

20.4 The Bauer–Fike Theorem

Section 3 bounded the movement of the eigenvalues of an arbitrary matrix by a constant times $\|\mathbf{E}\|_2^{1/n}$, and the Jordan block showed that nothing better holds in general. The culprit was the Jordan block itself. This section removes it by assuming that $\mathbf{A}$ is diagonalizable, and the exponent $1/n$ becomes 1: every eigenvalue of $\mathbf{A} + \mathbf{E}$ lies within $\kappa_2(\mathbf{X})\|\mathbf{E}\|_2$ of some eigenvalue of $\mathbf{A}$, where $\mathbf{X}$ is a matrix of eigenvectors.

This is the theorem several earlier chapters pointed to. Chapter 12 §11 described it as the statement “that for a diagonalizable $\mathbf{A}$ every eigenvalue of a perturbed $\mathbf{A} + \mathbf{E}$ lies within $\kappa\|\mathbf{E}\|$ of some eigenvalue of $\mathbf{A}$, where κ measures how far from unitary the diagonalizing matrix is, and that $\kappa = 1$ exactly in the normal case”, and said that it “needs matrix norms and a little analysis, and Chapter 20 develops both and proves it”. Chapter 16 §07 and Chapter 17 §03 promised a bound “for a diagonalizable matrix” carrying “an extra factor that measures how far the eigenvector basis of $\mathbf{A}$ is from orthonormal”. Chapter 17 §11 left the bounds for non-Hermitian matrices, “beginning with Bauer–Fike”, to this chapter, and Chapter 16 §08 recorded that “the bounds that do exist, Weyl’s and Bauer–Fike’s, are Chapters 17 and 20”. We prove the theorem, identify the factor, prove the “exactly”, and then return to the defective case. There the m -th root comes back, where m is the size of the largest Jordan block.

Throughout, $\mathbf{A} \in M_n(\mathbb{C})$, the perturbation is $\mathbf{E} \in M_n(\mathbb{C})$, and $\kappa_p(\mathbf{X}) = \|\mathbf{X}\|_p \|\mathbf{X}^{-1}\|_p$ is the condition number of Definition 16.8.1. No fact of analysis is imported in this section beyond those already used by the results it cites.

20.4.1 Factoring out the inverse

The whole section runs on one estimate. If μ is an eigenvalue of $\mathbf{A} + \mathbf{E}$ but not of $\mathbf{A}$, then $\mathbf{A} - \mu\mathbf{I}$ is invertible while $\mathbf{A} + \mathbf{E} - \mu\mathbf{I}$ is not. A small $\mathbf{E}$ cannot destroy invertibility (Corollary 16.5.2), so $\mathbf{E}$ must be large compared with the inverse of $\mathbf{A} - \mu\mathbf{I}$. The matrix $(\mathbf{A} - \mu\mathbf{I})^{-1}$, a function of μ , is called the **resolvent** of $\mathbf{A}$, up to a sign convention. Chapter 16 asked for this estimate in Exercise (C1: How close to singular a nearby eigenvalue forces you to be). Since the theorems below rest on it, we prove it here in full.

Lemma 20.4.1. A Perturbed Eigenvalue Forces a Large Resolvent

Let $\mathbf{A}, \mathbf{E} \in M_n(\mathbb{C})$, let $\|\cdot\|$ be an operator norm on $M_n(\mathbb{C})$ (Definition 16.3.2), and let μ be an eigenvalue of $\mathbf{A} + \mathbf{E}$ that is **not** an eigenvalue of $\mathbf{A}$. Then

$$1 \leq \|(\mathbf{A} - \mu\mathbf{I})^{-1}\mathbf{E}\| \leq \|(\mathbf{A} - \mu\mathbf{I})^{-1}\| \|\mathbf{E}\|.$$

Proof. Since $\mu \notin \text{spec}(\mathbf{A})$, the matrix $\mathbf{A} - \mu\mathbf{I}$ is invertible (Theorem 9.2.1), and

$$\mathbf{A} + \mathbf{E} - \mu\mathbf{I} = (\mathbf{A} - \mu\mathbf{I})(\mathbf{I} + (\mathbf{A} - \mu\mathbf{I})^{-1}\mathbf{E}).$$

The left side is singular, because μ is an eigenvalue of $\mathbf{A} + \mathbf{E}$. If the second factor on the right were invertible, the product of two invertible matrices would be invertible. So $\mathbf{I} - \mathbf{F}$ is singular, where $\mathbf{F} = -(\mathbf{A} - \mu\mathbf{I})^{-1}\mathbf{E}$. By Theorem 16.5.1 (a), $\|\mathbf{F}\| < 1$ would make $\mathbf{I} - \mathbf{F}$ invertible, hence $\|\mathbf{F}\| \geq 1$. Since $\|\mathbf{F}\| = \|(\mathbf{A} - \mu\mathbf{I})^{-1}\mathbf{E}\|$, this is the first inequality, and the second is submultiplicativity (Theorem 16.3.2 (d)).

■

The lemma holds for every matrix and needs no eigenvectors. To turn it into a bound on $|\lambda_i - \mu|$, we need to compute $\|(\mathbf{A} - \mu\mathbf{I})^{-1}\|$ in terms of the distances from μ to the eigenvalues. For a diagonal matrix that is immediate, and diagonalizable matrices are diagonal in the right coordinates.

20.4.2 The theorem

We need one fact about diagonal matrices. Let $\mathbf{D} = \text{diag}(d_1, \dots, d_n)$ and $p \in \{1, 2, \infty\}$. For every $\mathbf{x}$, each coordinate of $\mathbf{D}\mathbf{x}$ is $d_i x_i$, with $|d_i x_i| \leq (\max_j |d_j|)|x_i|$, so $\|\mathbf{D}\mathbf{x}\|_p \leq (\max_j |d_j|)\|\mathbf{x}\|_p$. Equality holds at $\mathbf{x} = \mathbf{e}_k$, where k is an index with $|d_k| = \max_j |d_j|$. Hence

$$\|\text{diag}(d_1, \dots, d_n)\|_p = \max_j |d_j| \quad (p = 1, 2, \infty). \quad (20.4.1)$$

Theorem 20.4.2. Bauer–Fike Theorem

Let $\mathbf{A} \in M_n(\mathbb{C})$ be diagonalizable. Write $\mathbf{A} = \mathbf{X}\mathbf{\Lambda}\mathbf{X}^{-1}$, where $\mathbf{X}$ is invertible and $\mathbf{\Lambda}$ is the diagonal matrix with diagonal entries $\lambda_1, \dots, \lambda_n$. Let $\mathbf{E} \in M_n(\mathbb{C})$ be arbitrary, and let $p \in \{1, 2, \infty\}$. Then every eigenvalue μ of $\mathbf{A} + \mathbf{E}$ satisfies

$$\min_{1 \leq i \leq n} |\lambda_i - \mu| \leq \kappa_p(\mathbf{X}) \|\mathbf{E}\|_p.$$

Idea. Change coordinates so that $\mathbf{A}$ becomes $\mathbf{\Lambda}$. Then $\mathbf{A} + \mathbf{E}$ becomes $\mathbf{\Lambda} + \mathbf{X}^{-1}\mathbf{E}\mathbf{X}$, with the same eigenvalues, and the perturbation has grown by at most the factor $\kappa_p(\mathbf{X})$. For the diagonal matrix $\mathbf{\Lambda}$, the resolvent $(\mathbf{\Lambda} - \mu\mathbf{I})^{-1}$ has norm exactly $1/\min_i |\lambda_i - \mu|$, and [Lemma 20.4.1](#) finishes the proof.

Proof. If $\mu = \lambda_i$ for some i , the left side is 0 and there is nothing to prove. So suppose $\mu \neq \lambda_i$ for every i , and put $\delta = \min_i |\lambda_i - \mu| > 0$.

Since $\mathbf{X}^{-1}(\mathbf{A} + \mathbf{E})\mathbf{X} = \mathbf{\Lambda} + \mathbf{F}$ with $\mathbf{F} = \mathbf{X}^{-1}\mathbf{E}\mathbf{X}$, the matrices $\mathbf{A} + \mathbf{E}$ and $\mathbf{\Lambda} + \mathbf{F}$ are similar, so μ is an eigenvalue of $\mathbf{\Lambda} + \mathbf{F}$ (Theorem 7.10.3). The eigenvalues of $\mathbf{\Lambda}$ are the λ_i , so $\mu \notin \text{spec}(\mathbf{\Lambda})$. By [Lemma 20.4.1](#) with the operator norm $\|\cdot\|_p$,

$$1 \leq \|(\mathbf{\Lambda} - \mu\mathbf{I})^{-1}\|_p \|\mathbf{F}\|_p.$$

Now $(\mathbf{\Lambda} - \mu\mathbf{I})^{-1} = \text{diag}(1/(\lambda_1 - \mu), \dots, 1/(\lambda_n - \mu))$, whose norm is $\max_i 1/|\lambda_i - \mu| = 1/\delta$ by [Eq. 20.4.1](#). By submultiplicativity (Theorem 16.3.2 (d)), $\|\mathbf{F}\|_p \leq \|\mathbf{X}^{-1}\|_p \|\mathbf{E}\|_p \|\mathbf{X}\|_p = \kappa_p(\mathbf{X}) \|\mathbf{E}\|_p$. Therefore $1 \leq \kappa_p(\mathbf{X}) \|\mathbf{E}\|_p / \delta$, that is, $\delta \leq \kappa_p(\mathbf{X}) \|\mathbf{E}\|_p$. This proves the theorem. ■

The proof used only two properties of $\|\cdot\|_p$: it is an operator norm, and it gives a diagonal matrix the norm $\max_j |d_j|$. So the theorem holds for every operator norm with that second property. The spectral norm is the one used in practice, and $\kappa_2(\mathbf{X})$ is the headline constant.

Read the theorem as a picture. Draw the closed disc of radius $\kappa_p(\mathbf{X}) \|\mathbf{E}\|_p$ about each λ_i ; every eigenvalue of $\mathbf{A} + \mathbf{E}$ lies in the union. The bound is **one-sided**, like Elsner's (Theorem 20.3.3).

It does not say that each disc holds one perturbed eigenvalue. When the discs are disjoint, Lemma 20.3.4 adds that, and Exercise C1 does it.

Quick check. Let $\mathbf{A} = \text{diag}(1, 2)$. Which matrices $\mathbf{X}$ satisfy $\mathbf{X}^{-1}\mathbf{A}\mathbf{X}$ diagonal, and what does Theorem 20.4.2 give for each of them?

Solution. The eigenvalues 1, 2 are distinct and each eigenspace is a line: $E_1(\mathbf{A}) = \text{span}(\mathbf{e}_1)$ and $E_2(\mathbf{A}) = \text{span}(\mathbf{e}_2)$. The columns of $\mathbf{X}$ must be eigenvectors (Theorem 9.4.1), so $\mathbf{X} = \text{diag}(a, b)$ or $\mathbf{X} = \begin{pmatrix} 0 & a \\ b & 0 \end{pmatrix}$ with $a, b \neq 0$. In both cases the singular values of $\mathbf{X}$ are $|a|$ and $|b|$, so $\kappa_2(\mathbf{X}) = \max(|a|, |b|) / \min(|a|, |b|)$ (Proposition 16.8.2 (d)). The theorem gives $\min_i |\lambda_i - \mu| \leq \kappa_2(\mathbf{X}) \|\mathbf{E}\|_2$. This is $\|\mathbf{E}\|_2$ when $|a| = |b|$, and arbitrarily weak as $|a|/|b|$ grows. The theorem is true for every choice, and one uses the best.

20.4.3 Normal matrices

For a normal matrix the eigenvectors can be chosen orthonormal, and then the factor is 1. The second half of the next result proves the converse, in the form Chapter 12 §11 stated it: $\kappa = 1$ **exactly** in the normal case.

Corollary 20.4.3. Bauer–Fike for Normal Matrices

- (a) Let $\mathbf{A} \in M_n(\mathbb{C})$ be normal with eigenvalues $\lambda_1, \dots, \lambda_n$, and let $\mathbf{E} \in M_n(\mathbb{C})$. Then every eigenvalue μ of $\mathbf{A} + \mathbf{E}$ satisfies $\min_i |\lambda_i - \mu| \leq \|\mathbf{E}\|_2$.
- (b) Let $\mathbf{A} \in M_n(\mathbb{C})$ be diagonalizable. Then

$$\inf \{ \kappa_2(\mathbf{X}) : \mathbf{X}^{-1}\mathbf{A}\mathbf{X} \text{ is diagonal} \} = 1,$$

the infimum over invertible $\mathbf{X}$, if and only if $\mathbf{A}$ is normal.

Idea. Part (a) is the theorem with a unitary $\mathbf{X}$. For (b), write $\mathbf{X} = \mathbf{W}\Sigma\mathbf{V}^*$ by the singular value decomposition. Then $\mathbf{A}$ is $\mathbf{W}\Sigma\mathbf{M}\Sigma^{-1}\mathbf{W}^*$ with $\mathbf{M}$ normal. If all the singular values of $\mathbf{X}$ are nearly equal, $\Sigma\mathbf{M}\Sigma^{-1}$ is nearly $\mathbf{M}$, so $\mathbf{A}$ is nearly the normal matrix $\mathbf{W}\mathbf{M}\mathbf{W}^*$, and its commutator $\mathbf{A}^*\mathbf{A} - \mathbf{A}\mathbf{A}^*$ is nearly $\mathbf{O}$. A commutator that is arbitrarily small is zero.

Proof.

- (a) By Corollary 12.4.3, $\mathbf{A} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^*$ with $\mathbf{U}$ unitary and $\mathbf{\Lambda} = \text{diag}(\lambda_1, \dots, \lambda_n)$, so $\mathbf{U}^{-1} = \mathbf{U}^*$. A unitary matrix preserves $\|\cdot\|_2$ (Theorem 11.7.3), so $\|\mathbf{U}\|_2 = \|\mathbf{U}^*\|_2 = 1$ and $\kappa_2(\mathbf{U}) = 1$. Now apply Theorem 20.4.2 with $\mathbf{X} = \mathbf{U}$ and $p = 2$.
- (b) ($\Leftarrow$) If $\mathbf{A}$ is normal, the $\mathbf{U}$ of (a) diagonalizes $\mathbf{A}$ with $\kappa_2(\mathbf{U}) = 1$, and every $\kappa_2(\mathbf{X}) \geq 1$ by Proposition 16.8.2 (a). So the infimum is 1.

($\Rightarrow$) First we prove that every invertible $\mathbf{X}$ with $\mathbf{X}^{-1}\mathbf{A}\mathbf{X} = \mathbf{\Lambda}$ diagonal satisfies

$$\begin{aligned} \|\mathbf{A}^* \mathbf{A} - \mathbf{A} \mathbf{A}^*\|_F &\leq 2(\kappa_2(\mathbf{X}) - 1) \|\mathbf{\Lambda}\|_F \\ &\quad \cdot (\|\mathbf{A}\|_F + \|\mathbf{\Lambda}\|_F). \end{aligned} \quad (20.4.2)$$

Write $\mathbf{X} = \mathbf{W} \mathbf{\Sigma} \mathbf{V}^*$ (Theorem 13.9.1), with $\mathbf{W}, \mathbf{V}$ unitary and $\mathbf{\Sigma} = \text{diag}(\sigma_1, \dots, \sigma_n)$. Here $\sigma_n > 0$, because $\mathbf{X}$ is invertible and so has rank n . Then $\mathbf{X}^{-1} = \mathbf{V} \mathbf{\Sigma}^{-1} \mathbf{W}^*$, and

$$\mathbf{A} = \mathbf{X} \mathbf{\Lambda} \mathbf{X}^{-1} = \mathbf{W} \mathbf{\Sigma} \mathbf{\Lambda} \mathbf{\Sigma}^{-1} \mathbf{W}^*, \quad \mathbf{M} = \mathbf{V}^* \mathbf{\Lambda} \mathbf{V}.$$

Put $\mathbf{N} = \mathbf{W} \mathbf{M} \mathbf{W}^* = (\mathbf{W} \mathbf{V}^*) \mathbf{\Lambda} (\mathbf{W} \mathbf{V}^*)^*$. It is unitarily diagonalizable, hence normal by Corollary 12.4.3. The (i, j) entry of $\mathbf{\Sigma} \mathbf{\Lambda} \mathbf{\Sigma}^{-1} - \mathbf{M}$ is $m_{ij}(\sigma_i/\sigma_j - 1)$. Write $\kappa = \kappa_2(\mathbf{X}) = \sigma_1/\sigma_n$ (Proposition 16.8.2 (d)). If $\sigma_i \geq \sigma_j$, then $0 \leq \sigma_i/\sigma_j - 1 \leq \kappa - 1$. If $\sigma_i < \sigma_j$, then $0 < 1 - \sigma_i/\sigma_j \leq 1 - 1/\kappa \leq \kappa - 1$, the last step because $(\kappa - 1) - (1 - 1/\kappa) = (\kappa - 1)^2/\kappa \geq 0$. So every entry is at most $(\kappa - 1)|m_{ij}|$ in modulus. In the display below, the first equality is Lemma 11.7.5 applied to $\mathbf{A} - \mathbf{N} = \mathbf{W}(\mathbf{\Sigma} \mathbf{\Lambda} \mathbf{\Sigma}^{-1} - \mathbf{M}) \mathbf{W}^*$; the inequality is the entrywise bound just proved, summed over the entries; and the last equality is the same lemma again, applied to $\mathbf{M} = \mathbf{V}^* \mathbf{\Lambda} \mathbf{V}$. So

$$\|\mathbf{A} - \mathbf{N}\|_F = \|\mathbf{\Sigma} \mathbf{\Lambda} \mathbf{\Sigma}^{-1} - \mathbf{M}\|_F \leq (\kappa - 1) \|\mathbf{M}\|_F = (\kappa - 1) \|\mathbf{\Lambda}\|_F.$$

Now $\mathbf{N}^* \mathbf{N} = \mathbf{N} \mathbf{N}^*$, and a direct expansion gives

$$\begin{aligned} \mathbf{A}^* \mathbf{A} - \mathbf{A} \mathbf{A}^* &= (\mathbf{A}^* \mathbf{A} - \mathbf{N}^* \mathbf{N}) - (\mathbf{A} \mathbf{A}^* - \mathbf{N} \mathbf{N}^*) \\ &= \mathbf{A}^* (\mathbf{A} - \mathbf{N}) + (\mathbf{A} - \mathbf{N})^* \mathbf{N} \\ &\quad - \mathbf{A} (\mathbf{A} - \mathbf{N})^* - (\mathbf{A} - \mathbf{N}) \mathbf{N}^*. \end{aligned}$$

The Frobenius norm is submultiplicative (Example (The Frobenius norm is submultiplicative)), $\|\mathbf{Y}^*\|_F = \|\mathbf{Y}\|_F$ for every $\mathbf{Y}$, since the entries of $\mathbf{Y}^*$ are the conjugates of those of $\mathbf{Y}$ and conjugation preserves the modulus, and $\|\mathbf{N}\|_F = \|\mathbf{\Lambda}\|_F$ by Lemma 11.7.5. So each of the four terms has norm at most $\|\mathbf{A} - \mathbf{N}\|_F$ times $\|\mathbf{A}\|_F$ or $\|\mathbf{\Lambda}\|_F$, and

$$\|\mathbf{A}^* \mathbf{A} - \mathbf{A} \mathbf{A}^*\|_F \leq 2 \|\mathbf{A} - \mathbf{N}\|_F (\|\mathbf{A}\|_F + \|\mathbf{\Lambda}\|_F),$$

which with the previous display gives Eq. 20.4.2.

Now suppose the infimum is 1. Every diagonalizing $\mathbf{X}$ produces the same $\|\mathbf{\Lambda}\|_F = (\sum_i |\lambda_i|^2)^{1/2}$, since $\mathbf{\Lambda}$ lists the eigenvalues of $\mathbf{A}$ in some order. So the right side of Eq. 20.4.2 can be made smaller than any positive number, while the left side does not depend on $\mathbf{X}$. Hence $\|\mathbf{A}^* \mathbf{A} - \mathbf{A} \mathbf{A}^*\|_F = 0$, so $\mathbf{A}^* \mathbf{A} = \mathbf{A} \mathbf{A}^*$ and $\mathbf{A}$ is normal. This proves the corollary. ■

Part (b) says more than the bare equivalence. By Eq. 20.4.2, a diagonalizable matrix whose eigenvector matrix is **nearly** unitary is **nearly** normal. It also shows that $\kappa_2(\mathbf{X}) = 1$ forces $\mathbf{A}$ to be normal. In fact $\kappa_2(\mathbf{X}) = 1$ says that $\sigma_1 = \sigma_n$, so $\mathbf{\Sigma} = \sigma_1 \mathbf{I}$ and $\mathbf{X} = \sigma_1 \mathbf{W} \mathbf{V}^*$, a nonzero multiple of a unitary matrix.

Part (a) is sharp, and the example that shows it is the one Chapter 17 §03 used to warn that Weyl's inequality needs Hermitian matrices. Take $\mathbf{A} = \mathbf{O}$, which is normal with the single eigenvalue 0, and $\mathbf{E} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, with $\|\mathbf{E}\|_2 = 1$ because $\mathbf{E}$ is unitary. The eigenvalues of $\mathbf{A} + \mathbf{E}$ are $\pm i$, since $p_{\mathbf{E}}(x) = x^2 + 1$. Weyl's comparison of the i -th eigenvalues cannot even be stated, because $\pm i$ are not real. The Bauer–Fike bound can: each of $\pm i$ is within $\|\mathbf{E}\|_2 = 1$ of 0, and the distance is exactly 1. So the constant 1 in (a) cannot be lowered.

20.4.4 Residuals

In practice one rarely knows $\mathbf{E}$. What one has is an approximate eigenpair $(\mu, \hat{\mathbf{x}})$, computed somehow, and the **residual** $\mathbf{r} = \mathbf{A}\hat{\mathbf{x}} - \mu\hat{\mathbf{x}}$, which can be evaluated. The residual turns into a perturbation: a small residual means that $(\mu, \hat{\mathbf{x}})$ is an exact eigenpair of a nearby matrix.

Theorem 20.4.4. Residual Bound

Let $\mathbf{A} = \mathbf{X}\mathbf{\Lambda}\mathbf{X}^{-1} \in M_n(\mathbb{C})$ be diagonalizable, with $\mathbf{\Lambda} = \text{diag}(\lambda_1, \dots, \lambda_n)$. Let $\mu \in \mathbb{C}$ and $\hat{\mathbf{x}} \in \mathbb{C}^n$ with $\hat{\mathbf{x}} \neq \mathbf{0}$, and let $\eta \geq 0$ satisfy

$$\|\mathbf{A}\hat{\mathbf{x}} - \mu\hat{\mathbf{x}}\|_2 \leq \eta \|\hat{\mathbf{x}}\|_2.$$

Then some eigenvalue λ_i of $\mathbf{A}$ satisfies $|\lambda_i - \mu| \leq \kappa_2(\mathbf{X})\eta$. If $\mathbf{A}$ is normal, some eigenvalue satisfies $|\lambda_i - \mu| \leq \eta$.

Proof. Put $\mathbf{r} = \mathbf{A}\hat{\mathbf{x}} - \mu\hat{\mathbf{x}}$ and $\mathbf{E} = -\mathbf{r}\hat{\mathbf{x}}^*/\|\hat{\mathbf{x}}\|_2^2$. Then

$$(\mathbf{A} + \mathbf{E})\hat{\mathbf{x}} = \mathbf{A}\hat{\mathbf{x}} - \mathbf{r} \frac{\hat{\mathbf{x}}^*\hat{\mathbf{x}}}{\|\hat{\mathbf{x}}\|_2^2} = \mathbf{A}\hat{\mathbf{x}} - \mathbf{r} = \mu\hat{\mathbf{x}},$$

so μ is an eigenvalue of $\mathbf{A} + \mathbf{E}$ with eigenvector $\hat{\mathbf{x}}$. For every $\mathbf{z}$, the Cauchy–Schwarz inequality (Theorem 11.1.5) gives $|\hat{\mathbf{x}}^*\mathbf{z}| \leq \|\hat{\mathbf{x}}\|_2 \|\mathbf{z}\|_2$, hence

$$\|\mathbf{E}\mathbf{z}\|_2 = \frac{|\hat{\mathbf{x}}^*\mathbf{z}| \|\mathbf{r}\|_2}{\|\hat{\mathbf{x}}\|_2^2} \leq \frac{\|\mathbf{r}\|_2}{\|\hat{\mathbf{x}}\|_2} \|\mathbf{z}\|_2,$$

so $\|\mathbf{E}\|_2 \leq \|\mathbf{r}\|_2/\|\hat{\mathbf{x}}\|_2 \leq \eta$. (The first inequality is an equality, since $\mathbf{E}\hat{\mathbf{x}} = -\mathbf{r}$.) By [Theorem 20.4.2](#) with $p = 2$, some λ_i satisfies $|\lambda_i - \mu| \leq \kappa_2(\mathbf{X})\|\mathbf{E}\|_2 \leq \kappa_2(\mathbf{X})\eta$. If $\mathbf{A}$ is normal, use [Corollary 20.4.3](#) (a) instead. ■

The theorem reverses the usual question. Instead of asking how far the eigenvalues of a given perturbed matrix have moved, it asks: of which nearby matrix is my computed answer exactly right? The answer is $\mathbf{A} + \mathbf{E}$ with $\|\mathbf{E}\|_2 = \|\mathbf{r}\|_2/\|\hat{\mathbf{x}}\|_2$. An error analysis of this kind, which measures an answer by a change in the data that makes it exact, is a **backward error analysis**. Chapter 24 builds on it. For a normal matrix, the scaled residual norm $\|\mathbf{r}\|_2/\|\hat{\mathbf{x}}\|_2$ is itself a certified error bar for μ .

Example. An error bar from one product

Let $\mathbf{A} = \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix}$, $\hat{\mathbf{x}} = (1, 2)$, and $\mu = \frac{7}{2}$. Compute the residual, and say what interval [Theorem 20.4.4](#) certifies around μ . Then compare it with the truth.

Solution. $\mathbf{A}\hat{\mathbf{x}} = (4, 7)$ and $\mu\hat{\mathbf{x}} = (\frac{7}{2}, 7)$, so $\mathbf{r} = (\frac{1}{2}, 0)$, with $\|\mathbf{r}\|_2 = \frac{1}{2}$ and $\|\hat{\mathbf{x}}\|_2 = \sqrt{5}$. Take $\eta = \|\mathbf{r}\|_2/\|\hat{\mathbf{x}}\|_2 = \sqrt{5}/10 \approx 0.2236$. The matrix $\mathbf{A}$ is real symmetric, hence normal, so [Theorem 20.4.4](#) gives an eigenvalue of $\mathbf{A}$ within η of $\frac{7}{2}$, and that eigenvalue is real (Theo-

rem 12.2.1). So $\mathbf{A}$ has an eigenvalue in $[\frac{7}{2} - \eta, \frac{7}{2} + \eta] \approx [3.276, 3.724]$, and no eigenvalue was computed to say so: one matrix–vector product did it.

The truth. $p_{\mathbf{A}}(x) = (x-2)(x-3) - 1 = x^2 - 5x + 5$, with roots $(5 \pm \sqrt{5})/2$. The larger is ≈ 3.618 , which does lie in the interval, at distance ≈ 0.118 from $\frac{7}{2}$ — about half of η , so the certificate costs a factor of 2 in sharpness.

20.4.5 A worked example

Chapter 12 §11 perturbed a corner of $\begin{pmatrix} 1 & M \\ 0 & 1 \end{pmatrix}$, with a double eigenvalue, and found the eigenvalues moving by $\sqrt{M}\varepsilon$ (Example (A perturbation that moves the eigenvalues a thousand times as far)). Its closing remark said that “a matrix with well-separated eigenvalues and a large defect behaves less dramatically”. Bauer–Fike says how much less.

Example. A large corner, separated eigenvalues

Let $M > 0$, $\mathbf{A} = \begin{pmatrix} 1 & M \\ 0 & 2 \end{pmatrix}$, and $\mathbf{E} = \varepsilon \mathbf{e}_2 \mathbf{e}_1^\top$ with $0 < \varepsilon \leq 1$. Diagonalize $\mathbf{A}$ in two ways, compute κ_2 of each eigenvector matrix, and compare the Bauer–Fike bound with the exact movement of the eigenvalues.

Solution. Eigenvectors. $\mathbf{A}$ is upper triangular with distinct diagonal entries 1, 2, which are its eigenvalues. For 1, $\mathbf{A} - \mathbf{I} = \begin{pmatrix} 0 & M \\ 0 & 1 \end{pmatrix}$ has kernel $\text{span}(\mathbf{e}_1)$. For 2, $\mathbf{A} - 2\mathbf{I} = \begin{pmatrix} -1 & M \\ 0 & 0 \end{pmatrix}$ has kernel spanned by $(M, 1)$. Two natural choices are

$$\mathbf{X}_1 = \begin{pmatrix} 1 & M \\ 0 & 1 \end{pmatrix}, \quad \mathbf{X}_2 = \begin{pmatrix} 1 & M/s \\ 0 & 1/s \end{pmatrix}, \quad s = \sqrt{1 + M^2}.$$

The second has columns of length 1. Both satisfy $\mathbf{X}^{-1}\mathbf{A}\mathbf{X} = \text{diag}(1, 2)$.

$\kappa_2(\mathbf{X}_1)$. $\mathbf{X}_1^* \mathbf{X}_1 = \begin{pmatrix} 1 & M \\ M & 1+M^2 \end{pmatrix}$ has trace $2 + M^2$ and determinant 1. Its eigenvalues $\sigma_1^2 \geq \sigma_2^2$ therefore have product 1, and $\kappa_2(\mathbf{X}_1) = \sigma_1/\sigma_2 = \sigma_1^2$. Solving the quadratic,

$$\kappa_2(\mathbf{X}_1) = \frac{2 + M^2 + M\sqrt{M^2 + 4}}{2},$$

which exceeds M^2 .

$\kappa_2(\mathbf{X}_2)$. $\mathbf{X}_2^* \mathbf{X}_2 = \begin{pmatrix} 1 & q \\ q & 1 \end{pmatrix}$ with $q = M/s$, since the columns are unit vectors whose inner product is M/s . Its eigenvalues are $1 \pm q$, so

$$\kappa_2(\mathbf{X}_2)^2 = \frac{1+q}{1-q} = \frac{s+M}{s-M} = (s+M)^2,$$

using $(s+M)(s-M) = s^2 - M^2 = 1$. Hence $\kappa_2(\mathbf{X}_2) = M + \sqrt{1 + M^2}$, which lies between $2M$ and $2M+1$.

The bound. $\|\mathbf{E}\|_2 = \varepsilon$, since $\mathbf{E}$ has a single nonzero entry. With $\mathbf{X}_2$, Theorem 20.4.2 gives movement at most $(M + \sqrt{1 + M^2})\varepsilon$, about $2M\varepsilon$. With $\mathbf{X}_1$ it gives about $M^2\varepsilon$.

The truth. $p_{\mathbf{A}+\mathbf{E}}(x) = (x-1)(x-2) - M\varepsilon = x^2 - 3x + 2 - M\varepsilon$, with roots $(3 \pm \sqrt{1 + 4M\varepsilon})/2$.

Each eigenvalue has moved by

$$\frac{\sqrt{1+4M\varepsilon}-1}{2} = \frac{2M\varepsilon}{1+\sqrt{1+4M\varepsilon}},$$

multiplying top and bottom by $\sqrt{1+4M\varepsilon}+1$. The denominator is at least 2, and at most $2+2M\varepsilon$ because $\sqrt{1+4M\varepsilon} \leq 1+2M\varepsilon$. So the movement lies between $M\varepsilon/(1+M\varepsilon)$ and $M\varepsilon$.

Numbers. For $M = 100$ and $\varepsilon = 10^{-6}$, the eigenvalues move by about 0.9999×10^{-4} , and the bound with $\mathbf{X}_2$ is about 2.0×10^{-4} . The bound with $\mathbf{X}_1$ is about 10^{-2} . By comparison, in Example (A perturbation that moves the eigenvalues a thousand times as far), the same corner and the same ε moved a **double** eigenvalue by $\sqrt{M\varepsilon} = 10^{-2}$. Separating the eigenvalues turns a square-root amplification into a linear one, with a constant of about M . The constant is still large, and $\kappa_2(\mathbf{X}_2)$ predicts it correctly to within a factor of about 2.

Warning. $\kappa_2(\mathbf{X})$ belongs to the choice of $\mathbf{X}$, not to $\mathbf{A}$. Eigenvectors can be rescaled, and rescaling changes κ_2 : in [Example \(A large corner, separated eigenvalues\)](#) the same matrix gave $\kappa_2(\mathbf{X}_1) \approx 10^4$ and $\kappa_2(\mathbf{X}_2) \approx 200$ at $M = 100$. The theorem holds for every choice, so use the best one you can find; normalizing the columns is a reasonable first step. And even the best $\mathbf{X}$ measures all the eigenvalues at once. Section 11 attaches a separate condition number to each simple eigenvalue, and one eigenvalue can be well conditioned while $\mathbf{X}$ is badly conditioned.

20.4.6 Defective matrices

When $\mathbf{A}$ is not diagonalizable, write $\mathbf{A} = \mathbf{XJX}^{-1}$ with $\mathbf{J}$ a Jordan form (Theorem 10.3.1). The letter $\mathbf{J}$ is local to this subsection; elsewhere in the book $\mathbf{J}$ is the all-ones matrix, which does not appear here. The proof of [Theorem 20.4.2](#) goes through verbatim, except that $(\mathbf{J} - \mu\mathbf{I})^{-1}$ is no longer diagonal. Its norm is no longer $1/\delta$; it grows like $1/\delta^m$ as $\delta \rightarrow 0$, where m is the size of the largest Jordan block. Taking m -th roots brings back the exponent of Section 3.

Theorem 20.4.5. Bauer–Fike for a Jordan Form

Let $\mathbf{A} = \mathbf{XJX}^{-1} \in M_n(\mathbb{C})$, where $\mathbf{X}$ is invertible and $\mathbf{J}$ is a Jordan form of $\mathbf{A}$ whose largest block has size m . Let $\lambda_1, \dots, \lambda_n$ be the eigenvalues of $\mathbf{A}$, let $\mathbf{E} \in M_n(\mathbb{C})$, and put $\theta = m\kappa_2(\mathbf{X})\|\mathbf{E}\|_2$. Then every eigenvalue μ of $\mathbf{A} + \mathbf{E}$ satisfies

$$\min_{1 \leq i \leq n} |\lambda_i - \mu| \leq \max(\theta, \theta^{1/m}).$$

Idea. Run the proof of [Theorem 20.4.2](#) with $\mathbf{J}$ in place of $\mathbf{A}$. The only new work is to bound $\|(\mathbf{J}_k(\lambda) - \mu\mathbf{I})^{-1}\|_2$ for one block. The block is $(\lambda - \mu)\mathbf{I} + \mathbf{N}$, with $\mathbf{N}$ nilpotent of norm at most 1, so its inverse is a **finite** Neumann series $\sum_{j < k} (-1)^j \mathbf{N}^j / (\lambda - \mu)^{j+1}$ with at most m terms. Bound it by $m\delta^{-1}$ when $\delta \geq 1$ and by $m\delta^{-m}$ when $\delta < 1$.

Proof. Put $\delta = \min_i |\lambda_i - \mu|$. If $\delta = 0$ there is nothing to prove, so suppose $\delta > 0$. The upper triangular $\mathbf{J}$ has its diagonal entries as eigenvalues (Theorem 9.7.4 (b)), and these are the λ_i because $\mathbf{J}$ and $\mathbf{A}$ are similar (Theorem 7.10.3); so $\mu \notin \text{spec}(\mathbf{J})$. As in the proof of Theorem 20.4.2, μ is an eigenvalue of $\mathbf{J} + \mathbf{X}^{-1}\mathbf{E}\mathbf{X}$, and Lemma 20.4.1 with submultiplicativity gives

$$1 \leq \|(\mathbf{J} - \mu\mathbf{I})^{-1}\|_2 \kappa_2(\mathbf{X}) \|\mathbf{E}\|_2.$$

One block. Let $\mathbf{J}_k(\lambda)$ be a block of $\mathbf{J}$, so $k \leq m$ and $|\lambda - \mu| \geq \delta$, and write $\mathbf{J}_k(\lambda) - \mu\mathbf{I} = a\mathbf{I} + \mathbf{N}$ with $a = \lambda - \mu \neq 0$ and $\mathbf{N} = \mathbf{J}_k(0)$. The matrix $\mathbf{N}$ sends $\mathbf{e}_1$ to $\mathbf{0}$ and $\mathbf{e}_j$ to $\mathbf{e}_{j-1}$ for $j \geq 2$, so $\mathbf{N}^k$ sends every $\mathbf{e}_j$ to $\mathbf{0}$ and $\mathbf{N}^k = \mathbf{0}$. Then

$$(a\mathbf{I} + \mathbf{N}) \sum_{j=0}^{k-1} \frac{(-1)^j \mathbf{N}^j}{a^{j+1}} = \sum_{j=0}^{k-1} \frac{(-1)^j \mathbf{N}^j}{a^j} - \sum_{j=1}^k \frac{(-1)^j \mathbf{N}^j}{a^j} = \mathbf{I} - \frac{(-1)^k \mathbf{N}^k}{a^k} = \mathbf{I},$$

so the sum is the inverse of $a\mathbf{I} + \mathbf{N}$. The same description of $\mathbf{N}$ gives $\|\mathbf{N}\mathbf{x}\|_2^2 = \sum_{j \geq 2} |x_j|^2 \leq \|\mathbf{x}\|_2^2$, so $\|\mathbf{N}\|_2 \leq 1$. Hence $\|\mathbf{N}^j\|_2 \leq 1$ for every $j \geq 0$ by submultiplicativity, and the triangle inequality (Theorem 16.3.2 (b)) gives

$$\|(\mathbf{J}_k(\lambda) - \mu\mathbf{I})^{-1}\|_2 \leq \sum_{j=1}^k |a|^{-j} \leq \sum_{j=1}^m \delta^{-j},$$

using $|a| \geq \delta$ and $k \leq m$.

All blocks. $(\mathbf{J} - \mu\mathbf{I})^{-1}$ is block diagonal, with the inverses of the blocks $\mathbf{J}_k(\lambda) - \mu\mathbf{I}$ on its diagonal. Split $\mathbf{x}$ into the matching pieces $\mathbf{x}_1, \dots, \mathbf{x}_r$ and let $\mathbf{C}_1, \dots, \mathbf{C}_r$ be the diagonal blocks. Then $\|(\mathbf{J} - \mu\mathbf{I})^{-1}\mathbf{x}\|_2^2 = \sum_l \|\mathbf{C}_l \mathbf{x}_l\|_2^2 \leq (\max_l \|\mathbf{C}_l\|_2)^2 \|\mathbf{x}\|_2^2$. So $\|(\mathbf{J} - \mu\mathbf{I})^{-1}\|_2 \leq \sum_{j=1}^m \delta^{-j}$, and

$$1 \leq \frac{\theta}{m} \sum_{j=1}^m \delta^{-j}.$$

Two cases. If $\delta \geq 1$, each $\delta^{-j} \leq \delta^{-1}$, so $1 \leq \theta/\delta$ and $\delta \leq \theta$. If $\delta < 1$, each $\delta^{-j} \leq \delta^{-m}$, so $1 \leq \theta\delta^{-m}$, that is $\delta^m \leq \theta$ and $\delta \leq \theta^{1/m}$. In both cases $\delta \leq \max(\theta, \theta^{1/m})$. This proves the theorem. ■

For $m = 1$, $\mathbf{J}$ is diagonal, $\theta = \kappa_2(\mathbf{X})\|\mathbf{E}\|_2$, and the theorem is Theorem 20.4.2 with $p = 2$. For $m \geq 2$ and small $\mathbf{E}$, the term $\theta^{1/m}$ dominates, and the eigenvalues move like the m -th root of the perturbation. That exponent is the Jordan block's signature, and it is the answer to Chapter 10 §03's warning that the Jordan form is not stable. Section 3's Example (The Jordan block meets Elsner's bound) shows it cannot be improved. There $\mathbf{A} = \mathbf{J}_n(0)$, $\mathbf{X} = \mathbf{I}$, $m = n$, and $\|\mathbf{E}\|_2 = \varepsilon \leq 1/n$, so $\theta = n\varepsilon \leq 1$ and the theorem gives $\max(\theta, \theta^{1/n}) = (n\varepsilon)^{1/n}$ while the eigenvalues actually move $\varepsilon^{1/n}$. The two differ by the factor $n^{1/n}$, which never exceeds $3^{1/3} < 1.45$.

Only the size of the largest block enters, and a diagonalizable part of $\mathbf{A}$ is charged the same m -th root as the defective part. Section 11 separates the eigenvalues. There a simple eigenvalue gets its own first-order condition number, and that number becomes infinite exactly when two eigenvalues merge into a Jordan block.

20.4.7 Exercises

20.4.7.1 A. Check your understanding

Exercise. A1

- (a) State [Theorem 20.4.2](#) with all its hypotheses.
- (b) True or false: the constant $\kappa_2(\mathbf{X})$ in [Theorem 20.4.2](#) is determined by $\mathbf{A}$. Justify your answer.
- (c) True or false: for every $\mathbf{A}, \mathbf{E} \in M_n(\mathbb{C})$, every eigenvalue of $\mathbf{A} + \mathbf{E}$ lies within $\|\mathbf{E}\|_2$ of an eigenvalue of $\mathbf{A}$. Justify your answer.
- (d) What does [Corollary 20.4.3](#) (a) give when $\mathbf{A}$ is normal, and which example shows that its constant cannot be lowered?
- (e) In [Theorem 20.4.5](#), what exponent appears for small $\|\mathbf{E}\|_2$, and what decides it?

Solution.

- (a) See [Theorem 20.4.2](#): $\mathbf{A} = \mathbf{X}\mathbf{\Lambda}\mathbf{X}^{-1}$ diagonalizable, $\mathbf{E}$ arbitrary, $p \in \{1, 2, \infty\}$; every eigenvalue μ of $\mathbf{A} + \mathbf{E}$ is within $\kappa_p(\mathbf{X})\|\mathbf{E}\|_p$ of some λ_i .
- (b) False. The columns of $\mathbf{X}$ can be rescaled, and κ_2 changes. The Quick check gave $\mathbf{X} = \text{diag}(a, b)$ for $\mathbf{A} = \text{diag}(1, 2)$, with $\kappa_2(\mathbf{X}) = \max(|a|, |b|)/\min(|a|, |b|)$, any number ≥ 1 .
- (c) False. By Example (The Jordan block meets Elsner's bound) with $n = 2$, $\mathbf{A} = \mathbf{J}_2(0)$ and $\mathbf{E} = \varepsilon \mathbf{e}_2 \mathbf{e}_1^\top$ with $0 < \varepsilon < 1$ give eigenvalues $\pm\sqrt{\varepsilon}$ of $\mathbf{A} + \mathbf{E}$, at distance $\sqrt{\varepsilon} > \varepsilon = \|\mathbf{E}\|_2$ from 0. The hypothesis that fails is diagonalizability.
- (d) Every eigenvalue of $\mathbf{A} + \mathbf{E}$ is within $\|\mathbf{E}\|_2$ of an eigenvalue of $\mathbf{A}$. The example $\mathbf{A} = \mathbf{O}$, $\mathbf{E} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ has perturbed eigenvalues $\pm i$ at distance exactly $1 = \|\mathbf{E}\|_2$.
- (e) The exponent $1/m$, where m is the size of the largest Jordan block of $\mathbf{A}$.

20.4.7.2 B. Practice

Exercise. B1: A Bauer–Fike region

Let $\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}$.

- (a) Find an eigenvector matrix $\mathbf{X}$ with columns of length 1, and compute $\kappa_2(\mathbf{X})$.
- (b) Hence find r such that every eigenvalue of $\mathbf{A} + \mathbf{E}$ lies within r of 1 or of 3 whenever $\|\mathbf{E}\|_2 \leq 0.01$.
- (c) Compute the eigenvalues of $\mathbf{A} + 0.01 \mathbf{e}_2 \mathbf{e}_1^\top$ exactly, and compare their movement with r .

Solution.

- (a) The eigenvalues are 1 and 3, with eigenvectors $\mathbf{e}_1$ and $(1, 1)$, since $(\mathbf{A} - 3\mathbf{I})(1, 1) = (-2 + 2, 0) = \mathbf{0}$. So $\mathbf{X} = \begin{pmatrix} 1 & 1/\sqrt{2} \\ 0 & 1/\sqrt{2} \end{pmatrix}$. Then $\mathbf{X}^* \mathbf{X} = \begin{pmatrix} 1 & q \\ q & 1 \end{pmatrix}$ with $q = 1/\sqrt{2}$, with

eigenvalues $1 \pm q$, so

$$\kappa_2(\mathbf{X})^2 = \frac{1+q}{1-q} = \frac{\sqrt{2}+1}{\sqrt{2}-1} = (\sqrt{2}+1)^2,$$

and $\kappa_2(\mathbf{X}) = 1 + \sqrt{2} \approx 2.414$.

- (b) By [Theorem 20.4.2](#) with $p = 2$, $r = (1 + \sqrt{2})(0.01) \approx 0.0241$ works.
- (c) $p(x) = (x-1)(x-3) - 0.02 = x^2 - 4x + 2.98$, with roots $2 \pm \sqrt{1.02}$. Each eigenvalue has moved by $\sqrt{1.02} - 1 \approx 0.00995$, well within $r \approx 0.0241$.

Exercise. B2: A residual certificate

Let $\mathbf{A} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$, $\hat{\mathbf{x}} = (10, 9)$, and $\mu = 3$.

- (a) Compute the residual $\mathbf{r} = \mathbf{A}\hat{\mathbf{x}} - \mu\hat{\mathbf{x}}$ and $\eta = \|\mathbf{r}\|_2 / \|\hat{\mathbf{x}}\|_2$.
- (b) Hence show that $\mathbf{A}$ has an eigenvalue in the interval $[3 - \eta, 3 + \eta]$, without computing the eigenvalues of $\mathbf{A}$.

Solution.

- (a) $\mathbf{A}\hat{\mathbf{x}} = (29, 28)$ and $3\hat{\mathbf{x}} = (30, 27)$, so $\mathbf{r} = (-1, 1)$, $\|\mathbf{r}\|_2 = \sqrt{2}$, $\|\hat{\mathbf{x}}\|_2 = \sqrt{181}$, and $\eta = \sqrt{2/181} \approx 0.105$.
- (b) $\mathbf{A}$ is real symmetric, hence normal. By [Theorem 20.4.4](#), some eigenvalue λ satisfies $|\lambda - 3| \leq \eta$. The eigenvalues of a real symmetric matrix are real ([Theorem 12.2.1](#)), so $\lambda \in [3 - \eta, 3 + \eta]$. (In fact $\lambda = 3$, with eigenvector $(1, 1)$.)

Exercise. B3: A defective bound

Let $\mathbf{A} = \mathbf{J}_2(3)$.

- (a) Explain why [Theorem 20.4.2](#) does not apply to $\mathbf{A}$.
- (b) Use [Theorem 20.4.5](#) to find r such that every eigenvalue of $\mathbf{A} + \mathbf{E}$ lies within r of 3 whenever $\|\mathbf{E}\|_2 \leq 10^{-4}$.
- (c) Compute the eigenvalues of $\mathbf{A} + 10^{-4}\mathbf{e}_2\mathbf{e}_1^\top$, and compare.

Solution.

- (a) $\mathbf{A}$ has the single eigenvalue 3 with algebraic multiplicity 2, and $\mathbf{A} - 3\mathbf{I} = \mathbf{J}_2(0)$ has rank 1, so the eigenspace has dimension 1. Hence $\mathbf{A}$ is not diagonalizable ([Theorem 9.4.1](#)), and there is no $\mathbf{X}$ to put in [Theorem 20.4.2](#).
- (b) $\mathbf{A}$ is its own Jordan form, so take $\mathbf{X} = \mathbf{I}$, with $\kappa_2(\mathbf{I}) = 1$, and $m = 2$. Then $\theta = 2 \cdot 1 \cdot 10^{-4} = 2 \times 10^{-4}$, and $\max(\theta, \theta^{1/2}) = \sqrt{2} \times 10^{-2} \approx 0.0141$. So $r = \sqrt{2} \times 10^{-2}$ works.
- (c) $p(x) = (x-3)^2 - 10^{-4}$, with roots 3 ± 10^{-2} . The movement 0.01 is below $r \approx 0.0141$, and it is a hundred times the size 10^{-4} of the perturbation: the square root at work.

20.4.7.3 C. Going deeper

Exercise. C1: One eigenvalue in each disc

Let $\mathbf{A} = \mathbf{X}\mathbf{\Lambda}\mathbf{X}^{-1}$ have n distinct eigenvalues $\lambda_1, \dots, \lambda_n$, put $g = \min_{i \neq j} |\lambda_i - \lambda_j|$, and let $\mathbf{E}$ satisfy $r = \kappa_2(\mathbf{X})\|\mathbf{E}\|_2 < g/2$.

- Prove that for each i , $\mathbf{A} + \mathbf{E}$ has exactly one eigenvalue in the disc $\{z : |z - \lambda_i| \leq r\}$, and that $\mathbf{A} + \mathbf{E}$ is diagonalizable.
- Deduce that if $\mathbf{A}$ and $\mathbf{E}$ are real and every λ_i is real, then every eigenvalue of $\mathbf{A} + \mathbf{E}$ is real.

Hint: apply Lemma 20.3.4 to the path $\mathbf{A} + t\mathbf{E}$.

Solution.

- Let $Q_i = \{z : |z - \lambda_i| \leq r\}$. For $i \neq j$, $z \in Q_i$, and $w \in Q_j$, the triangle inequality (Theorem 0.6.4) gives $|z - w| \geq |\lambda_i - \lambda_j| - 2r \geq g - 2r > 0$. So Q_i and the union L_i of the other Q_j are at distance at least $g - 2r > 0$. For $t \in [0, 1]$, Theorem 20.4.2 applied to $\mathbf{A}$ and $t\mathbf{E}$ puts every eigenvalue of $\mathbf{A} + t\mathbf{E}$ within $\kappa_2(\mathbf{X})\|t\mathbf{E}\|_2 \leq r$ of some λ_j , that is, in $Q_i \cup L_i$. The path $\mathbf{A} + t\mathbf{E} = (1 - t)\mathbf{A} + t(\mathbf{A} + \mathbf{E})$ is of the form in Lemma 20.3.4, so the number of eigenvalues in Q_i is the same at $t = 0$ and $t = 1$. At $t = 0$ it is 1: $\lambda_i \in Q_i$, and $\lambda_j \in Q_j \subseteq L_i$ is not in Q_i for $j \neq i$. So $\mathbf{A} + \mathbf{E}$ has exactly one eigenvalue in each Q_i . The Q_i are pairwise disjoint, so these n eigenvalues are distinct, and $\mathbf{A} + \mathbf{E}$ is diagonalizable by Corollary 9.4.2.
- Let μ be the eigenvalue of $\mathbf{A} + \mathbf{E}$ in Q_i . Since λ_i is real, $|\bar{\mu} - \lambda_i| = |\mu - \lambda_i| \leq r$, so $\bar{\mu} \in Q_i$. Since $\mathbf{A} + \mathbf{E}$ is real, $\bar{\mu}$ is also an eigenvalue (Theorem 9.2.5). By (a), Q_i holds only one eigenvalue, so $\bar{\mu} = \mu$, and μ is real. Every eigenvalue of $\mathbf{A} + \mathbf{E}$ lies in some Q_i , so all of them are real.

Exercise. C2: The best shift for a residual

Let $\mathbf{A} \in M_n(\mathbb{C})$ and $\hat{\mathbf{x}} \in \mathbb{C}^n$ with $\hat{\mathbf{x}} \neq \mathbf{0}$, and put $\rho = \hat{\mathbf{x}}^* \mathbf{A} \hat{\mathbf{x}} / \hat{\mathbf{x}}^* \hat{\mathbf{x}}$.

- Prove that $\|\mathbf{A}\hat{\mathbf{x}} - \mu\hat{\mathbf{x}}\|_2 \geq \|\mathbf{A}\hat{\mathbf{x}} - \rho\hat{\mathbf{x}}\|_2$ for every $\mu \in \mathbb{C}$, with equality only for $\mu = \rho$.
- Deduce that if $\mathbf{A}$ is normal, then $\mathbf{A}$ has an eigenvalue λ with $|\lambda - \rho| \leq \|\mathbf{A}\hat{\mathbf{x}} - \mu\hat{\mathbf{x}}\|_2 / \|\hat{\mathbf{x}}\|_2$ for every $\mu \in \mathbb{C}$.

Solution.

- Put $\mathbf{w} = \mathbf{A}\hat{\mathbf{x}} - \rho\hat{\mathbf{x}}$. Then $\langle \mathbf{w}, \hat{\mathbf{x}} \rangle = \hat{\mathbf{x}}^* \mathbf{w} = \hat{\mathbf{x}}^* \mathbf{A}\hat{\mathbf{x}} - \rho \hat{\mathbf{x}}^* \hat{\mathbf{x}} = 0$ by the choice of ρ . For any μ , $\mathbf{A}\hat{\mathbf{x}} - \mu\hat{\mathbf{x}} = \mathbf{w} + (\rho - \mu)\hat{\mathbf{x}}$ is a sum of two orthogonal vectors, so by the Pythagorean theorem (Theorem 11.1.3)

$$\|\mathbf{A}\hat{\mathbf{x}} - \mu\hat{\mathbf{x}}\|_2^2 = \|\mathbf{w}\|_2^2 + |\rho - \mu|^2 \|\hat{\mathbf{x}}\|_2^2.$$

This is at least $\|\mathbf{w}\|_2^2$, with equality exactly when $|\rho - \mu| \|\hat{\mathbf{x}}\|_2 = 0$, that is, when $\mu = \rho$, since $\hat{\mathbf{x}} \neq \mathbf{0}$.

- Apply Theorem 20.4.4 for normal $\mathbf{A}$ with the shift ρ and $\eta = \|\mathbf{A}\hat{\mathbf{x}} - \rho\hat{\mathbf{x}}\|_2 / \|\hat{\mathbf{x}}\|_2$. It

gives an eigenvalue λ with $|\lambda - \rho| \leq \eta$, and $\eta \leq \|\mathbf{A}\hat{\mathbf{x}} - \mu\hat{\mathbf{x}}\|_2 / \|\hat{\mathbf{x}}\|_2$ for every μ by (a).