ORNL/TM-12853

Engineering Physics and Mathematics Division

BOUNDS FOR DEPARTURE FROM NORMALITY AND THE FROBENIUS NORM OF MATRIX EIGENVALUES

Steven L. Lee

Mathematical Sciences Section
Oak Ridge National Laboratory
P.O. Box 2008, Bldg. 6012
Oak Ridge, TN 37831-6367
na.slee@na-net.ornl.gov

Date Published: December 1994

Research was supported by the Applied Mathematical Sciences Research Program of the Office of Energy Research, U.S. Department of Energy.

Prepared by the
Oak Ridge National Laboratory
Oak Ridge, Tennessee 37831
managed by
Martin Marietta Energy Systems, Inc.
for the
U.S. DEPARTMENT OF ENERGY
under Contract No. DE-AC05-84OR21400

Contents

BOUNDS FOR DEPARTURE FROM NORMALITY AND THE FROBENIUS NORM OF MATRIX EIGENVALUES

Steven L. Lee

Abstract

New lower and upper bounds for the departure from normality and the Frobenius norm of the eigenvalues of a matrix are given. The sign properties of these bounds are also described. For example, th bound for matrix eigenvalues improves upon the one derived by de Vries and Wegmann in [Lin. Alg. Appl., 8 (1974), pp. 109–120]. Th upper bound for departure from normality is sharp for any matri eigenvalues are collinear in the complex plane. Moreover, the la is a practical estimate that costs (at most) $2m$ multiplications, the number of nonzeros in the matrix. In terms of applications, th can be used to bound from above the sensitivity of eigenvalues t perturbations or bound from below the condition number of the ei of a matrix.

1. Introduction

The departure from normality of a matrix, like the condition number of a matrix, is a real scalar that can be used to compute various matrix bounds. If A is an $n \times n$ matrix, its departure from normality (in the Frobenius norm) is defined to be [8]

$$\mathrm{dep}_F(A) := \left(\|A\|_F^2 - \|\Lambda\|_F^2 \right)^{1/2}, \tag{1}$$

where Λ is a diagonal matrix whose entries are the eigenvalues, λ_i, of A. This measure of matrix nonnormality can be used to bound the spectral norm of matrix functions [5, 2], the sensitivity of eigenvalues to matrix perturbations [8, 16], and the distance to the closest normal matrix [16, 10], for example. It is impractical to compute $\mathrm{dep}_F(A)$ if A is large and its eigenvalues are unknown. This difficulty motivates us to seek lower and upper bounds for $\mathrm{dep}_F(A)$ that are practical to compute or optimal in some sense.

In terms of eigenvalues, bounds for $\mathrm{dep}_F(A)$ can be used to obtain lower and upper bounds for

$$\|\Lambda\|_F^2, \quad \|\mathrm{Re}(\Lambda)\|_F^2, \quad \text{and} \quad \|\mathrm{Im}(\Lambda)\|_F^2, \tag{2}$$

where $\mathrm{Re}(\Lambda)$ and $\mathrm{Im}(\Lambda)$ are the real and imaginary parts of Λ. In particular, such results can be obtained by substituting lower and upper bounds for $\mathrm{dep}_F(A)$ into [13]

$$\|\Lambda\|_F^2 = \|A\|_F^2 - \mathrm{dep}_F^2(A), \tag{3}$$

$$\|\mathrm{Re}(\Lambda)\|_F^2 = \|M\|_F^2 - \frac{1}{2}\mathrm{dep}_F^2(A), \tag{4}$$

$$\|\mathrm{Im}(\Lambda)\|_F^2 = \|N\|_F^2 - \frac{1}{2}\mathrm{dep}_F^2(A) \tag{5}$$

where

$$M = \frac{1}{2}\left(A + A^H\right) \tag{6}$$

and

$$N = \frac{1}{2}\left(A - A^H\right) \tag{7}$$

are the Hermitian and skew-Hermitian part of A, respectively. Upper bounds for $\|\Lambda\|_F^2$ can be used to bound the spectral radius [15] and the spread of a matrix [1]. Bounds for $\|A\|_F^2$ can also be used to compute or estimate lower bounds for the condition number of the eigenbasis of A [11].

The outline of this paper is as follows. In §2, we give the notation, definitions, and observations that will be needed in later sections. In §3, we present various bounds for $\|N\|_F^2$ and $\mathrm{dep}_F^2(A)$, and show how they can be improved. In §4, we describe the significant properties of the newly improved bounds. In §5, we group the currently known a priori bounds for $\|N\|_F^2$ and $\mathrm{dep}_F^2(A)$ into two main categories, and then show that the new bounds are among the best available.

2. Preliminaries

Let $A = (a_{ij})$ be an $n \times n$ matrix with conjugate transpose $A^H = (\bar{a}_{ij})$ and Frobenius norm

$$\|A\|_F^2 := \sum_{i,j} |a_{ij}|^2. \qquad (8)$$

Also, recall that A is normal if and only if (iff), for example, [7]

(9a) A has a complete, orthogonal set of eigenvectors,

(9b) $\|A\|_F = \|\Lambda\|_F = \sum |\lambda_k|^2$, or

(9c) $A^H A - A A^H = 0$.

The set of normal matrices includes the Hermitian, skew-Hermitian, and unitary matrices and, in general, any matrix that is unitarily similar to a diagonal matrix. It is easily seen that $\mathrm{dep}_F(A)$ is invariant with respect to complex shifts and rotations. That is,

$$\mathrm{dep}_F(A) = \mathrm{dep}_F(e^{-i\theta}(A - \alpha I)) \qquad (10)$$

for any complex scalar α and $0 \le \theta < 2\pi$. For the Frobenius norm, we note that

$$\|e^{-i\theta}(A - \alpha I)\|_F = \|A - \alpha I\|_F \qquad (11)$$

and

$$\|(A - \alpha I)^H (A - \alpha I) - (A - \alpha I)(A - \alpha I)^H\|_F = \|A^H A - A A^H\|_F. \qquad (12)$$

The simplification in (12) also holds when $A - \alpha I$ is replaced with $e^{-i\theta}(A - \alpha I)$. It is also easy to show that the quadratic function $\|A - \alpha I\|_F^2$ is minimized for $\alpha = \frac{\mathrm{tr}(A)}{n}$, where $\mathrm{tr}(A)$ is the trace of A. If $\mathrm{tr}(A) = 0$, we shall say A is a *centered* matrix. Centered matrices such as

$$\tilde{A} = A - \frac{\mathrm{tr}(A)}{n} I \qquad (13)$$

and

$$\tilde{\Lambda} = \Lambda - \frac{\mathrm{tr}(A)}{n} I \qquad (14)$$

will be denoted with a tilde accent. Finally, we give a lemma that relates the norm of the shifted matrices $A - \alpha I$ and $\Lambda - \alpha I$ to the norm of the centered matrices $\tilde{A}$ and $\tilde{\Lambda}$, respectively.

Lemma 2.1 *For any $n \times n$ matrix A and complex scalar α,*

$$\|A - \alpha I\|_F^2 = \|\tilde{A}\|_F^2 + \frac{|\mathrm{tr}(A - \alpha I)|^2}{n} \qquad (15)$$

and

$$\|\Lambda - \alpha I\|_F^2 = \|\tilde{\Lambda}\|_F^2 + \frac{|\mathrm{tr}(A - \alpha I)|^2}{n}. \qquad (16)$$

Proof: First, we relate the norm of A to the norm of $\tilde{A}$. For $\sigma = \frac{\mathrm{tr}(A)}{n}$, we have

$$\begin{aligned}
\|A\|_F^2 - \|\tilde{A}\|_F^2 &= \|A\|_F^2 - \left\| A - \frac{\mathrm{tr}(A)}{n} I \right\|_F^2 && (17)\\
&= \sum (|a_{ii}|^2) - \sum (|a_{ii} - \sigma|^2) && (18)\\
&= \sum (a_{ii}^H a_{ii}) - \sum \left[(a_{ii} - \sigma)^H (a_{ii} - \sigma) \right] && (19)\\
&= \sum (a_{ii}^H a_{ii}) - \sum \left[a_{ii}^H a_{ii} - a_{ii}^H \sigma - \sigma^H a_{ii} + \sigma^H \sigma \right] && (20)\\
&= \sigma \sum (a_{ii}^H) + \sigma^H \sum (a_{ii}) - n \sigma^H \sigma && (21)\\
&= \frac{\mathrm{tr}(A)}{n} \mathrm{tr}^H(A) + \frac{\mathrm{tr}^H(A)}{n} \mathrm{tr}(A) - n \left| \frac{\mathrm{tr}(A)}{n} \right|^2 && (22)\\
&= \frac{2 |\mathrm{tr}(A)|^2}{n} - \frac{|\mathrm{tr}(A)|^2}{n} && (23)\\
&= \frac{|\mathrm{tr}(A)|^2}{n}. && (24)
\end{aligned}$$

If we now replace A with $A - \alpha I$ on the right-hand side of (17) and (24), we obtain

$$\|A - \alpha I\|_F^2 - \left\| A - \alpha I - \frac{\mathrm{tr}(A - \alpha I)}{n} I \right\|_F^2 = \frac{|\mathrm{tr}(A - \alpha I)|^2}{n}, \qquad (25)$$

and the second term can be simplified via

$$\begin{aligned}
\left\| A-\alpha I - \frac{\mathrm{tr}(A-\alpha I)}{n} I \right\|_F &= \left\| A-\alpha I - \frac{[\mathrm{tr}(A) - \mathrm{tr}(\alpha I)]}{n} I \right\|_F^2 && (26)\\
&= \left\| A-\alpha I - \frac{\mathrm{tr}(A)}{n} I + \frac{n\alpha}{n} I \right\|_F^2 && (27)\\
&= \left\| A - \frac{\mathrm{tr}(A)}{n} I \right\|_F^2 && (28)\\
&= \|\tilde{A}\|_F^2 && (29)
\end{aligned}$$

to obtain (15). The second equality (16) can also be proved in this manner. ■

3. Bounds for eigenvalues and departure from normality

We now present several bounds for $\|\Lambda\|_F^2$ and $\mathrm{dep}_F^2(A)$, along with their important properties. An upper bound for $\|\Lambda\|_F^2$ is given by Kress, de Vries, and Wegmann in [9]. Moreover, the authors exhibit nonnormal matrices for which the bound is sharp, and prove that the upper bound is the best possible in terms of $\|A\|_F$ and $\|A^H A - AA^H\|_F$.

Theorem 3. *[9, Thm. 1] For nonnormal A there holds*

$$\|\Lambda\|_F^2 \le \left(\|A\|_F^4 - \frac{1}{2}\|A^H A - AA^H\|_F^2 \right)^{1/2} \qquad (30)$$

with equality iff

$$A = \gamma(vw^H + rwv^H), \qquad (31)$$

where γ is a nonzero complex scalar, $0 \le r < 1$ is a real scalar, and where v, w are orthonormal vectors.

A practical lower bound for $\|\Lambda\|_F^2$,

$$|\mathrm{tr}(A^2)| \le \|\Lambda\|_F^2, \qquad (32)$$

comes from the triangle inequality applied to the eigenvalues of A^2

$$\begin{aligned}
|\mathrm{tr}(A^2)| &= |\mathrm{tr}(\Lambda^2)| = \left|\sum \lambda_i^2\right| = \left|\lambda_1^2 + \cdots + \lambda_n^2\right| && (33)\\
&\le \left|\lambda_1^2\right| + \cdots + \left|\lambda_n^2\right| = |\lambda_1|^2 + \cdots + |\lambda_n|^2 = \|\Lambda\|_F^2. && (34)
\end{aligned}$$

The lower bound is sharp iff 0 and the eigenvalues of A are collinear. Moreover, the bound is cheap to compute since only the diagonal of A^2 is needed. This diagonal can be computed with (at most) m multiplications, where m is the number of nonzeros in A.

The lower bound [17] and the upper bound [5] for $\mathrm{dep}_F^2(A)$,

$$\|A\|_F^2 - \left(\|A\|_F^4 - \frac{1}{2}\|A^HA-AA^H\|_F^2\right)^{1/2} \le \mathrm{dep}_F^2(A) \le \|A\|_F^2 - |\mathrm{tr}(A^2)|, \quad (35)$$

can be obtained by substituting (30) and (32) into (1). The upper bound in (35) is sharp iff 0 and the eigenvalues of A are collinear, and it can be computed with (at most) $2m$ multiplications. The lower bound is an $O(n^3)$ computation that is sharp if A is normal, or satisfies condition (31). This lower bound inherits the properties of the upper bound (30) via (1); thus, it is the best possible in terms of $\|A\|_F$ and $\|A^HA-AA^H\|_F$.

In §3.1, we will strengthen the lower bounds for $\mathrm{dep}_F^2(A)$ and $\|\Lambda\|_F^2$ in (35) and (34). In §3.2, these improved lower bounds will be substituted into (3) and (1), respectively, to obtain tighter upper bounds for $\|\Lambda\|_F^2$ and $\mathrm{dep}_F^2(A)$.

3.1. Improved lower bounds

The value of $\mathrm{dep}_F(A)$ is invariant with respect to the shift parameter α; see (10). We now show how this free parameter can be used to maximize the lower bound in (35). For normal matrices, the lower bound

$$\|A-\alpha I\|_F^2 - \left(\|A-\alpha I\|_F^4 - \frac{1}{2}\|A^HA-AA^H\|_F^2\right)^{1/2} \le \mathrm{dep}_F^2(A-\alpha I) = \mathrm{dep}_F^2(A) \quad (36)$$

is zero for any choice of α. For nonnormal matrices, however, there is a unique value of α that maximizes (36). In particular, by substituting (15) into (36), we seek to maximize the function

$$f(z(\alpha)) = \left(\beta^2 + z^2(\alpha)\right) - \left[\left(\beta^2 + z^2(\alpha)\right)^2 - \frac{1}{2}K^2\right]^{1/2} \quad (37)$$

where

$$z^2(\alpha) = \frac{|\mathrm{tr}(A-\alpha I)|^2}{n} \quad (38)$$

and

$$\beta^2 = \|\tilde{A}\|_F^2, \quad K^2 = \|A^H A - AA^H\|_F^2 > 0. \tag{39}$$

By solving

$$\frac{df}{dz} = 2z\left(1 - (\beta^2 + z^2)\left[(\beta^2 + z^2)^2 - \frac{1}{2}K^2\right]^{-1/2}\right) = 0, \tag{40}$$

we find that the unique solution $z = 0$ is a global maximum since

$$\frac{d^2 f}{dz^2}(0) = 2\left(1 - \frac{\beta^2}{(\beta^4 - \frac{1}{2}K^2)^{1/2}}\right) < 0. \tag{41}$$

By solving

$$z(\alpha) = \frac{|\mathrm{tr}(A - \alpha I)|}{n^{1/2}} = 0, \tag{42}$$

we find that the lower bound is maximized for $\alpha = \frac{\mathrm{tr}(A)}{n}$.

Lemma 3.2. *For any $n \times n$ matrix A,*

$$\mathrm{dep}_F^2(A) \geq \|A - \alpha I\|_F^2 - \left(\|A - \alpha I\|_F^4 - \frac{1}{2}\|A^H A - AA^H\|_F^2\right)^{1/2}, \tag{43}$$

where the lower bound is maximized for $\alpha = \frac{\mathrm{tr}(A)}{n}$.

Proof: The lemma follows via (37)-(42). ■

An improved lower bound for $\|\Lambda\|_F^2$ is less troublesome to obtain.

Lemma 3.3. *For any $n \times n$ matrix A,*

$$\|\Lambda\|_F^2 \geq |\mathrm{tr}(\tilde{A}^2)| + \frac{|\mathrm{tr}(A)|^2}{n}, \tag{44}$$

where $\tilde{A} = A - \frac{\mathrm{tr}(A)}{n} I$.

Proof: As in (33)-(34), we begin by applying the triangle inequality to the eigenvalues of $\tilde{A}^2$,

$$|\mathrm{tr}(\tilde{A}^2)| = |\mathrm{tr}(\tilde{\Lambda}^2)| \leq \|\tilde{\Lambda}^2\|_F = \|\tilde{\Lambda}\|_F^2. \tag{45}$$

The lemma is obtained by substituting this lower bound for $\|\tilde{\Lambda}\|_F^2$ into

$$\|\Lambda\|_F^2 = \|\tilde{\Lambda}\|_F^2 + \frac{|\mathrm{tr}(A)|^2}{n}. \qquad (46)$$

Note that (46) is the same as (16) with $\alpha=0$. ■

3.2. Improved upper bounds

For the $\|\Lambda\|_F^2$ upper bound, we have

$$\begin{aligned}\|\Lambda\|_F^2 &= \|A\|_F^2 - \mathrm{dep}_F^2(A) && (47)\\ &\le \|A\|_F^2 - \left(\|\tilde{A}\|_F^2 - \left(\|\tilde{A}\|_F^4 - \frac{1}{2}\|A^HA - AA^H\|_F^2\right)^{1/2}\right). && (48)\end{aligned}$$

Equation (15), with $\alpha=0$, shows that $\|A\|_F^2 - \|\tilde{A}\|_F^2$ simplifies to $\frac{|\mathrm{tr}(A)|^2}{n}$.

Lemma 3.4. *For any $n\times n$ matrix A,*

$$\|\Lambda\|_F^2 \le \left(\|\tilde{A}\|_F^4 - \frac{1}{2}\|A^HA - AA^H\|_F^2\right)^{1/2} + \frac{|\mathrm{tr}(A)|^2}{n}, \qquad (49)$$

where $\tilde{A} = A - \frac{\mathrm{tr}(A)}{n}I$.

For the $\mathrm{dep}_F^2(A)$ upper bound, we can substitute $\|\tilde{\Lambda}\|_F^2 \ge |\mathrm{tr}(\tilde{A}^2)|$ into

$$\mathrm{dep}_F^2(A) = \mathrm{dep}_F^2(\tilde{A}) = \|\tilde{A}\|_F^2 - \|\tilde{\Lambda}\|_F^2 \qquad (50)$$

to obtain the following lemma.

Lemma 3.5. *For any $n\times n$ matrix A,*

$$\mathrm{dep}_F^2(A) \le \|\tilde{A}\|_F^2 - |\mathrm{tr}(\tilde{A}^2)|, \qquad (51)$$

where $\tilde{A} = A - \frac{\mathrm{tr}(A)}{n}I$.

4. Main results

In this section, we establish the significant properties of the four bounds given in §3.1 and §3.2. To begin, recall that the lower bound for $\mathrm{dep}_F^2(A)$, in turn, the upper bound for $\|\Lambda\|_F^2$ were optimized via the complex shift $\alpha = \frac{\mathrm{tr}(A)}{n}$. Moreover, the latter derivation (47)–(48) shows that if the $\mathrm{dep}_F^2(A)$ lower bound is sharp then so is the $\|\Lambda\|_F^2$ upper bound. The $\mathrm{dep}_F^2(A)$ lower bound in (35) is sharp for any nonnormal matrix that satisfies condition (31). The improved $\mathrm{dep}_F^2(A)$ lower bound (43) is unaffected by complex shifts; thus, it is sharp for

$$A = \gamma(vw^H + rwv^H) - \sigma I \tag{52}$$

for any choice of the scalar σ. Note that we have

$$\alpha = \frac{\mathrm{tr}(\gamma(vw^H + rwv^H) - \sigma I)}{n} = -\sigma, \tag{53}$$

and that the shift α in (43) cancels the arbitrary shift σ. The improved bound is also unaffected by rotations. We summarize the above results as follows.

Theorem 4. *For any $n \times n$ matrix A,*

$$\mathrm{dep}_F^2(A) \;\geq\; \|\tilde{A}\|_F^2 - \left(\|\tilde{A}\|_F^4 - \frac{1}{2}\|A^HA - AA^H\|_F^2\right)^{1/2} \tag{54}$$

and

$$\|\Lambda\|_F^2 \;\leq\; \left(\|\tilde{A}\|_F^4 - \frac{1}{2}\|A^HA - AA^H\|_F^2\right)^{1/2} + \frac{|\mathrm{tr}(A)|^2}{n}, \tag{55}$$

where $\tilde{A} = A - \frac{\mathrm{tr}(A)}{n}I$. The bounds are sharp iff

$$A = e^{-i\theta}(\gamma(vw^H + rwv^H) - \sigma I), \tag{56}$$

where γ, r and σ are complex scalars, $0 \leq \theta < 2\pi$, and where v, w are orthonormal vectors.

We will now prove that the other two bounds (44) and (51) are sharp iff the eigenvalues of A are collinear in the complex plane. Before doing so, we must establish a natural measure of the noncollinearity of matrix eigenvalues. One approach is to define "departure from collinearity" as

$$\mathrm{depcol}(A) := \sum |d_k|^2, \tag{57}$$

where $|d_k|$ is the perpendicular distance from λ_k to the total least squares (TLS) fit of the eigenvalues of A. Recall that a TLS fit minimizes the sum of the squares of the perpendicular distances from the points to the fitted line, $\sum |d_k|^2$, and that is the TLS error [16]. Given the definition (57), we find depcol(A) to be a sensible metric for quantifying departure from collinearity, especially since depcol(A) = 0 iff A has collinear eigenvalues.

A useful result concerning TLS error and departure from collinearity follows from [12, Thm 2.2].

Theorem 4.2 *Given the complex numbers z_k, $k = 1, \cdots, n$, let $\bar{z} = \frac{1}{n}\sum z_k$ so that*

$$\tilde{z}_k = z_k - \bar{z}. \tag{58}$$

The error for the total least squares fit is

$$\sum |d_k|^2 = \frac{1}{2}\left(\sum |\tilde{z}_k|^2 - \left|\sum \tilde{z}_k^2\right|\right), \tag{59}$$

where $|d_k|$ is the perpendicular distance from z_k to the fit.

In the context of matrix eigenvalues, (59) yields

$$\text{depcol}(A) = \frac{1}{2}\left(\|\tilde{\Lambda}\|_F^2 - |\text{tr}(\tilde{A}^2)|\right). \tag{60}$$

If we arrange (60) as

$$\|\tilde{\Lambda}\|_F^2 = |\text{tr}(\tilde{A}^2)| + 2\,\text{depcol}(A), \tag{61}$$

and substitute into (46) and (50), we obtain

$$\|\Lambda\|_F^2 = |\text{tr}(\tilde{A}^2)| + 2\,\text{depcol}(A) + \frac{|\text{tr}(A)|^2}{n} \tag{62}$$

and

$$\text{dep}_F^2(A) = \|\tilde{A}\|_F^2 - \left(|\text{tr}(\tilde{A}^2)| + 2\,\text{depcol}(A)\right). \tag{63}$$

Note that the bounds in Lemmas 3.3 and 3.5 are special cases of (62) and (63).

Theorem 4.3 *For any $n \times n$ matrix A,*

$$\|\Lambda\|_F^2 \geq |\text{tr}(\tilde{A}^2)| + \frac{|\text{tr}(A)|^2}{n} \tag{64}$$

and

$$\mathrm{dep}_F^2(A) \le \|\tilde{A}\|_F^2 - |\mathrm{tr}(\tilde{A}^2)|, \qquad (65)$$

where $\tilde{A} = A - \frac{\mathrm{tr}(A)}{n} I$. *The bounds are sharp iff the eigenvalues of* A *are collinear in the complex plane. Moreover,*

$$\|\Lambda\|_F^2 \approx |\mathrm{tr}(\tilde{A}^2)| + \frac{|\mathrm{tr}(A)|^2}{n} \qquad (66)$$

and

$$\mathrm{dep}_F^2(A) \approx \|\tilde{A}\|_F^2 - |\mathrm{tr}(\tilde{A}^2)| \qquad (67)$$

iff the eigenvalues of A *are relatively close to being collinear.*

Proof: The bounds (64)–(65) are obtained from (62)–(63) by dropping the term $2\,\mathrm{dep}_{col}(A)$. These bounds are sharp iff $\mathrm{dep}_{col}(A) = 0$; that is, iff the eigenvalues of A are collinear. Finally, the bounds are good estimates when the neglected term $\mathrm{dep}_{col}(A)$ is relatively small. ∎

5. Discussion and summary

To the best of our knowledge, a priori bounds for $\|\Lambda\|_F^2$ and $\mathrm{dep}_F^2(A)$ fall into one of two distinct categories. The bounds in the first category are based on computing the Frobenius norm of the commutator $A^H A - A A^H$ [3, 4, 8, 14, 17]. The bounds in the second category are based on inequalities that are sharp iff the eigenvalues of A have a certain alignment in the complex plane [13, 15]. For each of these categories, we now give the best available bounds known to us at this time.

Bounds based on $\|A^H A - A A^H\|_F$

$$\|\Lambda\|_F^2 \ge \|A\|_F^2 - \left(\frac{n^3 - n}{12}\right)^{1/2} \left(\|A^H A - A A^H\|_F\right) \qquad (68)$$

$$\|\Lambda\|_F^2 \le \left(\|\tilde{A}\|_F^4 - \frac{1}{2}\|A^H A - A A^H\|_F^2\right)^{1/2} + \frac{|\mathrm{tr}(A)|^2}{n} \qquad (69)$$

$$\mathrm{dep}_F^2(A) \ge \|\tilde{A}\|_F^2 - \left(\|\tilde{A}\|_F^4 - \frac{1}{2}\|A^H A - A A^H\|_F^2\right)^{1/2} \qquad (70)$$

$$\mathrm{dep}_F^2(A) \;\le\; \left(\frac{n^3-n}{12}\right)^{1/2}\left(\|A^HA-AA^H\|_F\right) \qquad (71)$$

Remarks. The lower bound (68) is the counterpart to the upper bound (71) of Henrici [8, Thm 1]. The bounds (69)–(70) are given in Theorem 4.1. Note that Sun's lower bound (35) is the best possible in terms of $\|A^HA-AA^H\|_F$ and $\|A\|_F$; thus it is stronger than the bounds in [4]. The bound (70) improves upon Sun's lower bound, and it is also stronger than the one in [14].

Bounds based on eigenvalue alignment

$$\|\Lambda\|_F^2 \;=\; |\mathrm{tr}(\tilde{A}^2)| + 2\,\mathrm{depco}(A) + \frac{|\mathrm{tr}(A)|^2}{n} \;\ge\; |\mathrm{tr}(\tilde{A}^2)| + \frac{|\mathrm{tr}(A)|^2}{n} \qquad (72)$$

$$\mathrm{dep}_F^2(A) \;=\; \|\tilde{A}\|_F^2 - \left(|\mathrm{tr}(\tilde{A}^2)| + 2\,\mathrm{depco}(A)\right) \;\le\; \|\tilde{A}\|_F^2 - |\mathrm{tr}(\tilde{A}^2)| \qquad (73)$$

Remarks. The new bounds (72)–(73) are sharp iff the eigenvalues of A are collinear. In contrast, note that for $\alpha = \frac{\mathrm{tr}(A)}{n}$, we have [13, Thm 3.2]

$$\mathrm{dep}_F^2(A) \;\le\; 2\,\min\left\{\|M-\mathrm{Re}(\alpha)I\|_F^2,\; \|N-i\,\mathrm{Im}(\alpha)I\|_F^2\right\} \qquad (74)$$

and its (unsimplified) counterpart

$$\|\Lambda\|_F^2 \;\ge\; \|A\|_F^2 - 2\,\min\left\{\|M-\mathrm{Re}(\alpha)I\|_F^2,\; \|N-i\,\mathrm{Im}(\alpha)I\|_F^2\right\}. \qquad (75)$$

Unfortunately, the bounds (74)–(75) are sharp only when the eigenvalues are horizontally or vertically aligned in the complex plane. Furthermore, the bounds in (72)–(73) are half as expensive to compute as those in (74)–(75). Despite these shortcomings, the latter bounds are useful and have some noteworthy properties. In particular, the bounds in (72)–(73) and those in (74)–(75) yield the same values if A is a real matrix. We also remark that (74) explicitly bounds matrix nonnormality in terms of the nonsymmetry of A.

Besides their practicality, the estimates (72)–(73) are also appealing because they sometimes enable us to precisely compute $\|\Lambda\|_F^2$ and $\mathrm{dep}_F^2(A)$ for matrices with extremely sensitive eigenvalues. For example, consider the $n\times n$ matrix

$$\widehat{W}_n = U^H W_n U \qquad (76)$$

where $\widehat{W}_n$ is dense and unitarily similar to the Wilkinson matrix [18, pg. 90],

$$W_n = \begin{bmatrix} n & n & & & \\ & n-1 & n & & \\ & & \ddots & \ddots & \\ & & & 2 & n \\ & & & & 1 \end{bmatrix}, \qquad (77)$$

where $n=20$. The eigenvalues of $\widehat{W}_n$ are real, and the interior eigenvalues are notoriously difficult to compute for $n \gg 20$. Thus, we cannot directly compute $\|\Lambda\|_F^2$ for, say, $\widehat{W}_{50}$ due to these eigenvalue sensitivities. However, we can accurately obtain $\|\Lambda\|_F^2$ and $\mathrm{dep}_F^2(A)$ for $\widehat{W}_{50}$ via (72)–(73) since the sharpness of these formulas (modulo rounding errors) only depends upon eigenvalue collinearity — not eigenvalue sensitivity.

To summarize, we have developed several new and improved bounds for $\mathrm{dep}_F^2(A)$ and $\|\Lambda\|_F^2$, and described their significant properties. We have also grouped these and the other known a priori bounds for $\mathrm{dep}_F^2(A)$ and $\|\Lambda\|_F^2$ into two categories. Within each category, we have given the best available bounds. The bounds based on $\|A^HA - AA^H\|_F$ have an important property: they reduce to zero if A is normal. Unfortunately, such bounds are often weak, and impractical to compute if A is large. On the other hand, the bounds based on eigenvalue alignment are often good estimates (e.g., Table 1 [31]), and they are practical to compute if A is large and sparse. A minor drawback is that these bounds only reduce to zero for normal matrices with collinear eigenvalues (e.g., Hermitian and skew-Hermitian matrices). Theorem 4.1, Theorem 4.3, and Theorem in [8] describe the nonnormal matrices for which the bounds in (68)–(73) are sharp. The significance of our results are described in §

6. References

[1] N. A. Derzko and A. M. Pfeffer. Bounds for the spectral radius of a matrix. *Math. Comp.*, 19:62–67, 1965.

[2] J. Descloux. Bounds for the spectral norm of functions of matrices. *Numer. Math.*, pages 185–190, 1963.

[3] P. J. Eberlein. On measures of non-normality for matrices. *Amer. Math. Mon.*, 72:995–996, 1965.

[4] L. Elsner and M. H. C. Paardekooper. On measures of nonnormality of matrices. *Linear Algebra Appl.*, 92:107–124, 1987.

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

[6] G. H. Golub and C. F. Van Loan. An analysis of the total least squares problem. *SIAM J. Numer. Anal.*, 17:883–893, 1980.

[7] R. Grone, C. R. Johnson, E. M. Sá, and H. Wolkowicz. Normal matrices. *Linear Algebra Appl.*, 87:213–225, 1987.

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

[9] R. Kress, H. L. de Vries, and R. Wegmann. On nonnormal matrices. *Linear Algebra Appl.*, 8:109–120, 1974.

[10] L. László. An attainable lower bound for the best normal approximation. *SIAM J. Matrix Anal. Appl.*, 15:1035–1043, 1994.

[11] S. L. Lee. On the eigenvalue sensitivity of nonnormal matrices. Technical report, Oak Ridge National Laboratory. In preparation.

[12] S. L. Lee. A note on the total least squares problems for coplanar points. Technical report TM-12852, Oak Ridge National Laboratory, September 1994.

[13] S. L. Lee. A practical upper bound for departure from normality. *SIAM J. Matrix Anal. Appl.*, 16(2), April 1995. To appear.

[14] G. Loizou. Nonnormality and Jordan condition numbers of matrices. *J. Assoc. Comput. Mach.*, 16(4):580–584, October 1969.

[15] L. Mirsky. The spread of a matrix. *Mathematika*, 3:127–130, 1956.

[16] A. Ruhe. Closest normal matrix finally found! *BIT*, 27:585–598, 1987.

[17] J.-G. Sun. *Matrix Perturbation Analysis*. Chinese Academic Press, 1987. In Chinese.

[18] J. H. Wilkinson. *The Algebraic Eigenvalue Problem*. Claredon Press, Oxford, England, 1965.