)'s is zero—i.e., if \( e \) is in general position with respect to \( \{v_1, v_2, \ldots, v_n\} \)—then it follows that the zeros of \( D_1(\lambda) \) lie strictly between the zeros of \( D(\lambda) \).

That is, \( \mu_k \in (\lambda_k, \lambda_{k+1}) \) (\( k = 1, 2, \ldots, n-1 \)). If \( e \) is not in general position, then one may choose a sequence \( \{u_j\} \) of vectors which are in general position, and which tend to \( e \); 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 \( \lambda \) is an \( m \)-fold eigenvalue of \( A \), then it is at least an \( (m-1) \)-fold eigenvalue of \( A_1 \) (\( m \geq 2 \)).

2. Lower bounds for the spectral radius. For any square matrix \( A \) we denote by \( \rho(A) \) its spectral radius
\[ \rho(A) = \max \{ |\lambda| : \lambda \text{ is an eigenvalue for } A \}. \quad (2.1) \]

In [2], the following result is proved.

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

If \( \text{tr}(A^2) \leq (\text{tr}(A))^2/m \), then \( \rho(A) \geq \sqrt{\text{tr}(A^2) - (\text{tr}(A))^2/(m(m-1))} \). \quad (2.2a)

If \( \text{tr}(A^2) \geq (\text{tr}(A))^2/m \), then
\[ \rho(A) \geq |\text{tr}(A)|/m + \sqrt{1/(m(m-1))}[\text{tr}(A^2) - (1/m)(\text{tr}(A))^2]. \quad (2.2b) \]

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 \( n \geq 3 \). Then
\[ \rho(A) \geq \frac{1}{2} \max_{1 \leq j < k \leq n} \left[ |a_{jj} + a_{kk}| + \sqrt{(a_{jj} - a_{kk})^2 + 4a_{jk}^2} \right]. \quad (2.3) \]
**Proof.** Delete from $A$ any $n - 2$ row-column pairs, leaving a $2 \times 2$ submatrix $B$. It has characteristic polynomial, say, $p(\lambda) = \lambda^2 + b\lambda + c$, where $b = -\text{tr}(B)$ and $2c = (\text{tr}(B))^2 - \text{tr}(B^2)$. As $B$ is also symmetric it has real roots, the larger of their magnitudes being

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

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. \(\square\)

**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>
<thead>
<tr>
<th>$n$</th>
<th>Theorem 2.1</th>
<th>Theorem 2.2</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>0.517802</td>
<td>0.802070</td>
</tr>
<tr>
<td>8</td>
<td>0.285717</td>
<td>0.739505</td>
</tr>
<tr>
<td>12</td>
<td>0.208946</td>
<td>0.694311</td>
</tr>
</tbody>
</table>

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 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 rank($A$) < $n$, then there is no simple relationship: the matrices (each with eigenvalue $\lambda = 0$)

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

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


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: mercepr@buffalostate.edu