GERŠGORIN THEOREMS BY HOUSEHOLDER'S PROOF

BY J. L. BRENNER

Communicated by I. Reiner, December 11, 1967

0. The method. Given an \( m \times m \) matrix \( A = [a_{ij}] \) of complex numbers, S. Geršgorin [4] proved that every proper value \( \lambda \) lies in the union of the \( m \) disks \( D_i \), where

\[
D_i = \{ \lambda \mid |\lambda - a_{ii}| < R_i, \quad R_i = \sum_{j \neq i} |a_{ij}| \}\.
\]

Generalizations of this theorem have appeared in several papers, see for example [1], [3], [5], [7], [8], and a convenient summary in [6]. The theorem is derivable from the following (older) result, if we set \( B = A - \lambda I \).

**Theorem 1.** Let \( B = [b_{ij}] \) be a matrix of complex numbers. If \( B \) is not invertible, then for some \( i \) we must have \(|b_{ii}| \leq \sum_{j \neq i} |b_{ij}| = R_i\).

**Corollary.** \( \forall_i \{ |b_{ii}| > R_i \} \Rightarrow B \) is invertible.

This is the contrapositive of Theorem 1. To prove Theorem 1, find \( x = \{x_1, x_2, \ldots, x_n\} \) so that \( Bx = 0 \); choose \( i \) so that \( x_i \neq 0 \) and \( \forall_j \{ |x_i| \geq |x_j| \} \). Then \(|b_{ii}| \leq \sum_{j \neq i} |b_{ij}| \cdot |x_j/x_i| \leq R_i\).

Householder [5, p. 66] looks at the theorem from a different point of view. He writes \( B = D - C \), where \( D \) is the diagonal part of \( B \), i.e. \( D = [d_{ij}] \), \( d_{ii} = \delta_i \cdot b_{ii} \), and \( C \) has zero diagonal. If \( \forall_i \{ b_{ii} \neq 0 \} \), then \( B = D(I - D^{-1}C) \). The condition \( \|D^{-1}C\| < 1 \) guarantees that \( B \) be invertible. The corollary follows on applying this condition and using the row-sum norm.

1. A new result. In the preceding paragraph, a known result was recovered by Householder's method. This does not demonstrate the full power of the method. In this section, we obtain a new result by the same method. (This result can be obtained also by other methods; see [2].)

**Definition.** The notation

\[
B \left( \begin{array}{c} 1 \cdots n \\ 1 \cdots n \end{array} \right)
\]

means the minor matrix obtained from the large matrix \( B \) by retaining only rows \( 1 \cdots n \) and columns \( 1 \cdots n \). The notation

\[
B \left( \begin{array}{c} 1 \cdots n \\ \{1 \cdots n\} \setminus \{i, j\} \end{array} \right)
\]
means the minor matrix based on rows $1 \cdots n$ and columns $1 \cdots n$, but with column $t$ omitted and column $j$ ($j > n$) appended.

**Lemma.** Let $c_{ik}$ be the $i$, $k$ element of

$$W = B \begin{pmatrix} 1 \cdots n \\ 1 \cdots n \end{pmatrix}^{-1}.$$

Then

$$\left| \det \left\{ WB \begin{pmatrix} 1 \cdots n \\ \{1 \cdots n\} \setminus \{t, j\} \end{pmatrix} \right\} \right| = | \sum c_{ik} b_{kj} |.$$

**Proof.** The matrix product $Q$ on the left side of the lemma is equal to the identity matrix except in the $t$th column, which is replaced by the $j$th column as shown. The determinant of $Q$ is therefore equal to the $t$, $t$ element of $Q$, i.e. the inner product of the $t$th row of $W$ by the $j$th column of $B$.

**Theorem 2.** Let $B$ be an $m \times m$ matrix of complex numbers; let $S(1)$, $S(2)$, $\cdots$ be a partitioning of $\{1 \cdots m\}$ into disjoint sets. Let

$$V(r) = \begin{pmatrix} S(r) \\ S(r) \end{pmatrix}$$

be the (principal) submatrix of $B$ on the rows and columns with indices in $S(r)$. Let

$$U(r, j, t) = A \begin{pmatrix} S(r) \\ S(r) \setminus \{j, t\} \end{pmatrix}$$

be the submatrix of $B$ that uses rows with indices in $S(r)$, and columns with indices from the same set, but with the column of index $j$ deleted and the column of index $t$ appended.

The matrix $B$ is nonsingular if the following $m$ inequalities hold among certain minor determinants of $B$:

$$\forall_{j \in S(r)} \forall_{r} \left\{ | \det V(r) | > \sum_{t \in S(r)} | \det U(r, j, t) | \right\}.$$

**Remark.** If $S(i) = \{i\}$, this theorem reduces to the Geršgorin corollary.

**Proof.** We write $B = D - C = D(I - D^{-1}C)$ as before, but interpret $D$ as the block diagonal $V(1) + V(2) + \cdots$ of $B$. If we apply the lemma (read from right to left) to the matrix $D^{-1}C$, and use row-sum norm in the condition $\|D^{-1}C\| < 1$, Theorem 2 follows.
Corollary. Every proper value of the matrix $A$ lies in one or another of the $m$ loci

$$
\left| \det \begin{pmatrix}
\lambda - a_{r,r} & a_{r,r+1} \\
a_{r+1,r} & \lambda - a_{r+1,r+1}
\end{pmatrix}\right| \leq \sum_{\delta \neq r, r+1} \left| \det \begin{pmatrix}
\lambda - a_{r,r} & a_{r,r+1} \\
a_{r+1,r} & \lambda - a_{r+1,r+1}
\end{pmatrix}\right|,
$$

$$
\left| \det \begin{pmatrix}
\lambda - a_{r,r} & a_{r,r+1} \\
a_{r+1,r} & \lambda - a_{r+1,r+1}
\end{pmatrix}\right| \leq \sum_{\delta \neq r} \left| \det \begin{pmatrix}
\lambda - a_{r,r} & a_{r,r+1} \\
a_{r+1,r} & \lambda - a_{r+1,r+1}
\end{pmatrix}\right|,
$$

$r = 1, 3, 5, \ldots, m-1$. (If $m$ is odd, the last value of $r$ is $m-2$, and the disk $|a_{mm}-\lambda| \leq R_m$ must be appended.)

This corollary has been used in numerical analysis, in a case in which complex numbers are replaced by $2 \times 2$ matrices.

References


Stanford Research Institute and University of British Columbia