INCLUSION REGIONS FOR MATRIX EIGENVALUES

CHRISTOPHER BEATTIE* AND ILSE C.F. IPSEN†

Abstract. We review Lehmann's inclusion bounds and provide extensions to general (non-normal) matrices. Each inclusion region has a diameter related to the singular values of a restriction of the matrix to a subspace and dependent on either an eigenvector condition number or the departure of the matrix from normality. The inclusion regions are optimal for normal matrices. Similar considerations lead to inclusion bounds based on relative distances expressed analogously in terms of appropriately defined generalized singular values.

Key words. eigenvalue bounds, singular values, inclusion regions

AMS subject classification. 15A18, 15A42, 15A57, 65F15

1. Introduction.

An "inclusion region" for matrix eigenvalues refers to a region in the complex plane guaranteed to contain a given number of eigenvalues. Geršgorin discs [5, §6.1] and residual bounds [5, Theorem 6.3.14] are well-known examples of such regions (containing respectively, all the eigenvalues and at least one). Generalizations of residual bounds that allow localization of a selected number of eigenvalues date back to Temple [10] and Lehmann [6] and similar notions have resurfaced recently with reference to harmonic Ritz values (see, e.g. [1]).

These generalizations have focussed on the self-adjoint (Hermitian) case. What extensions may be made to non-Hermitian problems ? Harrell [3], for example, has extended Temple's bounds to bounded normal operators. Here we focus on Lehmann's inclusion regions and extend them to general non-normal matrices (in the special case of Hermitian matrices, we recover Lehmann's bounds). We provide two types of inclusion regions, absolute (based on absolute distances of the eigenvalues from a given point) and relative (based on relative distances). Absolute inclusion regions will be disks whose radii are expressed in terms of singular values of the restriction of the matrix to a subspace, and the condition number of an eigenvector matrix or the departure of the matrix from normality (§2). Analogously, we derive in §4 relative inclusion regions whose size depend on generalized singular values and provide for example, relative inclusion intervals for eigenvalues of positive definite matrices with a known factorization. We also give expressions in §3 for absolute inclusion regions that are unbounded *exteriors* of disks and for annular *exclusion* regions (guaranteed to exclude a given number of eigenvalues).

Notation. The conjugate transpose of a complex matrix $\mathbf{A}$ is $\mathbf{A}^*$, and the identity matrix $\mathbf{I}$. The range of $\mathbf{A}$ is denoted by $Ran(\mathbf{A})$. The norm $\|\cdot\|$ is the Euclidean two-norm, and $cond(\mathbf{X}) \equiv \|\mathbf{X}\| \|\mathbf{X}^{-1}\|$ is the two-norm condition number with respect to inversion of a non-singular matrix $\mathbf{X}$.

A Schur decomposition of a complex square matrix $\mathbf{A}$ is $\mathbf{A} = \mathbf{V}(\Lambda + \mathbf{N})\mathbf{V}^*$, where $\mathbf{V}$ is unitary, $\mathbf{N}$ is strictly upper triangular, and Λ is a diagonal matrix whose diagonal elements λ_i are the eigenvalues of $\mathbf{A}$. The two-norm *departure of* $\mathbf{A}$ *from normality* is $\delta(\mathbf{A}) \equiv \min \|\mathbf{N}\|$, where the minimum ranges over all Schur decompositions of $\mathbf{A}$

*Department of Mathematics, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061-0123, USA (`beattie@math.vt.edu`).

†Center for Research in Scientific Computation, Department of Mathematics, North Carolina State University, P. O. Box 8205, Raleigh, NC 27695-8205, USA (`ipsen@math.ncsu.edu`, `http://www4.ncsu.edu/~ipsen/info.html`). This research was supported in part by NSF grant DMS 9714811.

[4, §1.2], [9, §IV.1.2]. If $\mathbf{A}$ is diagonalizable it also has an eigenvalue decomposition $\mathbf{A} = \mathbf{X}\Lambda\mathbf{X}^{-1}$, where Λ is a diagonal matrix whose diagonal elements λ_i are the eigenvalues of $\mathbf{A}$.

The singular values of a tall, skinny $m \times n$ matrix $\mathbf{B}$, $m \geq n$, are labeled in order of decreasing magnitude,

$$\sigma_1(\mathbf{B}) \geq \sigma_2(\mathbf{B}) \geq \cdots \geq \sigma_m(\mathbf{B}),$$

or increasing magnitude,

$$\sigma_{-1}(\mathbf{B}) \leq \sigma_{-2}(\mathbf{B}) \leq \cdots \leq \sigma_{-m}(\mathbf{B}).$$

2. Absolute Inclusion Regions. Although Lehmann's original work [6, 7], [8, §10-5] concerns optimal inclusion intervals for Hermitian matrices, much of his analysis applies to normal matrices as well. We extend Lehmann's results first to diagonalizable and then to general non-normal matrices, in such a way that our regions reduce to Lehmann's in the special case of Hermitian matrices. We show that our inclusion regions are optimal for normal matrices.

The singular values of a matrix, $\mathbf{B}$, have variational characterizations [5, Theorem 7.3.10] that underlie all of the results that follow:

$$\sigma_{-i}(\mathbf{B}) = \min_{\dim(\mathcal{P})=i,}\ \max_{\mathbf{x}\in\mathcal{P}} \frac{||\mathbf{Bx}||}{||\mathbf{x}||} \quad \text{and} \quad \sigma_i(\mathbf{B}) = \max_{\dim(\mathcal{P})=i,}\ \min_{\mathbf{x}\in\mathcal{P}} \frac{||\mathbf{Bx}||}{||\mathbf{x}||}, \tag{2.1}$$

where the leading minimum or maximum ranges over all i-dimensional subspaces $\mathcal{P}$ of $\mathbb{C}^n$.

We first derive inclusion regions for a diagonalizable matrix, $\mathbf{A}$, in terms of singular values of the restriction of $\mathbf{A}$ to a subspace determined by $\mathbf{S}$, and the eigenvector condition number, $cond(\mathbf{X}) \equiv ||\mathbf{X}||\,||\mathbf{X}^{-1}||$.

THEOREM 2.1 (Diagonalizable Matrices). *Let $\mathbf{A}$ be an $n \times n$ diagonalizable matrix with an eigenvalue decomposition $\mathbf{A} = \mathbf{X}\Lambda\mathbf{X}^{-1}$. For a given complex number ρ and an $n \times m$ matrix $\mathbf{S}$ with orthonormal columns, denote by $\tau_{-1} \leq \tau_{-2} \leq \tau_{-3} \leq \ldots$ the singular values of the $n \times m$ matrix $(\mathbf{A} - \rho\mathbf{I})\mathbf{S}$.*

Then each disk

$$\{z : \ |z - \rho| \leq cond(\mathbf{X})\, \tau_{-i}\}, \qquad 1 \leq i \leq m$$

contains at least i eigenvalues of $\mathbf{A}$.

Proof. Label the eigenvalues of $\mathbf{A}$ according to increasing distance from ρ:

$$|\lambda_1 - \rho| \leq |\lambda_2 - \rho| \leq \ldots .$$

The singular values of $\Lambda - \rho = \mathbf{X}^{-1}(\mathbf{A} - \rho\mathbf{I})\mathbf{X}$ are $|\lambda_i - \rho|$. The variational characterization (2.1) yields

$$\begin{aligned}
|\lambda_i - \rho| &= \min_{\dim(\mathcal{P})=i,}\ \max_{\mathbf{x}\in\mathcal{P}} \frac{||(\Lambda - \rho\mathbf{I})\mathbf{x}||}{||\mathbf{x}||} \\
&= \min_{\dim(\mathcal{Q})=i,}\ \max_{\mathbf{y}\in\mathcal{Q}} \frac{||\mathbf{X}^{-1}(\mathbf{A} - \rho\mathbf{I})\mathbf{y}||}{||\mathbf{X}^{-1}\mathbf{y}||} \\
&\leq ||\mathbf{X}^{-1}||\,||\mathbf{X}|| \min_{\dim(\mathcal{Q})=i,}\ \max_{\mathbf{y}\in\mathcal{Q}} \frac{||(\mathbf{A} - \rho\mathbf{I})\mathbf{y}||}{||\mathbf{y}||} = cond(\mathbf{X})\ \sigma_{-i}(\mathbf{A} - \rho\mathbf{I}).
\end{aligned}$$

The second line above follows from the substitutions $\mathbf{y} = \mathbf{X}\mathbf{x}$ and $\mathcal{Q} = \mathbf{X}\mathcal{P}$. Now let $\mathcal{S} \equiv Ran(\mathbf{S})$. Observe that $\dim(\mathcal{S}) = m$ and if the minimization characterizing σ_{-i} in (2.1) for $i \leq m$ occurs instead over the smaller class of subspaces $\mathcal{Q} \subset \mathcal{S}$, a potentially larger minimizing value is obtained:

$$
\begin{aligned}
\sigma_{-i}(\mathbf{A} - \rho\mathbf{I}) &\leq \min_{\substack{\dim(\mathcal{Q})=i \\ \mathcal{Q}\subset\mathcal{S}}} \max_{\mathbf{x}\in\mathcal{Q}} \frac{||(\mathbf{A} - \rho\mathbf{I})\mathbf{x}||}{||\mathbf{x}||} \\
&= \min_{\dim(\mathcal{R})=i} \max_{y\in\mathcal{R}} \frac{||(\mathbf{A} - \rho\mathbf{I})\mathbf{S}\mathbf{y}||}{||\mathbf{S}\mathbf{y}||} \\
&= \min_{\dim(\mathcal{R})=i} \max_{y\in\mathcal{R}} \frac{||(\mathbf{A} - \rho\mathbf{I})\mathbf{S}\mathbf{y}||}{||\mathbf{y}||} = \tau_{-i}.
\end{aligned}
$$

The second line above comes from the observation that if $\dim(\mathcal{Q}) = i \leq m$ and $\mathcal{Q} \subset \mathcal{S}$, then $\mathcal{Q} = \mathbf{S}\mathcal{R}$ for some i-dimensional subspace $\mathcal{R}$ of $\mathbb{C}^m$. The third line follows from the orthonormality of the columns of $\mathbf{S}$ which yields $||\mathbf{S}\mathbf{y}|| = ||\mathbf{y}||$. One may immediately conclude that

$$
|\lambda_1 - \rho| \leq |\lambda_2 - \rho| \leq \ldots \leq |\lambda_i - \rho| \leq cond(\mathbf{X})\, \sigma_{-i}(\mathbf{A} - \rho\mathbf{I}) \leq cond(\mathbf{X})\, \tau_{-i}.
$$

The i eigenvalues $\{\lambda_1, \lambda_2, \ldots, \lambda_i\}$ are in the disk $\{z : \ |z - \rho| \leq cond(\mathbf{X})\, \tau_{-i}\}$. □

Since Theorem 2.1 is based on an eigenvalue decomposition, it yields no information if $\mathbf{A}$ is not diagonalizable. Instead of an eigenvalue decomposition for $\mathbf{A}$, we can use a Schur decomposition of $\mathbf{A}$ to derive inclusion regions for general, possibly defective matrices. The departure from normality, $\delta(\mathbf{A})$, then plays a role in the bounds analogous to $cond(\mathbf{X})$ in Theorem 2.1.

THEOREM 2.2 (General Matrices). *Let $\mathbf{A}$ be an $n \times n$ matrix. For a given complex number ρ and an $n \times m$ matrix $\mathbf{S}$ with orthonormal columns, denote by $\tau_{-1} \leq \tau_{-2} \leq \tau_{-3} \leq \ldots$ the singular values of the $n \times m$ matrix $(\mathbf{A} - \rho\mathbf{I})\mathbf{S}$.*

Then each disk

$$
\{z : \ |z - \rho| \leq \tau_{-i} + \delta(\mathbf{A})\}, \qquad 1 \leq i \leq m,
$$

contains at least i eigenvalues of $\mathbf{A}$.

Proof. As before, label eigenvalues of $\mathbf{A}$ according to increasing distance from ρ:

$$
|\lambda_1 - \rho| \leq |\lambda_2 - \rho| \leq \ldots
$$

For any Λ and $\mathbf{N}$ arising from a Schur decomposition of $\mathbf{A} = \mathbf{V}(\Lambda + \mathbf{N})\mathbf{V}^*$,

$$
|\lambda_i(\mathbf{A}) - \rho| = \sigma_{-i}(\Lambda - \rho\mathbf{I}) = \sigma_{-i}(\Lambda + \mathbf{N} - \rho\mathbf{I} - \mathbf{N}) \leq \sigma_{-i}(\mathbf{A} - \rho\mathbf{I}) + ||\mathbf{N}||,
$$

where the inequality follows by observing that for any nontrivial $\mathbf{x}$

$$
\frac{||(\Lambda + \mathbf{N} - \rho\mathbf{I} - \mathbf{N})\mathbf{x}||}{||\mathbf{x}||} \leq \frac{||(\Lambda + \mathbf{N} - \rho\mathbf{I})\mathbf{x}||}{||\mathbf{x}||} + \frac{||\mathbf{N}\mathbf{x}||}{||\mathbf{x}||} \leq \frac{||(\Lambda + \mathbf{N} - \rho\mathbf{I})\mathbf{x}||}{||\mathbf{x}||} + ||\mathbf{N}||
$$

and applying the variational characterization (2.1) for $\sigma_{-i}(\Lambda + \mathbf{N} - \rho\mathbf{I}) = \sigma_{-i}(\mathbf{A} - \rho\mathbf{I})$ (c.f., [2, Corollary 8.6.2]).

Following the reasoning in the proof of Theorem 2.1, $\sigma_{-i}(\mathbf{A} - \rho\mathbf{I}) \leq \tau_{-i}$ and one may conclude

$$
|\lambda_1 - \rho| \leq |\lambda_2 - \rho| \leq \ldots \leq |\lambda_i - \rho| \leq \sigma_{-i}(\mathbf{A} - \rho\mathbf{I}) + ||\mathbf{N}|| \leq \tau_{-i} + ||\mathbf{N}||.
$$

Since this is true for any of the possible Schur decompositions for $\mathbf{A}$, it must happen that at least i eigenvalues $\{\lambda_1, \lambda_2, \ldots, \lambda_i\}$ lie in the disk $\{z: |z-\rho| \leq \tau_{-i} + \delta(\mathbf{A})\}$. $\square$

Theorems 2.1 and 2.2 imply, in particular, that the largest disks

$$\{z: |z-\rho| \leq \|(\mathbf{A}-\rho\mathbf{I})\mathbf{S}\| + \delta(\mathbf{A})\}$$

$$\{z: |z-\rho| \leq cond(\mathbf{X})\ \|(\mathbf{A}-\rho\mathbf{I})\mathbf{S}\|\} \quad \text{(if } \mathbf{A} \text{ is diagonalizable)}$$

each contain at least m eigenvalues of $\mathbf{A}$.

In the special case when $\mathbf{A}$ is normal, the bounds in Theorems 2.1 and 2.2 simplify.

COROLLARY 2.3 (Normal Matrices). *Let $\mathbf{A}$ be a normal $n \times n$ matrix. Let ρ be a complex number and $\mathbf{S}$ a $n \times m$ matrix with orthonormal columns, and $\tau_{-1} \leq \ldots \leq \tau_{-m}$ the singular values of $(\mathbf{A}-\rho\mathbf{I})\mathbf{S}$.*

Then each disk

$$\{z: |z-\rho| \leq \tau_{-i}\}, \qquad 1 \leq i \leq m$$

contains at least i eigenvalues of $\mathbf{A}$.

The following two examples show that the inclusion regions can be arbitrarily pessimistic when $\mathbf{S}$ is far away from an invariant subspace or when $\mathbf{A}$ is non-normal.

EXAMPLE 1 (Normal Matrices). *Inclusion regions for normal matrices can be arbitrarily pessimistic (i.e., large) when $\mathbf{S}$ is far away from an invariant subspace associated with m eigenvalues closest to ρ.*

For instance, let

$$\mathbf{A} = \begin{bmatrix} 1 & & \\ & 2 & \\ & & \eta \end{bmatrix}, \qquad \eta \geq 2.$$

Choose $\rho = 0$ and $\mathbf{S} = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}^T$. Then $|\lambda_1 - \rho| = 1$, but the bound for the inclusion region from Corollary 2.3 is $|z-\rho| \leq \eta$, which can become arbitrarily large as η increases.

EXAMPLE 2 (Non-Normal Matrices). *Inclusion regions for non-normal matrices can be arbitrarily pessimistic, even if $\mathbf{S}$ consists of eigenvectors for m eigenvalues closest to ρ.*

Let

$$\mathbf{A} = \begin{bmatrix} 1 & & \\ & 2 & \eta \\ & & 3 \end{bmatrix}, \qquad \eta \geq 0.$$

Here $\delta(\mathbf{A}) = \eta$ and $cond(\mathbf{X}) = \sqrt{1 + 2\eta(\eta + \sqrt{1+\eta^2})} \sim 2\eta + \mathcal{O}(1/\eta)$ as $\eta \to \infty$. Choose $\rho = 0$ and

$$\mathbf{S} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix},$$

whose columns are eigenvectors for the eigenvalues $\lambda_1 = 1$ and $\lambda_2 = 2$. Then $|\lambda_1 - \rho| = 1$ and $|\lambda_2 - \rho| = 2$, but the bounds for the inclusion regions from Theorems 2.1 and 2.2 are, respectively,

$$|z-\rho| \leq 4\eta + \mathcal{O}(1/\eta), \quad \text{as } \eta \to \infty \qquad \text{and} \qquad |z-\rho| \leq 2 + \eta,$$

which becomes arbitrarily large as the departure from normality η increases.

The inclusion regions in Corollary 2.3 are optimal for Hermitian matrices in the following sense: Given τ_{-i} there exists a Hermitian matrix whose eigenvalues lie on the boundaries of the inclusion regions [8, Theorem 10-5-3]. The result below shows that this is also true for normal matrices.

THEOREM 2.4 (Optimality). *Let $\mathbf{A}$ be a normal $n \times n$ matrix. Let ρ be a complex number and $\mathbf{S}$ a $n \times m$ matrix with orthonormal columns, and $\tau_{-1} \leq \ldots \leq \tau_{-m}$ the singular values of $(\mathbf{A} - \rho\mathbf{I})\mathbf{S}$.*

Then there exists a normal matrix $\mathbf{B}$ such that $(\mathbf{B} - \rho\mathbf{I})\mathbf{S}$ has singular values τ_{-i} and the eigenvalues $\lambda_i(\mathbf{B})$ of $\mathbf{B}$, ordered by increasing distance from ρ, satisfy

$$|\lambda_i(\mathbf{B}) - \rho| = \tau_{-i}, \qquad 1 \leq i \leq m,$$

while the remaining $n-m$ eigenvalues of $\mathbf{B}$ are not contained in any of these m disks.

Proof. Let $\begin{bmatrix}\mathbf{S} & \tilde{\mathbf{S}}\end{bmatrix}$ be unitary, $\mathbf{D}$ a diagonal matrix of order m whose ith diagonal element is $\rho + \tau_{-i}$, and $\tilde{\mathbf{D}} = (\rho + \tau_{-m} + \epsilon)\mathbf{I}_{n-m}$ for some $\epsilon > 0$. Set

$$B \equiv \begin{bmatrix}\mathbf{S} & \tilde{\mathbf{S}}\end{bmatrix} \begin{bmatrix}\mathbf{D} & \\ & \tilde{\mathbf{D}}\end{bmatrix} \begin{bmatrix}\mathbf{S} & \tilde{\mathbf{S}}\end{bmatrix}^* .$$

□

3. Extensions. We give expressions for absolute inclusion regions that are unbounded *exteriors* of disks, annular *exclusion* regions (guaranteed to not include a given number of eigenvalues), inclusion regions when the subspace basis $\mathbf{S}$ is not orthonormal, and finally, inclusion intervals for eigenvalues of positive definite matrices with a known factorization.

The first result presents inclusion regions complementary to the disks of Theorems 2.1 and 2.2.

THEOREM 3.1 (Exterior Inclusion). *Let $\mathbf{A}$ be an $n \times n$ matrix. Let ρ be a complex number and $\mathbf{S}$ a $n \times m$ matrix with orthonormal columns, and $\tau_1 \geq \ldots \geq \tau_m$ the singular values of $(\mathbf{A} - \rho\mathbf{I})\mathbf{S}$.*

Then each (unbounded) region

$$\{z : \ |z - \rho| \geq \tau_i - \delta(\mathbf{A})\}, \qquad 1 \leq i \leq m$$

contains at least i eigenvalues of $\mathbf{A}$.

If, in addition, $\mathbf{A}$ is diagonalizable, then each region

$$\{z : \ |z - \rho| \geq \tau_i / \, cond(\mathbf{X})\}, \qquad 1 \leq i \leq m$$

contains at least i eigenvalues of $\mathbf{A}$.

Proof. The key observation is that $\tau_i \leq \sigma_i(\mathbf{A} - \rho\mathbf{I})$ which may be deduced from the variational characterization for σ_i in (2.1) by restricting the maximization over subspaces, $\mathcal{P}$, of dimension i that additionally satisfy $\mathcal{P} \subset Ran(\mathbf{S})$. In other respects, the proof is similar to that of Theorems 2.1 and 2.2 and details are omitted. □

Combining these results with the inclusion regions of Theorems 2.1 and 2.2 yields annuli that contain no more than a specified number of eigenvalues.

COROLLARY 3.2 (Annular Exclusion). *Let $\mathbf{A}$ be an $n \times n$ matrix. Let ρ be a complex number and $\mathbf{S}$ a $n \times m$ matrix with orthonormal columns, and $\tau_1 = \tau_{-m} \geq \ldots \geq \tau_m = \tau_{-1}$ the singular values of $(\mathbf{A} - \rho\mathbf{I})\mathbf{S}$.*

Then each annulus

$$\{z:\ \tau_{-i}+\delta(\mathbf{A})<|z-\rho|<\tau_j-\delta(\mathbf{A})\},\qquad 1\leq i+j\leq m$$

contains at most $n-(i+j)$ *eigenvalues of* $\mathbf{A}$, *provided* $2\delta(\mathbf{A})<\tau_j-\tau_{-i}$.

If, in addition, $\mathbf{A}$ *is diagonalizable then each annulus*

$$\{z:\ cond(\mathbf{X})\ \tau_{-i}<|z-\rho|<\tau_j/\ cond(\mathbf{X})\},\qquad 1\leq i+j\leq m$$

contains at most $n-(i+j)$ *eigenvalues of* $\mathbf{A}$, *provided* $\tau_{-i}/\tau_j<1/\ cond(\mathbf{X})^2$.

Proof. An exterior region $|z-\rho|\geq\tau_j-\delta(\mathbf{A})$ in Theorem 3.1 contains at least j eigenvalues, hence its complement $|z-\rho|<\tau_j-\delta(\mathbf{A})$ contains at most $n-j$ eigenvalues. An interior region $|z-\rho|\leq\tau_{-i}+\delta(\mathbf{A})$ in Theorem 2.2 contains at least i eigenvalues. Since an annulus is the complement of the exterior region minus the interior region, it contains at most $(n-j)-i$ eigenvalues of $\mathbf{A}$. To make the lower bound strictly less than the upper we need two things. First $\tau_{-i}=\tau_{m+1-i}<\tau_j$, that is, $j<m+1-i$ or $i+j\leq m$; and second $\tau_{-i}+\delta(\mathbf{A})<\tau_j-\delta(\mathbf{A})$ or $2\delta(\mathbf{A})<\tau_j-\tau_{-i}$.

The proof for diagonalizable matrices is similar. □

Now we consider subspace bases $\mathbf{S}$ that are not orthonormal. In this case we have to resort to generalized singular values to express the inclusion regions. The *generalized singular values* σ of a matrix pair $(\mathbf{S},\,\mathbf{T})$ are the nonnegative square roots of the eigenvalues σ^2 of the generalized eigenvalue problem $\mathbf{S}^*\mathbf{S}\mathbf{z}=\sigma^2\mathbf{T}^*\mathbf{T}\mathbf{z}$ [2, §8.7].

THEOREM 3.3 (General Subspace Bases). *Let* $\mathbf{M}$ *be a nonsingular* $n\times n$ *matrix. Let* ρ *be a complex number and* $\mathbf{S}$ *a* $n\times m$ *matrix with* $\mathrm{rank}(\mathbf{S})=m$, *and* $\tau_{-1}\leq\ldots\leq\tau_{-m}$ *the generalized singular values of* $((\mathbf{A}-\rho\mathbf{I})\mathbf{M}\mathbf{S},\ \mathbf{M}\mathbf{S})$.

Then each disk

$$\{z:\ |z-\rho|\leq\tau_{-i}+\delta(\mathbf{A})\},\qquad 1\leq i\leq m$$

contains at least i *eigenvalues of* $\mathbf{A}$.

If, in addition, $\mathbf{A}$ *is diagonalizable then each disk*

$$\{z:\ |z-\rho|\leq cond(\mathbf{X})\,\tau_{-i}\},\qquad 1\leq i\leq m$$

contains at least i *eigenvalues of* $\mathbf{A}$.

Proof. Let σ be a generalized singular value of $((\mathbf{A}-\rho\mathbf{I})\mathbf{M}\mathbf{S},\ \mathbf{M}\mathbf{S})$. Then

$$[(\mathbf{A}-\rho\mathbf{I})\mathbf{M}\mathbf{S}]^*\,(\mathbf{A}-\rho\mathbf{I})\mathbf{M}\mathbf{S}\,z=\sigma^2[\mathbf{M}\mathbf{S}]^*\mathbf{M}\mathbf{S}\,z.$$

If $\mathbf{M}\mathbf{S}=\mathbf{Q}\mathbf{R}$ is a QR decomposition, where $\mathbf{Q}$ has orthonormal columns and $\mathbf{R}$ is non-singular then the above eigenvalue problem is equivalent to

$$\mathbf{Q}^*(\mathbf{A}-\rho\mathbf{I})^*(\mathbf{A}-\rho\mathbf{I})\mathbf{Q}y=\sigma^2y.$$

Hence the singular values of $(\mathbf{A}-\rho\mathbf{I})\mathbf{Q}$ are identical to the generalized singular values of $((\mathbf{A}-\rho\mathbf{I})\mathbf{M}\mathbf{S},\ \mathbf{M}\mathbf{S})$. Now apply Theorems 2.1 and 2.2 to $(\mathbf{A}-\rho\mathbf{I})\mathbf{Q}$. □

Different choices of $\mathbf{M}$ lead to different variants of Lehmann's bounds for Hermitian matrices. If $\mathbf{A}$ is Hermitian then the choice $\mathbf{M}=\mathbf{I}$ produces 'right-definite' Lehmann bounds, and when $\rho=0$, the τ_{-i} are harmonic Ritz values. If $\mathbf{A}$ is Hermitian positive definite, the choice $\mathbf{M}=\mathbf{A}^{-1/2}$ produces 'left-definite' Lehmann bounds, and when $\rho=0$, the τ_{-i} are dual harmonic Ritz values. Further discussion, comparisons, and references can be found in [1].

4. Relative Inclusion Regions. For simplicity, we present relative inclusion regions only for diagonalizable, non-singular matrices. We find inclusions for eigenvalues of $\mathbf{A}$ that are within the same *relative* distance, say $r > 0$, from ρ: that is, within the set

$$\mathcal{I}_r = \left\{ z : \frac{|z-\rho|}{|z|} \le r \right\}.$$

Unlike the disks of §2, the geometry of $\mathcal{I}_r$ can change change dramatically with r. If $0 < r < 1$ then $\mathcal{I}_r$ is a closed disk with radius $|\rho| r/(1-r^2)$ centered at $\rho/(1-r^2)$. If $r > 1$ then $\mathcal{I}_r$ is the (unbounded) complement of the open disk with radius $|\rho| r/(r^2-1)$ centered at $-\rho/(r^2-1)$. For the remaining exceptional case $r = 1$, $\mathcal{I}_r$ is a closed half plane containing ρ with boundary passing through $\rho/2$ and normal to the ray extending from 0 to ρ.

Factorizations of $\mathbf{A}$ play a role in our analysis. Notice that if an invertible matrix $\mathbf{A}$ is written as the product of square matrices $\mathbf{A} = \mathbf{B}_1\mathbf{B}_2$ then $\mathbf{A}$ is similar to $\mathbf{C} = \mathbf{B}_1^{-1}\mathbf{A}\mathbf{B}_1 = \mathbf{B}_2\mathbf{A}\mathbf{B}_2^{-1} = \mathbf{B}_2\mathbf{B}_1$. If $\mathbf{A}$ has an eigenvalue decomposition $\mathbf{A} = \mathbf{X}\Lambda\mathbf{X}^{-1}$ then $\mathbf{C}$ is diagonalized by $\tilde{\mathbf{X}} = \mathbf{B}_1^{-1}\mathbf{X}\Delta_1 = \mathbf{B}_2\mathbf{X}\Delta_2$. Δ_1 and Δ_2 are nonsingular diagonal matrices that are chosen so that $\Delta_1 = \Delta_2\Lambda$ and so as to scale the columns of $\mathbf{B}_1^{-1}\mathbf{X}$ and $\mathbf{B}_2\mathbf{X}$, respectively, to have unit norm. $cond(\tilde{\mathbf{X}})$ plays a role in the derived relative bounds.

THEOREM 4.1. *Let $\mathbf{A}$ be an $n \times n$ diagonalizable nonsingular matrix with an eigenvalue decomposition $\mathbf{A} = \mathbf{X}\Lambda\mathbf{X}^{-1}$. Suppose that a factorization of $\mathbf{A}$ into the product of square matrices $\mathbf{A} = \mathbf{B}_1\mathbf{B}_2$ is given. For a given complex number ρ and an $n \times m$ matrix $\mathbf{S}$ with $\operatorname{rank}(\mathbf{S}) = m$, let $\tau_{-1} \le \tau_{-2} \le \ldots$ denote the generalized singular values of the matrix pair*

$$(\mathbf{B}_1^{-1}(\mathbf{A}-\rho\mathbf{I})\mathbf{S}, \mathbf{B}_2\mathbf{S}). \tag{4.1}$$

For each $1 \le i \le m$, the region

$$\mathcal{I}_{cond(\tilde{\mathbf{X}})\tau_{-i}} = \{z : |z-\rho| \le cond(\tilde{\mathbf{X}})\, \tau_{-i}\, |z|\}$$

contains at least i eigenvalues of $\mathbf{A}$ where $\tilde{\mathbf{X}}$ is as described above.

Proof. Label the eigenvalues of $\mathbf{A}$ in order of increasing relative distance from ρ,

$$\left|\frac{\lambda_1-\rho}{\lambda_1}\right| \le \left|\frac{\lambda_2-\rho}{\lambda_2}\right| \le \ldots .$$

These are precisely the singular values of $\mathbf{I} - \rho\Lambda^{-1}$; $\left|\frac{\lambda_i-\rho}{\lambda_i}\right| = \left|1 - \frac{\rho}{\lambda_i}\right|$. For each $i \ge 1$, the pattern of proof of Theorem 2.1 can be repeated to show that each of the regions

$$\{z : |1 - \frac{\rho}{z}| \le cond(\tilde{\mathbf{X}})\, \sigma_{-i}(\mathbf{I} - \rho\mathbf{C}^{-1})$$

contains at least i eigenvalues of $\mathbf{C}$ (and hence of $\mathbf{A}$). $\mathbf{B}_1^{-1}(\mathbf{A}-\rho\mathbf{I})\mathbf{B}_2^{-1} = \mathbf{I} - \rho\mathbf{C}^{-1}$, so a change of variable argument as in Theorem 3.3 leads to $\sigma_{-i}(\mathbf{I} - \rho\mathbf{C}^{-1}) \le \tau_{-i}$ and the conclusion. □

Notice that the region we describe depends on $cond(\tilde{\mathbf{X}})$, which may be different from $cond(\mathbf{X})$. However, if either of the factors $\mathbf{B}_1$ or $\mathbf{B}_2$ is unitary then $\mathbf{C}$ is unitarily similar to $\mathbf{A}$ and $cond(\tilde{\mathbf{X}}) = cond(\mathbf{X})$. If, in addition, $\mathbf{A}$ is normal then $\mathbf{C}$ also is normal and $cond(\tilde{\mathbf{X}}) = cond(\mathbf{X}) = 1$.

The two cases of either $\mathbf{B}_1$ or $\mathbf{B}_2$ being unitary can produce only two sets of possible bounds, independent of the specific unitary factorization $\mathbf{A} = \mathbf{B}_1\mathbf{B}_2$. In particular, if $\mathbf{A}$ has been factored so that $\mathbf{B}_1$ is unitary, then the generalized singular values of (4.1) are exactly the generalized singular values of the matrix pair $((\mathbf{A} - \rho\mathbf{I})\mathbf{S}, \mathbf{AS})$. If instead, the factorization has been chosen so that $\mathbf{B}_2$ is unitary, then the generalized singular values of (4.1) will be the generalized singular values of the matrix pair $\left(\mathbf{A}^{-1}(\mathbf{A} - \rho\mathbf{I})\mathbf{S}, \mathbf{S}\right)$. Application of Theorem 4.1 in ether case leads to:

Corollary 4.2. *Let* $\mathbf{A}$ *be an* $n \times n$ *diagonalizable invertible matrix with an eigenvalue decomposition* $\mathbf{A} = \mathbf{X}\Lambda\mathbf{X}^{-1}$. *For a given complex number* ρ *and an* $n \times m$ *matrix* $\mathbf{S}$ *with* $\mathrm{rank}(\mathbf{S}) = m$, *let* $\tau_{-1} \le \tau_{-2} \le \ldots$ *denote the generalized singular values of either one of the matrix pairs*

$$((\mathbf{A} - \rho\mathbf{I})\mathbf{S}, \mathbf{AS}) \qquad or \qquad \left(\mathbf{A}^{-1}(\mathbf{A} - \rho\mathbf{I})\mathbf{S}, \mathbf{S}\right).$$

Then for each $1 \le i \le m$, *the region*

$$\mathcal{I}_{cond(\mathbf{X})\tau_{-i}} = \{z : \ |z - \rho| \le cond(\mathbf{X})\, \tau_{-i}\, |z|\}$$

contains at least i *eigenvalues of* $\mathbf{A}$.

One might consider preconditioning $\mathbf{X}$ to reduce the influence of $cond(\mathbf{X})$ in the bound. Say we know an approximate factorization of $\mathbf{X}$: $\mathbf{X} \approx \mathbf{RQ}$ where $\mathbf{R}$ is invertible and $\mathbf{Q}$ is unitary. Then one might expect that $cond(\mathbf{R}^{-1}\mathbf{X}) \ll cond(\mathbf{X})$ and factorization of $\mathbf{A}$ as $\mathbf{A} = \mathbf{R}(\mathbf{R}^{-1}\mathbf{A})$ could lead to tighter bounds:

Corollary 4.3. *Let* $\mathbf{A}$ *be an* $n \times n$ *nonsingular diagonalizable matrix with an eigenvalue decomposition* $\mathbf{A} = \mathbf{X}\Lambda\mathbf{X}^{-1}$. *Suppose that* $\mathbf{R}$ *is nonsingular. If* $\tau_{-1} \le \tau_{-2} \le \ldots$ *denote the generalized singular values of the matrix pair*

$$\left(\mathbf{R}^{-1}(\mathbf{A} - \rho\mathbf{I})\mathbf{S}, \mathbf{R}^{-1}\mathbf{AS}\right),$$

then for each $1 \le i \le m$, *the region*

$$\mathcal{I}_{cond(\mathbf{R}^{-1}\mathbf{X})\tau_{-i}} = \{z : \ |z - \rho| \le cond(\mathbf{R}^{-1}\mathbf{X})\, \tau_{-i}\, |z|\}$$

contains at least i *eigenvalues of* $\mathbf{A}$.

Of course, one needs some mechanism to estimate $cond(\mathbf{R}^{-1}\mathbf{X})$ in order to implement concrete bounds.

If $\mathbf{A}$ is positive definite and Hermitian, the additional structure can be exploited.

Theorem 4.4. *Suppose* $\mathbf{A}$ *is positive definite and decomposed as* $\mathbf{A} = \mathbf{T}^*\mathbf{T}$ *for some* $\mathbf{T} \in \mathbb{C}^{n \times n}$. *For a given real number* $\rho > 0$ *and an* $n \times m$ *matrix* $\mathbf{S}$ *with* $\mathrm{rank}(\mathbf{S}) = m$, *let* $\tau_{-1} \le \tau_{-2} \le \ldots$ *denote the generalized singular values of the matrix pair*

$$((\mathbf{T} - \rho\mathbf{T}^{-*})\mathbf{S}, \mathbf{TS}).$$

Then for each $1 \le i \le m$ *for which* $\tau_{-i} < 1$, *the interval* $\left[\dfrac{\rho}{1 + \tau_{-i}}, \dfrac{\rho}{1 - \tau_{-i}}\right]$ *contains at least* i *eigenvalues of* $\mathbf{A}$. *For each* $1 \le i \le m$ *for which* $\tau_{-i} \ge 1$, *the interval* $\left[\dfrac{\rho}{1 + \tau_{-i}}, \infty\right)$ *contains at least* i *eigenvalues of* $\mathbf{A}$.

Proof. Apply Theorem 4.1 with $\mathbf{B}_1 = \mathbf{T}^*$ and $\mathbf{B}_2 = \mathbf{T}$. Since $\mathbf{A}$ is positive definite and Hermitian the eigenvalues are real and positive. The intersection of the positive real half-line with the regions $\mathcal{I}_r$ for $r = \tau_{-i}$ as described at the beginning of the section yields the intervals given and must then include at least i eigenvalues. □

REFERENCES

[1] C. BEATTIE, *Harmonic Ritz and Lehmann bounds*, Electron. Trans. Numer. Anal., 7 (1998), pp. 18–39.

[2] G. GOLUB AND C. VAN LOAN, *Matrix Computations*, The Johns Hopkins University Press, Baltimore, third ed., 1996.

[3] E. HARRELL, *Generalizations of Temple's inequality*, Proc. Amer. Math. Soc., 69 (1978), pp. 271–6.

[4] P. HENRICI, *Bounds for iterates, inverses, spectral variation and fields of values of non-normal matrices*, Numer. Math., 4 (1962), pp. 24–40.

[5] R. HORN AND C. JOHNSON, *Matrix Analysis*, Cambridge University Press, Cambridge, 1985.

[6] N. LEHMANN, *Calculation of eigenvalue bounds in linear problems*, Arch. Math., 2 (1949), pp. 139–47.

[7] ———, *On optimal eigenvalue localization in the solution of symmetric matrix problems*, Num. Math., 8 (1966), pp. 42–55.

[8] B. PARLETT, *The Symmetric Eigenvalue Problem*, Prentice Hall, Englewood Cliffs, 1980.

[9] G. STEWART AND J. SUN, *Matrix Perturbation Theory*, Academic Press, San Diego, 1990.

[10] G. TEMPLE, *The computation of characteristic numbers and characteristic functions*, Proc. London Math. Soc., 29 (1928), pp. 257–80.