Numer. Math. 16, 141–144 (1970)

A Note on Gerschgorin's Inclusion Theorem for Eigenvalues of Matrices*

ALSTON S. HOUSEHOLDER, RICHARD S. VARGA, and JAMES H. WILKINSON

Received January 24, 1970

Let A be any $n\times n$ complex matrix, and let $\sigma(A)$ denote the set of its eigenvalues. For ϕ any fixed vector norm on $V_n(C)$, let $\|E\|_\phi \equiv \sup_{\phi(x)=1} \phi(Ex)$ denote the associated induced operator norm of any $n\times n$ matrix E. The well known Gerschgorin argument for obtaining inclusion regions for the eigenvalues of A can be described as follows (cf. Householder [1, p. 65]). For $\lambda\in\sigma(A)$, there is evidently a vector $x\neq 0$ such that $Ax=\lambda x$. Thus, if D is any fixed $n\times n$ matrix, this can be written equivalently as

$$(A-D+(z-\lambda)I)\,x=(zI-D)\,x$$

for any scalar z. In particular, if $z\notin\sigma(D)$, then

$$(zI-D)^{-1}\{A-D+(z-\lambda)I\}\,x=x,$$

which implies that

$$\|(zI-D)^{-1}\{A-D+(z-\lambda)I\}\|_\phi \geqq 1 \qquad \forall z\notin\sigma(D). \tag{1}$$

In particular, if $\lambda\notin\sigma(D)$, z can be chosen equal to λ in (1), which gives

$$\|(\lambda I-D)^{-1}(A-D)\|_\phi \geqq 1 \qquad \lambda\notin\sigma(D). \tag{2}$$

If $\lambda\in\sigma(D)$, then (1) is of course valid. Thus, if the Gerschgorin set $G_\phi(A)$ in the complex plane is defined as

$$G_\phi(A)=G_\phi^1(A)\cup G_\phi^2(A) \tag{3}$$

where

$$G_\phi^1(A)\equiv\{z:\ z\notin\sigma(D)\ \text{and}\ \|(zI-D)^{-1}(A-D)\|_\phi\geqq 1\}, \tag{4}$$

and where

$$G_\phi^2(A)\equiv\{w:\ w\in\sigma(D)\ \text{and}\ \|(zI-D)^{-1}\{A-D+(z-w)I\}\|_\phi\geqq 1\ \ \forall z\notin\sigma(D)\} \tag{5}$$

is a subset of $\sigma(D)$, then the above argument gives us

Theorem 1. For any fixed vector norm ϕ, and a fixed $n\times n$ matrix D,

$$\sigma(A)\subset G_\phi(A). \tag{6}$$

As is readily verified, $G_\phi(A)$ is a bounded set in the complex plane.

* This research was supported in part by the Atomic Energy Commission, the National Science Foundation, and the Office of Naval Research.

All this may seem rather familiar to the reader, and at first glance, it would seem that the closure of $G^1_\phi(A)$, $\overline{G^1_\phi(A)}$, would also give an inclusion region for $\sigma(A)$, and this is tacitly assumed in [1, p. 66], and in other places as well. One purpose of this note is to show that $\overline{G^1_\phi(A)}$ does *not* in general include $\sigma(A)$, and it thus is necessary to include $G^2_\phi(A)$ in $G_\phi(A)$ to have the desired inclusion of (6). As a counterexample, consider

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 5 & 1 \\ 1 & 1 & 5 \end{bmatrix}, \qquad D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 5 \end{bmatrix} \tag{7}$$

and let $\phi((x_1, x_2, x_3)^T) \equiv \max(|x_1|, |x_2|, |x_3|)$. As is easily verified for this case, $G^1_\phi(A) = \{z\colon |z-5| \leqq 2\}$, which does *not* include $\sigma(A) = \{1\} \cup \{4\} \cup \{6\}$.

It is logical to ask when $\overline{G^1_\phi(A)} = G_\phi(A)$. Certainly if for each eigenvalue λ_i of D in $G^2_\phi(A)$, one has that $\|(zI-D)^{-1}(A-D)\|_\phi \to \infty$ as $z \to \lambda_i$, then each such λ_i is in the closure of $G^1_\phi(A)$, and consequently, $\overline{G^1_\phi(A)} = G_\phi(A)$. This brings us to

Theorem 2. Let $\lambda \in \sigma(D)$. Then, $(zI-D)^{-1}(A-D)$ remains bounded (i.e., in every element) as $z \to \lambda$ if and only if every left-hand principal vector of D belonging to λ is a left-hand eigenvector of $A-D$ corresponding to the eigenvalue zero. In the contrary case, $\|(zI-D)^{-1}(A-D)\|_\phi \to +\infty$ as $z \to \lambda$ where the norm ϕ is arbitrary, but preassigned.

Proof. First, let

$$y_1^H D = \lambda y_1^H, \qquad y_1^H \neq 0. \tag{8}$$

Then,

$$y_1^H (zI-D)^{-1} = (z-\lambda)^{-1} y_1^H, \tag{9}$$

and hence,

$$y_1^H (zI-D)^{-1}(A-D) = (z-\lambda)^{-1} y_1^H (A-D).$$

Clearly, the right member becomes infinite as $z \to \lambda$ unless

$$y_1^H (A-D) = 0, \tag{10}$$

i.e., y_1^H is a left-hand eigenvector of $A-D$ corresponding to the eigenvalue zero. With the assumption of (10), suppose next that there is a $y_2 \neq 0$ such that

$$y_2^H (D-\lambda I) = y_1^H, \tag{11}$$

so that y_2^H is a left-hand principal vector of D belonging to λ (cf. [1, p. 3]). Then, after some minor manipulations of the identities (8)—(11), we have

$$y_2^H (zI-D)^{-1}(A-D) = (z-\lambda)^{-1} y_2^H (A-D). \tag{12}$$

Again, the right member becomes infinite unless

$$y_2^H (A-D) = 0. \tag{13}$$

Next, letting $y_3^H(D-\lambda I)=y_2^H$, a similar argument shows that $y_3^H(zI-D)^{-1}(A-D)$ becomes infinite unless $y_3^H(A-D)=0$. In this way, $(zI-D)^{-1}(A-D)$ is unbounded as $z\to\lambda$ unless each left principal vector of D corresponding to λ is a left-hand eigenvector of $A-D$ corresponding to the eigenvalue zero. This establishes the first part of the theorem.

Conversely, suppose for $\lambda\in\sigma(D)$, every left-hand principal vector of D belonging to λ is a left-hand eigenvector of $A-D$ corresponding to the eigenvalue zero. Let E_1 be the invariant subspace spanned by the principal vectors of D belonging to the eigenvalue $\lambda\in\sigma(D)$, and let E_2 analogously be the invariant subspace belonging to all other eigenvalues of D. Then, any vector y is uniquely expressible in the form $y=y_1+y_2$ where $y_i\in E_i$, $i=1, 2$. The hypothesis is that $y_1^H(A-D)=0$. But,

$$y_2^H(zI-D)^{-1}\in E_2$$

is a linear combination of principal vectors of D belonging to eigenvalues $\mu\neq\lambda$, with multipliers equal to powers of $(z-\mu)^{-1}$. Hence, $y_2^H(zI-D)^{-1}$ remains bounded as $z\to\lambda$, and so, too, does $y_2^H(zI-D)^{-1}(A-D)$. It follows that for every fixed y, $y^H(zI-D)^{-1}(A-D)$ is bounded as $z\to\lambda$, and thus every element of $(zI-D)^{-1}(A-D)$ remains bounded as $z\to\lambda$. The assertion about the norm follows from the fact that if any element of the matrix becomes infinite, so does the value of the norm. Q.E.D.

As an obvious consequence of Theorem 2, we have the

Corollary. If A and D have no common left-hand eigenvector corresponding to any common eigenvalue λ_i they may have, then

$$\lim_{\substack{z\to\lambda_i\\ z\notin\sigma(D)}}\|(zI-D)^{-1}(A-D)\|_\phi=\infty.$$

It is also interesting to give the result of Theorem 2 in terms of the resolvent operator $R_z\equiv(zI-D)^{-1}$. If λ is any (isolated) point of $\sigma(D)$, then it is known (cf. [2, p. 305]) that R_z has a Laurent expansion in powers of $z-\lambda$,

$$R_z=\sum_{k=0}^{\infty}(z-\lambda)^k A_k+\sum_{k=1}^{n}(z-\lambda)^{-k}B_k \tag{14}$$

where A_k and B_k are $n\times n$ matrices. Then, the following analog of Theorem 2 is evident.

Theorem 3. Let $\lambda\in\sigma(D)$, and let the resolvent operator $R_z=(zI-D)^{-1}$ have the Laurent expansion of (14). Then $R_z(A-D)$ is a bounded operator as $z\to\lambda$ if and only if

$$B_k(A-D)=0,\qquad k=1, 2, \ldots, n. \tag{15}$$

We remark that for the counterexample given in (7), for $\lambda_1=1$, the vector $(1, 0, 0)$ is a left eigenvector of D which is a left eigenvector of $A-D$ corresponding to the eigenvalue zero. Also, for the analog of Theorem 3 in this case, all B_k of (14) are null, except for

$$B_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

and it is clear that (15) is satisfied.

References

1. Householder, Alston S.: The Theory of Matrices in Numerical Analysis. New York: Blaisdell Publishing Co. 1964, 257 pp.
2. Taylor, Angus E.: Introduction to Functional Analysis. New York: John Wiley and Sons, Inc. 1961, 423 pp.

A. S. Householder
University of Tennessee
Knoxville, Tenn.
U.S.A.

J. H. Wilkinson
National Physical Lab.
Teddington, Middlesex
U.K.

R. S. Varga
Kent State University
Kent, Ohio
U.S.A.