DEPARTURE FROM NORMALITY AND EIGENVALUE PERTURBATION BOUNDS

ILSE C.F. IPSEN*

Abstract. Eigenvalue perturbation bounds for the optimal matching distance in the two-norm and in the Frobenius norm are presented. The original matrix must be diagonalizable, but there are no restrictions on the perturbed matrix, it can be defective. Furthermore, Gerschgorin-like inclusion regions in the Frobenius norm are derived, as well as bounds on the departure from normality.

Key words. diagonalizable matrix, normal matrix, optimal matching distance, departure from normality

AMS subject classification. 15A42, 15A18, 15A12, 65F15, 65F35

1. Introduction.

The results in this paper are motivated by two eigenvalue perturbation bounds for normal and Hermitian matrices by Sun and Kahan, respectively.

Sun [11] presents a bound for the eigenvalues of a perturbed normal matrix that does not contain the conditioning of the perturbed eigenvectors. Kahan [8] presents a similar, but better bound for Hermitian matrices.

We use the ideas in Sun's and Kahan's proofs to derive bounds on the departure from normality of complex square matrices $A + E$, where A is normal or Hermitian (§2).

Our eigenvalue perturbation results (§§3, 5, 6) bound the optimal matching distance between eigenvalues λ_j of A and eigenvalues μ_j of $A + E$, in the two-norm [2, §VI.1]

$$\min_\sigma \max_j |\mu_j - \lambda_{\sigma(j)}|,$$

and in the Frobenius norm [2, §VI.6]

$$\min_\sigma \left(\sum_j |\mu_j - \lambda_{\sigma(j)}|^2 \right)^{1/2},$$

where the minimum ranges over all permutations σ. The bounds require that A be diagonalizable, but E can be arbitrary, in particular, $A + E$ can be defective. The bounds do not contain the conditioning (with respect to inversion) of the eigenvectors of $A + E$. We give examples to illustrate why this can be advantageous (§4).

In particular we derive bounds when A is a normal matrix whose eigenvalues lie on a line in the complex plane (§3) and when A is a diagonalizable matrix (§5), and Gerschgorin-like inclusion regions in the Frobenius norm for general complex square matrices (§6).

We also present Frobenius norm bounds for the departure of a matrix from normality in terms of the diagonal and off-diagonal elements (§7). For real matrices, the departure from normality can be bounded by the norm of the off-diagonal elements. Our bounds are simple and cheaper to determine than many existing bounds, cf.

*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/). This research was supported in part by NSF grants DMS-0209931 and DMS-0209695.

[9, 10] and the references in there. We end the paper with a lower bound on eigenvalue perturbations for complex square matrices and for Hermitian matrices (§8).

Notation. We denote by $\|\cdot\|_F$ the Frobenius norm and by $\|\cdot\|_2$ the Euclidean 2-norm. The identity matrix is I, and the imaginary unit is $\imath \equiv \sqrt{-1}$.

2. Departure from Normality of Perturbed Normal Matrices. We bound the departure from normality of perturbed normal and Hermitian matrices.

To this end, partition the complex matrix A into Hermitian and skew-Hermitian parts, $A = \Re(A) + \imath\Im(A)$, where $\imath = \sqrt{-1}$ and

$$\Re(A) \equiv \frac{1}{2}(A + A^*), \quad \Im(A) \equiv \frac{\imath}{2}(A^* - A).$$

THEOREM 2.1. *Let $A + E$ be of order n with Schur decomposition $A + E = Q(M + N)Q^*$, where Q is unitary, M diagonal and N strictly upper triangular.*

If A is normal then

$$\|N\|_F \leq (\sqrt{n-1} + 1)\, \|E\|_F.$$

If A is Hermitian then

$$\|N\|_F \leq \sqrt{2}\|\Im(E)\|_F.$$

If, in addition, $A+E$ is real or $A+E$ has real eigenvalues then $\|N\|_F = \sqrt{2}\|\Im(E)\|_F$.

Proof. In $Q^*AQ + Q^*EQ = M + N$, partition

$$Q^*AQ = A_L + A_D + A_U, \qquad Q^*EQ = E_L + E_D + E_U,$$

where A_U and E_U are strictly upper triangular, A_D and E_D diagonal, and A_L and E_L strictly lower triangular. From $Q^*AQ + Q^*EQ = M + N$ upper triangular follows $A_L + E_L = 0$ and $A_U + E_U = N$. Hence $\|N\|_F \leq \|A_U\|_F + \|E_U\|_F$. Since Q^*AQ is normal, we have [11, Lemma 3.1], [2, Problem VI.8.10] $\|A_U\|_F \leq \sqrt{n-1}\|A_L\|_F$. Therefore

$$\|N\|_F \leq \sqrt{n-1}\|A_L\|_F + \|E_U\|_F = \sqrt{n-1}\|E_L\|_F + \|E_U\|_F \leq (\sqrt{n-1} + 1)\, \|E\|_F.$$

If A is Hermitian then $N = A_U + E_U = A_L^* + E_U = -E_L^* + E_U$. Because the matrix

$$N^* - N = (E_U + E_L)^* - (E_U + E_L)$$

has a zero diagonal, adding a diagonal cannot decrease its norm,

$$\|N^* - N\|_F = \|(E_U + E_L)^* - (E_U + E_L)\|_F \leq \|E^* - E\|_F.$$

Therefore, $\|\Im(N)\|_F \leq \|\Im(E)\|_F$. Since N is strictly upper triangular,

$$\|N\|_F^2 = \|\Re(N)\|_F^2 + \|\Im(N)\|_F^2 = 2\|\Im(N)\|_F^2 \leq 2\|\Im(E)\|_F^2.$$

If $A + E$ has real eigenvalues then $M - Q^*AQ = Q^*EQ - N$ is Hermitian, so $(Q^*EQ - N)^* = Q^*EQ - N$ and $Q^*(E^* - E)Q = N^* - N$. Hence $\|\Im(N)\|_F = \|\Im(E)\|_F$.

If $A + E$ is real then there exists a real orthogonal Q, so that E_D is real. Hence $\Im(E_D) = 0$ and $\|\Im(E_U + E_L)\|_F = \|\Im(E)\|_F$. □

It is not possible to bound $\|N\|_F$ from below by $\|\Im(E)\|_F$. For instance, if A is real diagonal and E is imaginary diagonal then $N = 0$ but $\Im(E) > 0$.

3. Eigenvalue Bounds for Normal Matrices. We use the bounds in §2 to derive eigenvalue perturbations for normal and Hermitian matrices. There are no restrictions on the perturbations, and the bounds do not depend on the conditioning of perturbed quantities. Although our bounds are slightly worse than those by Sun and Kahan, the derivations are simple and illustrate why the non-normality does not appear in the bounds. We also improve a Frobenius-norm bound for normal matrices whose eigenvalues lie on a line in the complex plane.

The bound below for eigenvalues of an arbitrarily perturbed normal matrix does not depend on the conditioning of the perturbed eigenvectors.

THEOREM 3.1. *Let A and $A+E$ be complex matrices of order n with eigenvalues λ_j and μ_j, respectively.*

If A is normal then there is a permutation σ so that

$$\left(\sum_{j=1}^{n} |\mu_j - \lambda_{\sigma(j)}|^2\right)^{1/2} \leq (\sqrt{n-1}+2)\,\|E\|_F.$$

Proof. Let $A+E = Q(M+N)Q^*$ be a Schur decomposition of $A+E$. Applying the Hoffman-Wielandt inequality [6, Theorem 1], [2, Theorem VI.4.1] to A and QMQ^* gives

$$\left(\sum_{j=1}^{n} |\mu_j - \lambda_{\sigma(j)}|^2\right)^{1/2} \leq \|QMQ^* - A\|_F = \|E - QNQ^*\|_F \leq \|E\|_F + \|N\|_F$$

for some permutation σ. Apply Theorem 2.1 to $\|N\|_F$. ◻

Sun presents a slightly better bound, by taking advantage of structure in $E - QNQ^*$.

THEOREM 3.2 (Theorem 1.1 in [11], Problem VI.8.11 in [2]). *If A is normal then there is a permutation σ so that*

$$\left(\sum_{j=1}^{n} |\mu_j - \lambda_{\sigma(j)}|^2\right)^{1/2} \leq \sqrt{n}\|E\|_F.$$

For Hermitian matrices the bounds can be tightened.

THEOREM 3.3. *Let A and $A+E$ be complex matrices of order n with eigenvalues λ_j and μ_j, respectively.*

If A is Hermitian then there is a permutation σ so that

$$\left(\sum_{j=1}^{n} |\mu_j - \lambda_{\sigma(j)}|^2\right)^{1/2} \leq (1+\sqrt{2})\,\|E\|_F.$$

Proof. Proceed as in the proof of Theorem 3.1, but use the bound for Hermitian matrices from Theorem 2.1. ◻

Kahan presents an even tighter bound, where eigenvalues are ordered according to size. Label the eigenvalues of a Hermitian matrix A in ascending order, $\lambda_1 \leq \ldots \leq \lambda_n$.

THEOREM 3.4 (§0, §3 in [8], Problem VI.8.7 in [2]). *If A is Hermitian then*

$$\left(\sum_{j=1}^{n} |\mu_j - \lambda_j|^2\right)^{1/2} \leq \sqrt{2}\|E\|_F,$$

where $\Re(\mu_1) \leq \ldots \leq \Re(\mu_n)$.

Below we extend Kahan's bound (Theorem 3.4) to normal matrices whose eigenvalues lie on a line in the complex plane, and thereby improve Sun's bound (Theorem 3.2) for this class of matrices.

COROLLARY 3.5. *If $A = e^{\imath\theta}H + \gamma I$, where H is Hermitian, θ a real scalar and γ a complex scalar, then*

$$\left(\sum_{j=1}^{n} |\mu_j - \lambda_j|^2\right)^{1/2} \leq \sqrt{2}\|E\|_F,$$

where $\Re(\mu_1) \leq \ldots \leq \Re(\mu_n)$.

Proof. Denote by $\lambda_j(H)$ the eigenvalues of H. Then the eigenvalues of A are $\lambda_j = e^{\imath\theta}\lambda_j(H) + \gamma$, and the eigenvalues of $A + E$ are $\mu_k = e^{\imath\theta}\lambda_k(H + F) + \gamma$, where $F \equiv e^{-\imath\theta}E$. Hence $|\lambda_j - \mu_k| = |\lambda_j(H) - \lambda_k(H + F)|$. Now apply Theorem 3.4 to the eigenvalues of H and $H + F$. $\square$

Kahan's bound for Hermitian matrices also implies a two-norm bound.

COROLLARY 3.6 (Problem VI.8.8 in [2]). *If A is Hermitian then*

$$\max_{1\leq j\leq n} |\mu_j - \lambda_j| \leq \sqrt{2n}\|E\|_2,$$

where $\Re(\mu_1) \leq \ldots \leq \Re(\mu_n)$.

A tighter version of this bound replaces $\sqrt{2n}$ by $\gamma_n + 2$ where $\gamma_n/\ln n \to 2/\pi$ as $n \to \infty$ [2, Problem VI.8.8].

4. Existing Eigenvalue Bounds for Diagonalizable Matrices. We review existing eigenvalue perturbation bounds in the Frobenius norm and the two-norm.

Let A be a complex diagonalizable matrix of order n with eigenvalues λ_j and an eigenvector matrix X, and $A + E$ a complex diagonalizable matrix with eigenvalues μ_j and eigenvector matrix Y. Real eigenvalues are labeled $\lambda_1 \leq \ldots \leq \lambda_n$ and $\mu_1 \leq \ldots \leq \mu_n$. Let $\kappa(X) = \|X\|_2\,\|X^{-1}\|_2$ be the two-norm condition number with respect to inversion of the non-singular matrix X.

Frobenius Norm Bounds. The tightest and most general Frobenius norm bound for diagonalizable matrices is [3, Theorem 3.3]

$$\left(\sum_{j=1}^{n} |\mu_j - \lambda_{\sigma(j)}|^2\right)^{1/2} \leq \kappa(X)^{1/2}\kappa(Y)^{1/2}\,\|E\|_F, \tag{4.1}$$

where σ is a permutation. If, in addition, both A and $A + E$ have real eigenvalues then the eigenvalue ordering is known [3, Theorem 3.1]

$$\left(\sum_{j=1}^{n} |\mu_j - \lambda_j|^2\right)^{1/2} \leq \kappa(X)^{1/2}\kappa(Y)^{1/2}\,\|E\|_F. \tag{4.2}$$

Two-Norm Bounds. The tightest and most general two-norm bound for diagonalizable matrices is the Bauer-Fike theorem [1, Theorem III],

$$\max_{1\le j\le n}\ \min_{1\le k\le n} |\mu_j - \lambda_k| \le \kappa(X)\|E\|_2. \tag{4.3}$$

With only a slight sacrifice, a bound on the optimal matching distance can be obtained [2, Exercise VIII.3.2]

$$\max_{1\le j\le n} |\mu_j - \lambda_{\sigma(j)}| \le (2n-1)\,\kappa(X)\|E\|_2 \tag{4.4}$$

for some perturbation σ. If, in addition both A and $A+E$ have real eigenvalues then [3, Theorem 3.1]

$$\max_{1\le j\le n} |\mu_j - \lambda_j| \le \kappa(X)^{1/2}\kappa(Y)^{1/2}\,\|E\|_2. \tag{4.5}$$

However, the presence of the condition number $\kappa(Y)$ of the perturbed eigenvectors, like in (4.1), (4.2) and (4.5), can be of disadvantage. We illustrate this with two examples below.

Example 1. *The matrix*

$$A = \begin{pmatrix} 1 & \epsilon \\ 0 & 1-\epsilon \end{pmatrix}, \qquad 0 < \epsilon < 1$$

is diagonalizable with an eigenvector matrix

$$X = X^{-1} = \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix}.$$

The eigenvalues of A are well-conditioned because the eigenvector matrix X is well-conditioned with respect to inversion, i.e. $\kappa(X) \le 2.7$. If the perturbed matrix is

$$A+E = \begin{pmatrix} 1 & \epsilon \\ 0 & 1 \end{pmatrix}, \qquad \text{where} \quad E = \begin{pmatrix} 0 & 0 \\ 0 & \epsilon \end{pmatrix},$$

then the bounds (4.1), (4.2) and (4.5) don't apply because $A+E$ is defective.

If the perturbed matrix is

$$A+E = \begin{pmatrix} 1 & \epsilon \\ 0 & 1-\epsilon^2 \end{pmatrix}, \qquad \text{where} \quad E = \begin{pmatrix} 0 & 0 \\ 0 & \epsilon-\epsilon^2 \end{pmatrix},$$

then the bounds (4.1), (4.2) and (4.5) are unnecessarily pessimistic, because the perturbed eigenvector matrix

$$Y = \begin{pmatrix} 1 & 1/\epsilon \\ 0 & -1 \end{pmatrix}$$

has $\kappa(Y) \ge 1/\epsilon$, while the eigenvalue error is bounded by $\|E\|_F$. In this case, $\kappa(Y) \ge 1/\epsilon$ can be much larger than $\kappa(X) \le 2.7$.

In comparison, the Bauer-Fike theorem (4.3) implies the bound $\kappa(X)\|E\| \le 2.7\epsilon$ for both of the above perturbed matrices.

5. New Eigenvalue Bounds. We use the bounds in §3 to derive eigenvalue bounds for diagonalizable matrices that do not depend on the conditioning $\kappa(Y)$ of the perturbed eigenvectors. In particular, the perturbed matrix can be defective. Our bounds are tighter than those in §4 if Y is worse conditioned than X with respect to inversion.

THEOREM 5.1. *If A is diagonalizable then there is a permutation σ so that*

$$\left(\sum_{j=1}^{n} |\mu_j - \lambda_{\sigma(j)}|^2\right)^{1/2} \leq \sqrt{n}\kappa(X)\,\|E\|_F.$$

If, in addition, $\lambda_1 \leq \ldots \leq \lambda_n$ are real then

$$\left(\sum_{j=1}^{n} |\mu_j - \lambda_j|^2\right)^{1/2} \leq \sqrt{2}\kappa(X)\,\|E\|_F,$$

where $\Re(\mu_1) \leq \ldots \leq \Re(\mu_n)$.

Proof. Let $A = X\Lambda X^{-1}$ be an eigenvalue decomposition where Λ is diagonal with diagonal elements λ_i. Substitute this into $(A+E) - A = E$ to get

$$X^{-1}(A+E)X - \Lambda = X^{-1}EX,$$

where $X^{-1}(A+E)X$ has the same eigenvalues μ_i as $A+E$ and Λ is normal. Apply Theorem 3.2 to Λ and $X^{-1}(A+E)X$.

If λ_i are also real then Λ is Hermitian. Apply Theorem 3.4 to Λ and $X^{-1}(A+E)X$. □

When $A+E$ is diagonalizable the first bound in Theorem 5.1 improves (4.1) if $\kappa(Y) > n\kappa(X)$. When $A+E$ is diagonalizable with real eigenvalues the second bound in Theorem 5.1 improves (4.2) if $\kappa(Y) > 2\kappa(X)$. Therefore, Theorem 5.1 is tighter than existing bounds when the perturbed eigenvectors Y are worse-conditioned than the exact eigenvectors X with respect to inversion.

For the matrices in Example 1 Theorem 5.1 implies an eigenvalue bound of $3.71\|E\|_F$, which is within a factor 3.71 of the exact error.

The Frobenius norm bounds in Theorem 5.1 imply the following weaker two-norm bounds, of which the second one represents an extension of Corollary 3.6 to non-Hermitian matrices.

COROLLARY 5.2 (Two-Norm). *If A is diagonalizable then there is a permutation σ so that*

$$\max_{1\leq j\leq n} |\mu_j - \lambda_{\sigma(j)}| \leq n\kappa(X)\,\|E\|_2.$$

If, in addition, $\lambda_1 \leq \ldots \leq \lambda_n$ are real then

$$\max_{1\leq j\leq n} |\mu_j - \lambda_j| \leq \sqrt{2n}\kappa(X)\,\|E\|_2,$$

where $\Re(\mu_1) \leq \ldots \leq \Re(\mu_n)$.

Proof. The proof follows from Theorem 5.1 and the fact that $\|E\|_F \leq \sqrt{n}\,\|E\|_2$ for a square matrix E of order n. □

Although the Bauer-Fike theorem (4.3) is always tighter than Corollary 5.2, it does not bound the optimal matching distance. The first bound in Corollary 5.2

improves (4.4) by about a factor of 2. When $A + E$ is diagonalizable with real eigenvalues then the second bound in Corollary 5.2 is tighter than (4.5) if $\kappa(Y) > 2n\kappa(X)$.

The bound below improves Corollary 5.2 and is as good as the Bauer-Fike theorem (4.3). However the perturbation must be sufficiently small compared to the eigenvalue separation.

THEOREM 5.3 (Two-Norm). *If A is diagonalizable and*

$$\kappa(X)\, \|E\|_2 < \frac{1}{2} \min_{\lambda_k \neq \lambda_j} |\lambda_k - \lambda_j|$$

then there is a permutation σ so that

$$\max_{1 \leq j \leq n} |\mu_j - \lambda_{\sigma(j)}| \leq \kappa(X)\, \|E\|_2.$$

Proof. As in the proof of Theorem 5.1 show that $X^{-1}(A+E)X - \Lambda = X^{-1}EX$, and apply to the normal matrix Λ and to $A+E$ the bound [2, Theorem VI.5.1] which says that

$$\max_{1 \leq j \leq n} |\lambda_j(B+F) - \lambda_{\sigma(j)}(B)| \leq \|F\|_2,$$

for any normal matrix B, provided $\|F\|_2$ is smaller than half of the distance between any two distinct eigenvalues of B. □

To estimate the accuracy of computed eigenvalues, one may want to reverse the roles of A and $A + E$ because the exact eigenvectors are not known, but the conditioning of the computed ones can be estimated.

6. Inclusion Regions in the Frobenius Norm. We use the bounds in §3 to determine Gerschgorin-like inclusion regions for eigenvalues of complex square matrices.

Gerschgorin's disk theorem [7, Theorem 6.1.1.1] states that for a complex square matrix A with elements a_{ij} the union of the disks

$$\bigcup_i \{z :\ |z - a_{ii}|\} \leq \sum_{j \neq i} |a_{ij}|$$

contains all eigenvalues of A. If A is Hermitian then [2, Problem III.6.7] each interval

$$\left\{ z :\ |z - a_{ii}| \leq \left(\sum_{j \neq i} |a_{ij}|^2 \right)^{1/2} \right\} \tag{6.1}$$

contains at least one eigenvalue of A.

The result below can be viewed as an extension of (6.1) to non-Hermitian matrices.

THEOREM 6.1. *Let A be a complex square matrix with elements a_{ij}. There is a permutation σ so that*

$$\sum_{j=1}^{n} |a_{jj} - \lambda_{\sigma(j)}|^2 \leq n \sum_{k \neq j} |a_{kj}|^2.$$

If all diagonal elements a_{ii} are real, and A is symmetrically permuted so that $a_{11} \leq \ldots \leq a_{nn}$, then

$$\sum_{j=1}^{n} |a_{jj} - \lambda_j|^2 \leq 2 \sum_{k \neq j} |a_{kj}|^2,$$

where $\Re(\lambda_1) \leq \ldots \leq \Re(\lambda_n)$.

Proof. Partition $A = A_D + A_O$, where A_D is a diagonal matrix consisting of the diagonal elements and A_O consists of the off-diagonal elements of A. Since A_D is normal, Theorem 3.2 implies

$$\sum_{j=1}^{n} |a_{jj} - \lambda_{\sigma(j)}|^2 \leq n\, \|A_D - A\|_F^2 = n\, \|A_O\|_F^2.$$

If A_D is real, it is Hermitian, and Theorem 3.4 applies. Note that a symmetric permutation PAP^T, where P is a permutation matrix, does not change the eigenvalues of A. $\square$

For Hermitian matrices the bound in Theorem 6.1 is a factor of 2 larger than the one implied by (6.1).

7. Departure from Normality. We use the bounds in §3 to estimate the departure of a matrix from normality in terms of the diagonal and off-diagonal elements. Intuitively, one would expect that if the off-diagonal 'mass' of a matrix is small, so should its departure from normality. We confirm this for real matrices, by bounding the departure from normality in terms of the off-diagonal elements.

Let A be a complex square matrix of order n with Schur decomposition $A = Q(\Lambda + N)Q^*$, where Q is unitary, N strictly upper triangular and Λ diagonal with the diagonal elements λ_i being the eigenvalues of A. We measure the departure of A from normality by $\|N\|_F$ [5, (1.4)].

Partition $A = A_D + A_O$, where A_D is a diagonal matrix consisting of the diagonal elements and A_O consists of the off-diagonal elements of A.

Theorem 7.1.

$$\|N\|_F \leq \sqrt{\|A_O\|_F^2 + 2\sqrt{n}\, \|A_D\|_F \|A_O\|_F}.$$

Proof. From

$$\|A_D\|_F^2 + \|A_O\|_F^2 = \|A\|_F^2 = \|\Lambda\|_F^2 + \|N\|_F^2$$

follows $\|N\|_F^2 = \|A_O\|_F^2 + \|A_D\|_F^2 - \|\Lambda\|_F^2$. If $\|A_D\|_F \leq \|\Lambda\|_F$ then $\|N\|_F^2 \leq \|A_O\|_F^2$ and the desired bound holds. If $\|A_D\|_F > \|\Lambda\|_F$ then

$$\begin{aligned} \|A_D\|_F^2 - \|\Lambda\|_F^2 &= (\|A_D\|_F + \|\Lambda\|_F)\,(\|A_D\|_F - \|\Lambda\|_F) \\ &\leq 2\|A_D\|_F \, \min_P \|A_D - P^*\Lambda P\|_F, \end{aligned}$$

where the minimum ranges over all permutation matrices P. Theorem 6.1 implies

$$\min_P \|A_D - P^*\Lambda P\|_F \leq \sqrt{n}\|A_O\|_F.$$

$\square$

In the special case when all diagonal elements of A are zero, i.e. $A_D = 0$, Theorem 7.1 implies $\|N\|_F \leq \|A_O\|_F$. Theorem 7.1 is not tight for normal matrices. That's why we derive another bound which is tighter for a larger class of matrices.

THEOREM 7.2. *If A_D is real then*

$$\|N\|_F^2 \leq \|A_O\|_F^2 - \mathrm{trace}(A_O^2).$$

Consequently, $\|N\|_F \leq \sqrt{2}\|A_O\|_F$.

Proof. From $\|N\|_F^2 = \|A\|_F^2 - \|\Lambda\|_F^2$ and $\|A\|_F^2 = \|A_O\|_F^2 + \|A_D\|_F^2$ follows

$$\|N\|_F^2 = \|A_O\|_F^2 + \|A_D\|_F^2 - \|\Lambda\|_F^2.$$

With, see also [4, §2],

$$\|\Lambda\|_F^2 = \sum_{j=1}^n |\lambda_j|^2 \geq |\sum_{j=1}^n \lambda_j^2| = |\,\mathrm{trace}(A^2)|$$

and $\mathrm{trace}(A^2) = \mathrm{trace}(A_O^2) + \mathrm{trace}(A_D^2)$ we get

$$\|N\|_F^2 \leq \|A_O\|_F^2 + \|A_D\|_F^2 - \mathrm{trace}(A_O^2) - \mathrm{trace}(A_D^2) = \|A_O\|_F^2 - \mathrm{trace}(A_O^2).$$

The last equality follows because A_D is real, so $\|A_D\|^2 = \mathrm{trace}(A_D^2)$ for real A_D.

The second bound follows from $|\,\mathrm{trace}(A_O^2)| \leq \|A_O\|_F^2$. □

The second bound in Theorem 7.2 shows that the departure from normality is small if the off-diagonal 'mass' is small. The next corollary presents matrices for which the first bound in Theorem 7.2 holds with equality.

COROLLARY 7.3. *If A is a complex square matrix all of whose diagonal elements and eigenvalues are real, or if A is Hermitian, then*

$$\|N\|_F^2 = \|A_O\|_F^2 - \mathrm{trace}(A_O^2).$$

Proof. In the proof of Theorem 7.2, if all λ_j are real, then

$$\|\Lambda\|_F^2 = \sum_{j=1}^n |\lambda_j|^2 = \sum_{j=1}^n \lambda_j^2 = \mathrm{trace}(A^2).$$

□

8. Lower Bounds for Eigenvalue Perturbations.

We present a lower bound for eigenvalue perturbations of general, complex matrices and of Hermitian matrices.

THEOREM 8.1. *For any ordering of the eigenvalues λ_j of A and μ_j of $A+E$,*

$$|\,\mathrm{trace}(E)| \leq \sum_{j=1}^n |\mu_j - \lambda_j|.$$

Proof.

$$|\,\mathrm{trace}(E)| = |\,\mathrm{trace}(A+E) - \mathrm{trace}(A)| = \left|\sum_{j=1}^n (\mu_j - \lambda_j)\right| \leq \sum_{j=1}^n |\mu_j - \lambda_j|.$$

□

The bounds in Theorem 8.1 hold with equality if $E = \gamma I + QUQ^*$, where γ is a complex scalar, U is strictly upper triangular and Q is a Schur vector matrix of A.

As a consequence of Theorem 8.1, the Cauchy-Schwartz inequality implies

$$\frac{1}{\sqrt{n}}\,|\operatorname{trace}(E)| \leq \left(\sum_{j=1}^{n} |\mu_j - \lambda_j|^2\right)^{1/2}.$$

Below we bound the imaginary parts of eigenvalues of perturbed Hermitian matrices (this represents an absolute error since the eigenvalues of Hermitian matrices have zero imaginary part).

THEOREM 8.2. *If A is Hermitian then the imaginary parts $\Im(\mu_j)$ of the eigenvalues μ_j of $A+E$ satisfy*

$$\operatorname{trace}(\Im(E)) = \sum_{j=1}^{n} \Im(\mu_j),$$

and

$$\frac{1}{\sqrt{n}}\,|\operatorname{trace}(\Im(E))| \leq \left(\sum_{j=1}^{n} |\Im(\mu_j)|^2\right)^{1/2} \leq \|\Im(E)\|_F.$$

Proof. Let $A+E = Q(M+N)Q^*$ be a Schur decomposition of $A+E$ where Q is unitary, N strictly upper triangular, and M diagonal with the eigenvalues μ_i on the diagonal. Then $QNQ^* = A + E - QMQ^*$, and $\Im(N) = \Im(E - QMQ^*)$ because A is Hermitian. From N nilpotent follows $0 = \operatorname{trace}(\Im(N)) = \operatorname{trace}(E - QMQ^*)$, and $\operatorname{trace}(E) = \operatorname{trace}(\Im(M))$.

From Theorem 8.2 and the upper bound [8, §0], [12, (1.7)]

$$\left(\sum_{j=1}^{N} |\Im(\mu_j)|^2\right)^{1/2} \leq \|\Im(A+E)\|_F = \|\Im(E)\|_F$$

follows

$$|\operatorname{trace}(\Im(E))| \leq \sqrt{n}\left(\sum_{j=1}^{n} |\Im(\mu_j)|^2\right)^{1/2}.$$

□

REFERENCES

[1] F. Bauer and C. Fike, *Norms and exclusion theorems*, Numer. Math., 2 (1960), pp. 137–41.

[2] R. Bhatia, *Matrix Analysis*, Springer Verlag, New York, 1997.

[3] R. Bhatia, F. Kittaneh, and R. Li, *Some inequalities for commutators and an application to spectral variation. II*, Linear and Multilinear Algebra, 43 (1997), pp. 207–19.

[4] M. Gil', *Estimate for the norm of matrix-valued functions*, Linear Multilinear Algebra, 35 (1993), pp. 65–73.

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

[6] A. Hoffman and H. Wielandt, *The variation of the spectrum of a normal matrix*, Duke Math J., 20 (1953), pp. 37–9.

[7] R. Horn and C. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, 1985.

[8] W. Kahan, *Spectra of nearly Hermitian matrices*, Proc. Amer. Math. Soc., 48 (1975), pp. 11–7.

[9] S. Lee, *A practical upper bound for the departure from normality*, SIAM J. Matrix Anal. Appl., 16 (1995), pp. 462–68.

[10] ———, *Best available bounds for departure from normality*, SIAM J. Matrix Anal. Appl., 17 (1996), pp. 984–91.

[11] J. Sun, *On the variation of the spectrum of a normal matrix*, Linear Algebra Appl., 246 (1996), pp. 215–23.

[12] H. Wolkowicz and G. Styan, *Bounds for eigenvalues using traces*, Linear Algebra Appl., 29 (1980), pp. 471–506.