CHARLES H. ROMINE

ON AN UNSOLVED PROBLEM OF OLGA TAUSSKY

This article summarizes recent investigations about an unsolved problem in the distribution of eigenvalues over a class of matrices. Several elementary results in matrix theory are used to generate initial results. Empirical results based on a computational study are also presented. The empirical study was motivated by the ready availability of high-performance computers that can quickly compute thousands of eigenvalues.

INTRODUCTION

While doing some background reading on the theory of matrices, I recently read a survey paper by Olga Taussky[1] that is a written record of a lecture given by her at the second advanced seminar conducted by the Mathematics Research Center of the U.S. Army. Many results summarized in that paper concern the perturbation theory of matrix eigenvalues: the analysis of the effect that small perturbations in the elements of the matrix have on the eigenvalues (characteristic roots). After presenting several interesting theorems in perturbation theory, she asked the following intriguing question:

> An unsolved problem is: Take a matrix A, replace its elements in all possible ways by numbers which have the same absolute value. What is the region in the complex plane covered by all their roots?

The purpose of this article is to describe some investigations into Taussky's question. I do not claim that the results presented here are new, although I have been unable to locate research that has cited Taussky's paper. On the other hand, her paper appeared more than twenty-five years ago, and it is likely that the problem has been addressed. Moreover, although the question is not addressed in the paper, she was surely aware of some of the elementary results given here. The advances in computing resources since the paper appeared, however, make it possible to study the problem empirically. In fact, a primary motivation for working on the problem was the ability to compute and plot thousands of eigenvalues with almost instant turnaround in an interactive session. The results presented should give the reader an appreciation for the role that computing can play in the investigation of theoretical questions in numerical analysis.

BACKGROUND

In this section, let us review some notational conventions and well-known theorems of matrix theory that we will use in the sequel.

We denote the real numbers by $\mathbf{R}$, the complex numbers by $\mathbf{C}$, and the set of all $n \times n$ matrices with real (complex) entries as $\mathbf{R}^{n\times n}$ ($\mathbf{C}^{n\times n}$). Unless otherwise specified, all matrices under consideration are assumed to lie in $\mathbf{C}^{n\times n}$.

Definition 1. A matrix A is singular if there exists some nonzero vector $\mathbf{x}$ for which $A\mathbf{x} = 0$ is satisfied. Otherwise, it is called nonsingular.

Equivalently, the matrix A is singular if and only if $\det A = 0$. (Recall that $\det A$ denotes the determinant of the matrix A.)

Definition 2. A number $\lambda \in \mathbf{C}$ is an eigenvalue of the matrix A if it satisfies the equation

$$A\mathbf{x} = \lambda\mathbf{x} \tag{1}$$

for some complex vector $\mathbf{x}$. The set of all eigenvalues of A is the spectrum of A, denoted $\lambda(A)$.

Equation 1 can be rewritten as $A\mathbf{x} - \lambda\mathbf{x} = 0$ or $(A - \lambda I)\mathbf{x} = 0$, where I is the $n \times n$ identity matrix

$$I = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix} .$$

In particular, λ is an eigenvalue of A if and only if $\det(A - \lambda I) = 0$. It is easy to show that $\det(A - \lambda I)$ is a polynomial of degree n in λ; hence, by the Fundamental Theorem of Algebra, A has exactly n eigenvalues, although they need not be distinct.

Definition 3. The nth degree polynomial $\det(A - \lambda I)$ is called the characteristic polynomial of the matrix A.

Hence, the eigenvalues of A are simply the roots of its characteristic polynomial.

The characteristic polynomial is clearly a function of the entries of the matrix; hence, the eigenvalues (roots of the characteristic polynomial) are also functions of the entries of the matrix. How "well-behaved" a function the eigenvalues are with respect to the entries of the matrix is described in the following theorem.

Theorem 1. The eigenvalues of a matrix A are continuous functions of the entries of A.

Loosely speaking, continuity of the eigenvalues means that a small enough change in the entries of A will ensure that the eigenvalues of A change very little, a fact that will have important consequences later.

Another elementary theorem in matrix theory that will be useful to us is the following.

Theorem 2. For any scalar $z \in \mathbf{C}$ and matrix A, if λ is an eigenvalue of A, then $z\lambda$ is an eigenvalue of zA.

Finally, we present a theorem on the "approximate" location of the eigenvalues of a matrix.

Theorem 3 (Gerschgorin, part I): If $A = D + F$, where $D = \text{diag}(d_1, d_2, \ldots, d_n)$ is the diagonal of A, then

$$\lambda(A) \subseteq \bigcup_{i=1}^{n} D_i , \tag{2}$$

where

$$D_i = \left\{ z \in \mathbf{C}: |z - d_i| \leq \sum_{j=1}^{n} |f_{ij}| \right\} . \tag{3}$$

The D_i are called the Gerschgorin disks, and $\lambda(A)$ is simply the region of the complex plane formed by the union $\cup D_i$ of these disks. For simplicity, the Gerschgorin disks are represented by circles in the figures. The reader should bear in mind, however, that the Gerschgorin disks consist of the entire interior of these circles, not just the boundary.

For a proof of Gerschgorin's theorem, see Golub and Van Loan.[2] For our purposes, only the moduli of the entries of F (i.e., the off-diagonal entries of A) are used in the theorem, not the entries themselves. Hence, the disks D_i are invariant under replacement of the off-diagonal elements of A by elements of the same modulus. (Recall that the modulus, or absolute value of a complex number $z = a + ib$ is its distance to the origin in the complex plane $|z| = \sqrt{a^2 + b^2}$.) To simplify the discussion of such replacements of elements of A, we present the following definition.

Definition 4. The replacement of an element (or elements) of A by complex numbers with the same modulus (moduli) as the original elements will be called a unit change in A.

The term unit change denotes that the element (or elements) of A is multiplied by a complex number of modulus one, that is, a complex number on the unit circle in the complex plane.

To describe precisely the region in the complex plane under consideration, we present the following definitions and notation.

Definition 5. We denote the set of all complex matrices whose entries have the same modulus as the entries of A by $\mathcal{M}(A)$.

Definition 6. We denote the set of all eigenvalues of all matrices in $\mathcal{M}(A)$ by $\Lambda(A)$. In other words, $\Lambda(A)$ is the union of the eigenvalues of all possible matrices in $\mathcal{M}(A)$:

$$\Lambda(A) = \bigcup_{M \in \mathcal{M}(A)} \lambda(M) .$$

It is the region in the complex plane defined by $\Lambda(A)$ that we wish to identify.

LOCI OF EIGENVALUES

The problem of determining the region $\Lambda(A)$ is difficult, but several elementary observations can be made. First, if λ is an eigenvalue of $M \in \mathcal{M}(A)$, then every point on the entire circle of radius $|\lambda|$ centered at (0, 0) is an eigenvalue of some matrix in $\mathcal{M}(A)$. For a given $\mu \in \mathbf{C}$ with $|\mu| = |\lambda|$, we write $\mu = e^{i\theta}\lambda$ for some $\theta \in [0, 2\pi]$. Then by Theorem 2, μ is an eigenvalue of $e^{i\theta}A$, which is in $\mathcal{M}(A)$.

Thus, we can imagine the region $\Lambda(A)$ in the plane to be an uncountable union of circles centered at the origin. Moreover, by Theorem 1, the eigenvalues of A are continuous in the entries of A. Hence, since there are n eigenvalues of the original matrix, by continuity there are at most n disconnected regions in the plane covered by $\Lambda(A)$. Thus, $\Lambda(A)$ consists of n concentric (possibly overlapping) annuli centered about the origin. The remaining (and most difficult) problem is to determine the inner and outer radii of the annuli.

One bound can be derived immediately from Gerschgorin's theorem; that is, form the region obtained by revolving the Gerschgorin disks D_i about the origin. The region $\mathcal{D}(A)$ swept out by the union of these disks will form concentric (possibly overlapping) annuli, and Gerschgorin's theorem (together with the observation made after the theorem) guarantees that

$$\Lambda(A) \subseteq \mathcal{D}(A) . \tag{4}$$

A simple example will illustrate the conclusions reached thus far. Suppose the original entries of a given matrix A are multipled by $e^{i\theta}$, for random θ uniformly distributed on the interval $[0, 2\pi]$, and the eigenvalues of the resulting matrix plotted. The graph thus generated should give an indication of the region $\Lambda(A)$. Here, a 2×2 matrix is sufficient, so let

$$A_i = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} .$$

If the procedure is repeated 1000 times, the result is the graph in Figure 1. (In each figure that follows, the Gerschgorin disks of the original matrix are plotted, along with the eigenvalues of 1000 random matrices in $\mathcal{M}(A)$. The region $\mathcal{D}(A)$ can be generated by revolving the disks about the origin.)

In this case, the two Gerschgorin disks D_1 and D_2 of A_1 coincide, with $D_1 = D_2 = \{z : |z - 2| \leq 1\}$, which is the disk of radius 1 centered at 2. Thus, $\mathcal{D}(A_1)$ is the annulus with inner radius 1 and outer radius 3.

Note that the distribution of the eigenvalues plotted in Figure 1 is densely concentrated around a central band in the interior of the annulus. Of course, we have no reason to believe that a uniform distribution of θ on $[0, 2\pi]$ should result in a uniform distribution of eigenvalues in $\Lambda(A_1)$. What initial distribution of θ in the random multiplication procedure described above would result in a uniform distribution of eigenvalues in $\Lambda(A)$? Does this distribution depend on A?

The plot in Figure 1 makes it appear that $\Lambda(A_1) = \mathcal{D}(A_1)$. On this basis, one might be tempted to conjecture that equality holds in general in Equation 4; that is, $\Lambda(A) = \mathcal{D}(A)$ for all A. But consider what happens to $\mathcal{D}(A_1)$ if we replace the (2, 1) element of A_1 by some smaller value, for example,

$$A_2 = \begin{bmatrix} 2 & 1 \\ 0.1 & 2 \end{bmatrix} .$$

Figure 2 shows the result of repeating the computational experiment on A_2. Note that the Gerschgorin disks are now distinct, but $D_2 \subset D_1$. Hence, the region $\mathcal{D}(A_2)$ is the same as $\mathcal{D}(A_1)$. However, Figure 2 reveals quite a different distribution of eigenvalues than in Figure 1. It appears that $\Lambda(A_2)$ is strictly smaller than $\mathcal{D}(A_1)$.

Might the computational experiment simply fail to discover some of the points in $\Lambda(A_2)$? The answer is no. Consider the limiting case of shrinking the (2, 1) element, namely, the 2×2 matrix

$$A_2 = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix} .$$

Since the eigenvalues of a triangular matrix are simply the diagonal elements, the region $\Lambda(A)$ is just the circle of radius 2 centered at the origin, whereas $\mathcal{D}(A)$ is still unchanged.

One final example will indicate the importance of continuity of the eigenvalues. As discussed previously, continuity implies that the region $\Lambda(A)$ is the union of n annuli centered at the origin. As we have just seen in the last example, these annuli may in fact degenerate to circles. Nevertheless, a second part of Gerschgorin's theorem will allow us to make further claims concerning the number of disjoint annuli making up the region $\Lambda(A)$.

Theorem 4 (Gerschgorin, part II). If the union of m of the disks D_i is disjoint from the remainder of the disks, that union contains exactly m eigenvalues of A.

Corollary 1. If the region $\mathcal{D}(A)$ consists of k disjoint annuli, the region $\Lambda(A)$ must consist of at least k disjoint annuli.

To see this, recall that an annulus of $\mathcal{D}(A)$ is formed by sweeping the disks D_i around the origin. Each disjoint annulus of $\mathcal{D}(A)$ must have been formed from the union S of a subset of disks that are disjoint from the remainder of the disks. Hence, S must contain at least one eigenvalue of A (indeed, by Theorem 4, part II, it must contain exactly m eigenvalues, where m is the number of disks in the union S). Thus, $\Lambda(A)$ must have a nonempty component (annulus) contained within the annulus formed by rotating S about the origin.

Figure 3 shows the result of the computational experiment on a matrix with three disjoint annuli in $\mathcal{D}(A_3)$, corresponding to the matrix

$$A_3 = \begin{bmatrix} 4 & 1 & 1 \\ 1 & 9 & 1 \\ 1 & 1 & 16 \end{bmatrix} .$$

As seen from Figure 3, the region $\Lambda(A_3)$ consists of three disjoint components, as required by the previous discussion.

In contrast, Figure 4 shows the results on a matrix with three overlapping annuli in $\mathcal{D}$, corresponding to the matrix

$$A_4 = \begin{bmatrix} 4 & 1 & 1 \\ 1 & 5 & 1 \\ 1 & 1 & 6 \end{bmatrix} .$$

Evidently, the region $\Lambda(A_4)$ consists of a single annulus, consistent with the fact that the Gerschgorin disks of A_4 all overlap.

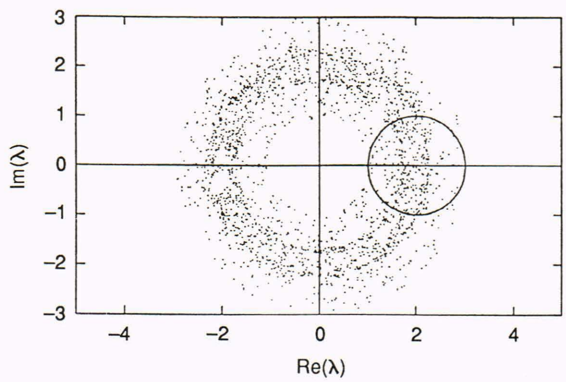

Figure 1. Results of the computational experiment with matrix A_1.

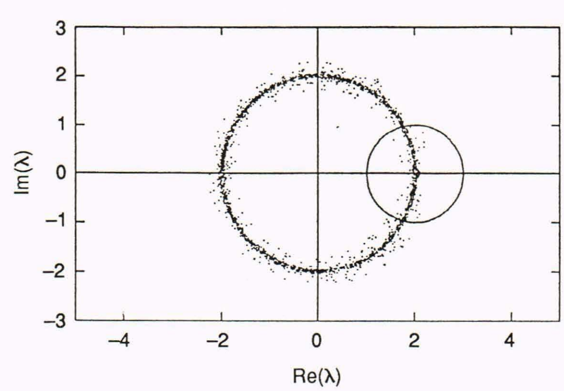

Figure 2. Results of the computational experiment with matrix A_2.

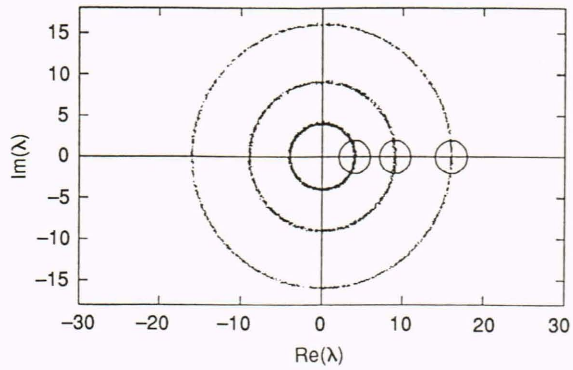

Figure 3. Results of the computational experiment with matrix A_3.

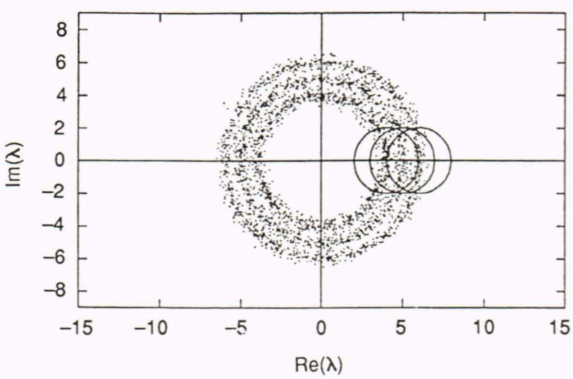

Figure 4. Results of the computational experiment with matrix A_4.

Since those disks overlap—and hence the region $\mathcal{D}(A_4)$ has only one connected component—the corollary to Theorem 4, part II, says nothing about how many disjoint annuli $\Lambda(A_4)$ can contain. This leads naturally to the following question: Is the converse of Corollary 1 true? That is, if $\mathcal{D}(A)$ consists of k disjoint annuli, must $\Lambda(A)$ also have exactly k disjoint components?

A simple counterexample shows that the answer is no. Consider the matrix

$$A_5 = \begin{bmatrix} 1.9 & 1 & 0 \\ 0 & 2.1 & 1 \\ 0 & 0 & 2 \end{bmatrix} .$$

Since A_5 is upper triangular, the region $\Lambda(A_5)$ is just the union of the three circles centered at the origin with radii 1.9, 2.1, and 2. These are clearly disjoint, so $\Lambda(A_5)$ has three disjoint components. The three Gerschgorin disks corresponding to A_5 are, however, the disk centered at 1.9 with radius 1, the disk centered at 2.1 with radius 1, and the disk centered at 2 with radius 0 (i.e., the point 2). Clearly, the region $\mathcal{D}(A_5)$ formed by rotating these disks about the origin forms a single annulus with inner radius 0.9 and outer radius 3.1.

CONCLUSION

Sharp bounds on the eigenvalues of a general complex matrix in terms of its elements do not now exist. Although Gerschgorin's theorem is not the sharpest known bound, other bounds (such as the 1-norm and ∞-norm bounds) that are also invariant with respect to unit changes in the entries of the original matrix are also insufficient to categorize $\Lambda(A)$. We claim that empirical investigations such as those described in the preceding sections, however, can offer insight into the behavior of the eigenvalues of a matrix as it undergoes unit changes. Future, more methodical computational investigations into this distribution of eigenvalues may yield stronger results for bounding the spectrum of a general complex matrix.

REFERENCES

[1] Taussky, O., "On the Variation of the Characteristic Roots of a Finite Matrix Under Various Changes of Its Elements," in *Recent Advances in Matrix Theory*, Schneider, H. (ed.), The University of Wisconsin Press, Madison (1964).

[2] Golub, G., and Van Loan, C., *Matrix Computations*, 2nd Ed., The Johns Hopkins University Press, Baltimore (1989).

THE AUTHOR

CHARLES H. ROMINE is a visiting scientist at APL's Milton S. Eisenhower Research Center in the Computational Physics Group, where his responsibilities are to promote the use of advanced-architecture computing for computational physics problems such as electromagnetic wave scattering. He received a B.A. in mathematics in 1979 and a Ph.D. in applied mathematics in 1986, both from the University of Virginia. He has recently returned to the Oak Ridge National Laboratory, where he is a research staff member in the Computer Science Group of the Engineering Physics and Mathematics Division. Dr. Romine's specialties are in numerical analysis (particularly numerical linear algebra) and parallel computation. He also holds an adjunct associate professorship at the University of Tennessee in the Computer Science Department.