

17.3 How Far Eigenvalues Move

Chapter 16 §07 asked what happens to the eigenvalues of a matrix when the matrix is nudged, and came back with a discouraging answer: they move continuously, and nothing better. Example (A Jordan block's eigenvalues move like $\varepsilon^{1/k}$) exhibits a perturbation of size ε that moves every eigenvalue a distance $\varepsilon^{1/k}$, so no Lipschitz constant survives. That section also promised that the Hermitian case is different, and left the proof here.

The reason the Hermitian case is different is now available. Theorem 17.2.3 describes $\lambda_k(\mathbf{A})$ as an optimization over subspaces, and an optimization over subspaces reacts to a change in $\mathbf{A}$ the way any optimization reacts to a change in the objective: by at most the change in the objective. Turning that sentence into inequalities is the business of this section, and the inequalities are sharper than the sentence suggests, because the index may be shifted as well.

Throughout this section the field is $\mathbb{R}$ or $\mathbb{C}$, and every matrix whose eigenvalues are written down is Hermitian, that is, $\mathbf{A}^* = \mathbf{A}$ with $\mathbf{A} \in M_n(F)$ and $n \geq 1$. Its eigenvalues are real and are always indexed **decreasingly**, $\lambda_1(\mathbf{A}) \geq \lambda_2(\mathbf{A}) \geq \cdots \geq \lambda_n(\mathbf{A})$, as they have been since Chapter 13. Sums of Hermitian matrices are Hermitian, so $\mathbf{A} + \mathbf{B}$ is again a matrix this section may write λ_i of.

17.3.1 Two small facts about eigenvalue lists

One small fact is used in every proof below, and it is the reason the index bookkeeping works out. If a vector lies in the span of the *bottom* eigenvectors of $\mathbf{A}$, its Rayleigh quotient cannot exceed the largest eigenvalue in that span; if it lies in the span of the *top* ones, the quotient cannot fall below the smallest eigenvalue there.

Lemma 17.3.1. The Quotient on a Span of Eigenvectors

Let $\mathbf{A} \in M_n(F)$ be Hermitian and let $(\mathbf{u}_1, \dots, \mathbf{u}_n)$ be an orthonormal basis of F^n with $\mathbf{A}\mathbf{u}_p = \lambda_p(\mathbf{A})\mathbf{u}_p$ for every p . Fix i with $1 \leq i \leq n$.

- (a) If $\mathbf{x} \neq \mathbf{0}$ lies in $\text{span}\{\mathbf{u}_i, \mathbf{u}_{i+1}, \dots, \mathbf{u}_n\}$, then $R_{\mathbf{A}}(\mathbf{x}) \leq \lambda_i(\mathbf{A})$.
- (b) If $\mathbf{x} \neq \mathbf{0}$ lies in $\text{span}\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_i\}$, then $R_{\mathbf{A}}(\mathbf{x}) \geq \lambda_i(\mathbf{A})$.

Proof. Such a basis exists by the spectral theorem: Corollary 12.4.3 over $\mathbb{C}$, where a Hermitian matrix is normal and its eigenvalues may be listed in any prescribed order, and Corollary 12.5.2 over $\mathbb{R}$, where Hermitian means symmetric.

- (a) Write $\mathbf{x} = \sum_{p \geq i} c_p \mathbf{u}_p$, with the c_p not all zero. Since the $\mathbf{u}_p$ are orthonormal,

$$\begin{aligned} \langle \mathbf{A}\mathbf{x}, \mathbf{x} \rangle &= \left\langle \sum_{p \geq i} c_p \lambda_p(\mathbf{A}) \mathbf{u}_p, \sum_{q \geq i} c_q \mathbf{u}_q \right\rangle = \sum_{p \geq i} \lambda_p(\mathbf{A}) |c_p|^2, \\ \|\mathbf{x}\|^2 &= \sum_{p \geq i} |c_p|^2. \end{aligned}$$

Every index in these sums satisfies $p \geq i$, so $\lambda_p(\mathbf{A}) \leq \lambda_i(\mathbf{A})$ by the decreasing indexing, and therefore $\langle \mathbf{A}\mathbf{x}, \mathbf{x} \rangle \leq \lambda_i(\mathbf{A}) \|\mathbf{x}\|^2$. Since $\mathbf{x} \neq \mathbf{0}$ we have $\|\mathbf{x}\|^2 > 0$ and may divide, which gives $R_{\mathbf{A}}(\mathbf{x}) \leq \lambda_i(\mathbf{A})$ by Definition 17.1.1.

- (b) The same computation with $\mathbf{x} = \sum_{p \leq i} c_p \mathbf{u}_p$ gives $\langle \mathbf{A}\mathbf{x}, \mathbf{x} \rangle = \sum_{p \leq i} \lambda_p(\mathbf{A}) |c_p|^2$, and now every index satisfies $p \leq i$, so $\lambda_p(\mathbf{A}) \geq \lambda_i(\mathbf{A})$ and the inequality reverses. This proves

the lemma. ■

The two halves are the two ways of using an eigenvector span: (a) caps a quotient from above, (b) props it up from below. The proof of Weyl's inequalities needs only (a), applied to a single vector for two different matrices at once; (b) is the form in which the same fact enters the witness half of a min-max argument.

A second small fact converts statements about the top of an eigenvalue list into statements about the bottom. Negating a matrix negates its eigenvalues, and negation reverses the order of real numbers.

Lemma 17.3.2. Negation Reverses the List

Let $\mathbf{A} \in M_n(F)$ be Hermitian. Then $-\mathbf{A}$ is Hermitian and

$$\lambda_p(-\mathbf{A}) = -\lambda_{n+1-p}(\mathbf{A}) \quad (1 \leq p \leq n).$$

Proof. $(-\mathbf{A})^* = -\mathbf{A}^* = -\mathbf{A}$, so $-\mathbf{A}$ is Hermitian. By the spectral theorem quoted in the proof of Lemma 17.3.1, $\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{U}^*$ with $\mathbf{D} = \text{diag}(\lambda_1(\mathbf{A}), \dots, \lambda_n(\mathbf{A}))$ and $\mathbf{U}$ unitary, whence $-\mathbf{A} = \mathbf{U}(-\mathbf{D})\mathbf{U}^*$. So the eigenvalue list of $-\mathbf{A}$, with multiplicity, is $-\lambda_1(\mathbf{A}), \dots, -\lambda_n(\mathbf{A})$. Multiplying by -1 reverses the order of real numbers, so listing these decreasingly gives

$$-\lambda_n(\mathbf{A}) \geq -\lambda_{n-1}(\mathbf{A}) \geq \dots \geq -\lambda_1(\mathbf{A}),$$

whose p -th entry is $-\lambda_{n+1-p}(\mathbf{A})$. This proves the lemma. ■

17.3.2 The two families of Weyl inequalities

Here is the question. Given the eigenvalues of $\mathbf{A}$ and the eigenvalues of $\mathbf{B}$, separately, what can be said about the eigenvalues of $\mathbf{A} + \mathbf{B}$? Nothing exact: $\text{diag}(1, 0)$ and $\text{diag}(0, 1)$ have the same eigenvalue list as $\text{diag}(1, 0)$ and $\text{diag}(1, 0)$, while the two sums $\mathbf{I}$ and $\text{diag}(2, 0)$ do not. But a great deal inexactly, and the shape of the answer is dictated by the index arithmetic.

Theorem 17.3.3. Weyl's Inequalities

Let $\mathbf{A}, \mathbf{B} \in M_n(F)$ be Hermitian and let $1 \leq i, j \leq n$.

(a) If $i + j - 1 \leq n$, then

$$\lambda_{i+j-1}(\mathbf{A} + \mathbf{B}) \leq \lambda_i(\mathbf{A}) + \lambda_j(\mathbf{B}).$$

(b) If $i + j - n \geq 1$, then

$$\lambda_{i+j-n}(\mathbf{A} + \mathbf{B}) \geq \lambda_i(\mathbf{A}) + \lambda_j(\mathbf{B}).$$

Idea. For (a), read $\lambda_{i+j-1}(\mathbf{A} + \mathbf{B})$ through the **max–min** half of Theorem 17.2.3: it is the best guaranteed value of $R_{\mathbf{A}+\mathbf{B}}$ on a subspace of dimension $i+j-1$. So take an arbitrary such subspace W and find one bad vector inside it — a vector on which $R_{\mathbf{A}}$ is capped by $\lambda_i(\mathbf{A})$ and $R_{\mathbf{B}}$ is capped by $\lambda_j(\mathbf{B})$ simultaneously. The span of the bottom $n-i+1$ eigenvectors of $\mathbf{A}$ is a *subspace* on which $R_{\mathbf{A}}$ is capped by $\lambda_i(\mathbf{A})$, and the span of the bottom $n-j+1$ eigenvectors of $\mathbf{B}$ is one on which $R_{\mathbf{B}}$ is capped by $\lambda_j(\mathbf{B})$. (The set of *all* vectors capped by $\lambda_i(\mathbf{A})$ is larger and is not a subspace; the spans are what the dimension count can use.) Three subspaces must meet, and the dimensions have been chosen so that they do — twice, by Lemma 17.2.1. Since $R_{\mathbf{A}+\mathbf{B}} = R_{\mathbf{A}} + R_{\mathbf{B}}$, that one vector finishes the argument.

Part (b) is not a new proof. Negating a Hermitian matrix reverses its eigenvalue list (Lemma 17.3.2), so (a) applied to $-\mathbf{A}$ and $-\mathbf{B}$ at the reflected indices $n+1-i$ and $n+1-j$ says exactly (b), once every index is translated back.

Proof. (a) Put $k = i + j - 1$, so $1 \leq k \leq n$. By the spectral theorem choose an orthonormal basis $(\mathbf{u}_1, \dots, \mathbf{u}_n)$ of eigenvectors of $\mathbf{A}$ with $\mathbf{A}\mathbf{u}_p = \lambda_p(\mathbf{A})\mathbf{u}_p$, and an orthonormal basis $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ of eigenvectors of $\mathbf{B}$ with $\mathbf{B}\mathbf{v}_q = \lambda_q(\mathbf{B})\mathbf{v}_q$. Set

$$U = \text{span}\{\mathbf{u}_i, \dots, \mathbf{u}_n\}, \quad V = \text{span}\{\mathbf{v}_j, \dots, \mathbf{v}_n\},$$

so that $\dim U = n - i + 1$ and $\dim V = n - j + 1$, the two lists being independent as parts of a basis.

Let $W \leq F^n$ be **any** subspace with $\dim W = k$. Then

$$\dim W + \dim U = (i + j - 1) + (n - i + 1) = n + j > n,$$

since $j \geq 1$, so $W \cap U \neq \{\mathbf{0}\}$ by Lemma 17.2.1. More than that is needed, namely how large $W \cap U$ is, and the dimension formula Theorem 1.7.3 supplies it:

$$\dim(W \cap U) = \dim W + \dim U - \dim(W + U) \geq (n + j) - n = j,$$

because $W + U$ is a subspace of F^n . Therefore

$$\dim(W \cap U) + \dim V \geq j + (n - j + 1) = n + 1 > n,$$

and a second application of Lemma 17.2.1, to the subspaces $W \cap U$ and V , produces a vector

$$\mathbf{x} \in W \cap U \cap V, \quad \mathbf{x} \neq \mathbf{0}.$$

Now evaluate. Since $\mathbf{x} \in U$, part (a) of Lemma 17.3.1 gives $R_{\mathbf{A}}(\mathbf{x}) \leq \lambda_i(\mathbf{A})$; since $\mathbf{x} \in V$, the same part applied to $\mathbf{B}$ gives $R_{\mathbf{B}}(\mathbf{x}) \leq \lambda_j(\mathbf{B})$. The inner product is additive in its first slot, so $\langle (\mathbf{A} + \mathbf{B})\mathbf{x}, \mathbf{x} \rangle = \langle \mathbf{A}\mathbf{x}, \mathbf{x} \rangle + \langle \mathbf{B}\mathbf{x}, \mathbf{x} \rangle$ and hence $R_{\mathbf{A}+\mathbf{B}}(\mathbf{x}) = R_{\mathbf{A}}(\mathbf{x}) + R_{\mathbf{B}}(\mathbf{x})$. Since $\mathbf{x}$ is a nonzero vector of W ,

$$\min_{\mathbf{0} \neq \mathbf{y} \in W} R_{\mathbf{A}+\mathbf{B}}(\mathbf{y}) \leq R_{\mathbf{A}+\mathbf{B}}(\mathbf{x}) \leq \lambda_i(\mathbf{A}) + \lambda_j(\mathbf{B}).$$

The subspace W of dimension k was arbitrary, so the maximum of the left-hand side over all such W obeys the same bound; and that maximum is $\lambda_k(\mathbf{A} + \mathbf{B})$ by the max–min half of Theorem 17.2.3. This proves (a).

(b) Suppose $i + j - n \geq 1$ and set

$$i' = n + 1 - i, \quad j' = n + 1 - j,$$

both of which lie in $\{1, \dots, n\}$. Then

$$i' + j' - 1 = 2n + 1 - i - j \leq n$$

precisely because $i + j - n \geq 1$, so part (a) applies to the Hermitian matrices $-\mathbf{A}$ and $-\mathbf{B}$ at the indices i', j' :

$$\lambda_{i'+j'-1}(-\mathbf{A} - \mathbf{B}) \leq \lambda_{i'}(-\mathbf{A}) + \lambda_{j'}(-\mathbf{B}).$$

Translate each term with [Lemma 17.3.2](#). On the right, $\lambda_{i'}(-\mathbf{A}) = -\lambda_{n+1-i'}(\mathbf{A}) = -\lambda_i(\mathbf{A})$, and likewise $\lambda_{j'}(-\mathbf{B}) = -\lambda_j(\mathbf{B})$. On the left, $-\mathbf{A} - \mathbf{B} = -(\mathbf{A} + \mathbf{B})$ with $\mathbf{A} + \mathbf{B}$ Hermitian, and $i' + j' - 1 = 2n + 1 - i - j$, so

$$\lambda_{i'+j'-1}(-(\mathbf{A} + \mathbf{B})) = -\lambda_{n+1-(2n+1-i-j)}(\mathbf{A} + \mathbf{B}) = -\lambda_{i+j-n}(\mathbf{A} + \mathbf{B}).$$

The displayed inequality therefore reads $-\lambda_{i+j-n}(\mathbf{A} + \mathbf{B}) \leq -\lambda_i(\mathbf{A}) - \lambda_j(\mathbf{B})$. Multiplying by -1 reverses it and gives (b). This proves the theorem. ■

Two remarks on the bookkeeping, because it is the whole content. First, the index on the left is *not* i or j but $i + j - 1$, so a fixed eigenvalue $\lambda_k(\mathbf{A} + \mathbf{B})$ receives k upper bounds from (a), one for each way of splitting $k = i + j - 1$, and the best of them is their minimum. Second, (a) only ever bounds from above and (b) only from below. To trap one eigenvalue $\lambda_i(\mathbf{A} + \mathbf{B})$ from both sides with the index of $\mathbf{A}$ unchanged, take $j = 1$ in (a) and $j = n$ in (b); that is the next corollary.

Quick check. Take $i = j = 1$ in (a) and $i = j = n$ in (b). What do the two inequalities say, and which earlier result are they?

Solution. They say $\lambda_1(\mathbf{A} + \mathbf{B}) \leq \lambda_1(\mathbf{A}) + \lambda_1(\mathbf{B})$ and $\lambda_n(\mathbf{A} + \mathbf{B}) \geq \lambda_n(\mathbf{A}) + \lambda_n(\mathbf{B})$. For $i = j = n$ the index on the left of (b) is $n + n - n = n$, as required. These are the statements that the largest eigenvalue is subadditive and the smallest superadditive, which one can also read straight off Lemma 13.5.2: $\langle (\mathbf{A} + \mathbf{B})\mathbf{x}, \mathbf{x} \rangle = \langle \mathbf{A}\mathbf{x}, \mathbf{x} \rangle + \langle \mathbf{B}\mathbf{x}, \mathbf{x} \rangle \leq \lambda_1(\mathbf{A}) + \lambda_1(\mathbf{B})$ for every unit $\mathbf{x}$. Weyl's inequalities are what happens when this easy observation is run at every index instead of only at the two ends.

Warning. The eigenvalues are not subadditive index by index. It is tempting to read Weyl as $\lambda_i(\mathbf{A} + \mathbf{B}) \leq \lambda_i(\mathbf{A}) + \lambda_i(\mathbf{B})$, and that statement is false. Take $\mathbf{A} = \text{diag}(1, 0)$ and $\mathbf{B} = \text{diag}(0, 1)$, so that $\lambda(\mathbf{A}) = \lambda(\mathbf{B}) = (1, 0)$ and $\mathbf{A} + \mathbf{B} = \mathbf{I}_2$ with $\lambda(\mathbf{A} + \mathbf{B}) = (1, 1)$. At $i = 2$ the proposed inequality reads $1 \leq 0 + 0$. The reverse reading $\lambda_i(\mathbf{A} + \mathbf{B}) \geq \lambda_i(\mathbf{A}) + \lambda_i(\mathbf{B})$ fails on the same pair at $i = 1$, where it reads $1 \geq 2$. Only the shifted indices $i + j - 1$ and $i + j - n$ are correct.

Here is the theorem in full on one pair of matrices, so that the index arithmetic can be seen

doing its work.

Example. Both families on one pair

Let

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

Find the eigenvalues of $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{A} + \mathbf{B}$, and check every inequality of [Theorem 17.3.3](#).

Solution. All three matrices are real symmetric. Their characteristic polynomials factor over $\mathbb{Z}$:

$$p_{\mathbf{A}}(x) = x(x - 4)(x + 2),$$

$$p_{\mathbf{B}}(x) = x(x - 5)(x - 1),$$

$$p_{\mathbf{A}+\mathbf{B}}(x) = (x - 8)(x - 1)(x + 1),$$

so, in decreasing order,

$$\lambda(\mathbf{A}) = (4, 0, -2), \quad \lambda(\mathbf{B}) = (5, 1, 0), \quad \lambda(\mathbf{A} + \mathbf{B}) = (8, 1, -1).$$

For $\mathbf{A}$ one can see the list without computing: the rows $(1, 1, 2)$ and $(1, 1, 2)$ are equal, so $\text{rank } \mathbf{A} = 2$ and 0 is an eigenvalue; $\text{tr } \mathbf{A} = 2$ and $\text{tr}(\mathbf{A}^2) = 20$ then force the other two to be 4 and -2 .

With $n = 3$, family (a) applies when $i + j \leq 4$ and family (b) when $i + j \geq 4$.

(i, j)	family (a): $\lambda_{i+j-1}(\mathbf{A} + \mathbf{B}) \leq \lambda_i(\mathbf{A}) + \lambda_j(\mathbf{B})$
(1, 1)	$8 \leq 4 + 5 = 9$
(1, 2)	$1 \leq 4 + 1 = 5$
(2, 1)	$1 \leq 0 + 5 = 5$
(1, 3)	$-1 \leq 4 + 0 = 4$
(2, 2)	$-1 \leq 0 + 1 = 1$
(3, 1)	$-1 \leq -2 + 5 = 3$

(i, j)	family (b): $\lambda_{i+j-3}(\mathbf{A} + \mathbf{B}) \geq \lambda_i(\mathbf{A}) + \lambda_j(\mathbf{B})$
(1, 3)	$8 \geq 4 + 0 = 4$
(3, 1)	$8 \geq -2 + 5 = 3$
(2, 2)	$8 \geq 0 + 1 = 1$
(2, 3)	$1 \geq 0 + 0 = 0$
(3, 2)	$1 \geq -2 + 1 = -1$
(3, 3)	$-1 \geq -2 + 0 = -2$

Every inequality holds, and every one of them is strict. At (1, 1) one can see why. Equality there would mean a unit $\mathbf{x}$ with $R_{\mathbf{A}+\mathbf{B}}(\mathbf{x}) = \lambda_1(\mathbf{A}) + \lambda_1(\mathbf{B})$; since $R_{\mathbf{A}}(\mathbf{x}) \leq \lambda_1(\mathbf{A})$ and $R_{\mathbf{B}}(\mathbf{x}) \leq \lambda_1(\mathbf{B})$ by [Lemma 17.3.1](#) (a) at $i = 1$, both would have to be equalities, and the computation in that proof shows that equality in $\sum_p \lambda_p |c_p|^2 \leq \lambda_1 \sum_p |c_p|^2$ forces $c_p = 0$ at every p with $\lambda_p < \lambda_1$. So $\mathbf{x}$ would lie in the λ_1 -eigenspace of $\mathbf{A}$ and in that of $\mathbf{B}$. Those are the lines through (1, 1, 1) and through (1, 1, 0) — as $\mathbf{A}(1, 1, 1) = 4(1, 1, 1)$ and $\mathbf{B}(1, 1, 0) = 5(1, 1, 0)$, and both eigenvalues are simple — and two distinct lines meet only at $\mathbf{0}$.

17.3.3 Perturbing one matrix

Setting $j = 1$ in family (a) and $j = n$ in family (b) leaves the index on the left equal to i . That is the case in which Weyl becomes a statement about *perturbation*: one matrix, moved.

Corollary 17.3.4. Perturbation, Two-Sided

Let $\mathbf{A}, \mathbf{B} \in M_n(F)$ be Hermitian. Then for every i ,

$$\lambda_i(\mathbf{A}) + \lambda_n(\mathbf{B}) \leq \lambda_i(\mathbf{A} + \mathbf{B}) \leq \lambda_i(\mathbf{A}) + \lambda_1(\mathbf{B}).$$

Proof. For the upper bound take $j = 1$ in [Theorem 17.3.3](#) (a): the hypothesis $i+1-1 = i \leq n$ holds, and the conclusion is $\lambda_i(\mathbf{A} + \mathbf{B}) \leq \lambda_i(\mathbf{A}) + \lambda_1(\mathbf{B})$. For the lower bound take $j = n$ in [Theorem 17.3.3](#) (b): the hypothesis $i+n-n = i \geq 1$ holds, and the conclusion is $\lambda_i(\mathbf{A} + \mathbf{B}) \geq \lambda_i(\mathbf{A}) + \lambda_n(\mathbf{B})$. This proves the corollary. ■

In words: adding $\mathbf{B}$ shifts each eigenvalue of $\mathbf{A}$, in order, by an amount trapped between the smallest and the largest eigenvalue of $\mathbf{B}$. Taking $\mathbf{B} \succeq \mathbf{O}$, so that $\lambda_n(\mathbf{B}) \geq 0$ by [Theorem 13.1.4](#) (b), the lower bound becomes $\lambda_i(\mathbf{A} + \mathbf{B}) \geq \lambda_i(\mathbf{A})$ — which is [Corollary 17.2.4](#) again, now as a special case of Weyl's inequalities.

To turn [Corollary 17.3.4](#) into a bound of the kind Chapter 16 wanted, the two numbers $\lambda_1(\mathbf{B})$ and $\lambda_n(\mathbf{B})$ must be replaced by one, and the right one is a norm.

Lemma 17.3.5. The Spectral Norm of a Hermitian Matrix

Let $\mathbf{E} \in M_n(F)$ be Hermitian. Then

$$\|\mathbf{E}\|_2 = \max_{1 \leq i \leq n} |\lambda_i(\mathbf{E})| = \max(\lambda_1(\mathbf{E}), -\lambda_n(\mathbf{E})).$$

Proof. By the spectral theorem, $\mathbf{E} = \mathbf{U}\mathbf{D}\mathbf{U}^*$ with $\mathbf{U}$ unitary and $\mathbf{D} = \text{diag}(\lambda_1(\mathbf{E}), \dots, \lambda_n(\mathbf{E}))$ real. Then

$$\mathbf{E}^*\mathbf{E} = \mathbf{E}^2 = \mathbf{U}\mathbf{D}^2\mathbf{U}^*,$$

so the eigenvalues of $\mathbf{E}^*\mathbf{E}$ are the numbers $\lambda_i(\mathbf{E})^2$, and the singular values of $\mathbf{E}$, which are their nonnegative square roots listed decreasingly (Definition 13.9.1), are the numbers $|\lambda_i(\mathbf{E})|$ listed decreasingly. Hence $\sigma_1(\mathbf{E}) = \max_i |\lambda_i(\mathbf{E})|$, and $\|\mathbf{E}\|_2 = \sigma_1(\mathbf{E})$ by Theorem 16.3.3 (c).

For the second equality, $\lambda_n(\mathbf{E}) \leq \lambda_i(\mathbf{E}) \leq \lambda_1(\mathbf{E})$ for all i , so $|\lambda_i(\mathbf{E})| \leq \max(|\lambda_1(\mathbf{E})|, |\lambda_n(\mathbf{E})|)$, so the maximum is attained at one end of the list. Now $\lambda_1(\mathbf{E}) \leq |\lambda_1(\mathbf{E})|$ and $-\lambda_n(\mathbf{E}) \leq |\lambda_n(\mathbf{E})|$, which gives one inequality between the two expressions. Conversely, $|\lambda_1(\mathbf{E})|$ is either $\lambda_1(\mathbf{E})$ or $-\lambda_1(\mathbf{E}) \leq -\lambda_n(\mathbf{E})$, and $|\lambda_n(\mathbf{E})|$ is either $-\lambda_n(\mathbf{E})$ or $\lambda_n(\mathbf{E}) \leq \lambda_1(\mathbf{E})$; in every case it is at most $\max(\lambda_1(\mathbf{E}), -\lambda_n(\mathbf{E}))$. This proves the lemma. ■

Corollary 17.3.6. Hermitian Eigenvalues Are 1-Lipschitz

Let $\mathbf{A}, \mathbf{E} \in M_n(F)$ be Hermitian. Then

$$|\lambda_i(\mathbf{A} + \mathbf{E}) - \lambda_i(\mathbf{A})| \leq \|\mathbf{E}\|_2 \quad \text{for every } i = 1, \dots, n.$$

Proof. By Corollary 17.3.4 applied with $\mathbf{B} = \mathbf{E}$,

$$\lambda_n(\mathbf{E}) \leq \lambda_i(\mathbf{A} + \mathbf{E}) - \lambda_i(\mathbf{A}) \leq \lambda_1(\mathbf{E}).$$

By Lemma 17.3.5, $\lambda_1(\mathbf{E}) \leq \|\mathbf{E}\|_2$ and $-\lambda_n(\mathbf{E}) \leq \|\mathbf{E}\|_2$, that is, $-\|\mathbf{E}\|_2 \leq \lambda_n(\mathbf{E})$. So the difference lies in $[-\|\mathbf{E}\|_2, \|\mathbf{E}\|_2]$, which is the assertion. This proves the corollary. ■

This is the sentence Chapter 16 §07 promised: **the eigenvalues of a Hermitian matrix move by at most $\|\mathbf{E}\|_2$ under a Hermitian perturbation.** It is worth being precise about what it adds to Corollary 16.7.4, the continuity statement of that section, because the gain is not continuity — that was already had — but a constant.

- Corollary 16.7.4 says: for each $\mathbf{A}$ and each $\varepsilon > 0$ there is a $\delta > 0$ such that every matrix within δ of $\mathbf{A}$ has eigenvalues that can be numbered to lie within ε of those of $\mathbf{A}$. The δ depends on $\mathbf{A}$, and Example (A Jordan block's eigenvalues move like $\varepsilon^{1/k}$) shows that it must.
- Corollary 17.3.6 says, for Hermitian $\mathbf{A}$ and Hermitian perturbations measured in $\|\cdot\|_2$: $\delta = \varepsilon$ works, for every $\mathbf{A}$ at once, and the numbering is simply the decreasing order. The movement is bounded by the perturbation itself, with **Lipschitz constant 1**.

The constant does not depend on the size n of the matrices, it does not depend on how ill-conditioned $\mathbf{A}$ is, and it does not depend on how close together the eigenvalues of $\mathbf{A}$ are —

three quantities that every reader of Chapter 16 has learned to expect in an error bound. It also cannot be improved: $\mathbf{E} = t\mathbf{I}$ with $t \in \mathbb{R}$ gives $\lambda_i(\mathbf{A} + t\mathbf{I}) = \lambda_i(\mathbf{A}) + t$ for every i , while $\|t\mathbf{I}\|_2 = |t|$, so the inequality is an equality at every index. Finally, the eigenvalues are matched **by index**: the i -th of $\mathbf{A} + \mathbf{E}$ is compared with the i -th of $\mathbf{A}$, not with whichever eigenvalue of $\mathbf{A}$ happens to be nearest, and no search for a good numbering is needed.

Corollary 17.3.7. The Hermitian Eigenvalue Problem Is Perfectly Conditioned in Absolute Terms

Let $\mathbf{A}, \mathbf{E} \in M_n(F)$ be Hermitian and write $\lambda(\mathbf{A}) = (\lambda_1(\mathbf{A}), \dots, \lambda_n(\mathbf{A}))$ for the decreasing eigenvalue list. Then

$$\|\lambda(\mathbf{A} + \mathbf{E}) - \lambda(\mathbf{A})\|_\infty \leq \|\mathbf{E}\|_2,$$

and the constant 1 is the smallest one for which this holds.

Proof. The ∞ -norm of a vector is the largest modulus of its entries, so the left-hand side is $\max_i |\lambda_i(\mathbf{A} + \mathbf{E}) - \lambda_i(\mathbf{A})|$, and the inequality is [Corollary 17.3.6](#) read over all i at once. For the last clause, $\mathbf{E} = t\mathbf{I}$ with $t > 0$ makes both sides equal to t , so no constant below 1 can work. This proves the corollary. ■

Chapter 16 §08 measured a linear system by $\kappa(\mathbf{A}) = \|\mathbf{A}\| \|\mathbf{A}^{-1}\|$ (Definition 16.8.1), the worst factor by which $\mathbf{A}$ magnifies a relative error in the data. The corollary above is the corresponding statement for a different problem — not “solve $\mathbf{A}\mathbf{x} = \mathbf{b}$ ” but “report the eigenvalues of $\mathbf{A}$ ” — and the factor there is 1, in absolute rather than relative terms. The difference between absolute and relative matters here, and it is worth being exact. $\mathbf{A} = \text{diag}(10^6, 10^{-6})$ is Hermitian with $\kappa_2(\mathbf{A}) = 10^{12}$, so solving with it is a delicate business, and yet its eigenvalues move by at most $\|\mathbf{E}\|_2$ under any Hermitian $\mathbf{E}$, exactly as those of every other Hermitian matrix do. **Ill-conditioning of the linear system says nothing about the *absolute* conditioning of the eigenvalues.** In *relative* terms the two numbers meet again. The Hermitian perturbation $\mathbf{E} = -10^{-6} \mathbf{e}_2 \mathbf{e}_2^\top$ has relative size $\|\mathbf{E}\|_2 / \|\mathbf{A}\|_2 = 10^{-12}$ and sends the eigenvalue 10^{-6} to 0, a relative change of 100%: an amplification of exactly $10^{12} = \kappa_2(\mathbf{A})$. In general, for an invertible Hermitian $\mathbf{A}$, [Corollary 17.3.6](#) bounds the relative change of the eigenvalue of smallest modulus by $\|\mathbf{E}\|_2 / \sigma_n(\mathbf{A})$, which is $\kappa_2(\mathbf{A})$ times the relative size of $\mathbf{E}$, and a rank-one $\mathbf{E}$ along its eigenvector attains it. So for Hermitian matrices, κ_2 is the relative condition number of the smallest eigenvalue.

The contrast that makes the corollary worth having is with Example (A Jordan block’s eigenvalues move like $\varepsilon^{1/k}$). There $\mathbf{A} = \mathbf{J}_k(0)$ and $\mathbf{E} = \varepsilon \mathbf{E}_{k1}$, and the eigenvalues move a distance $\varepsilon^{1/k}$ while $\|\mathbf{E}\|_2 = \varepsilon$, an amplification of $\varepsilon^{-(k-1)/k}$ that is unbounded as $\varepsilon \rightarrow 0^+$. Neither matrix there is Hermitian, and [Corollary 17.3.6](#) explains which hypothesis is doing the work: for $k = 2$ the perturbed matrix $\begin{pmatrix} 0 & 1 \\ \varepsilon & 0 \end{pmatrix}$ has the perfectly real eigenvalues $\pm\sqrt{\varepsilon}$, so the conclusion of [Corollary 17.3.6](#) can be *stated* for it, and it is false — $\sqrt{\varepsilon} > \varepsilon$ for $0 < \varepsilon < 1$, by a factor of 1000 at $\varepsilon = 10^{-6}$.

Warning. Both matrices must be Hermitian, and “small” is no substitute. If $\mathbf{A}$ is not Hermitian, the corollary fails even when $\mathbf{E}$ is: take $\mathbf{A} = \mathbf{J}_2(0)$ and the Hermitian $\mathbf{E} = \varepsilon \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, with $\|\mathbf{E}\|_2 = \varepsilon$. Then $\mathbf{A} + \mathbf{E} = \begin{pmatrix} 0 & 1+\varepsilon \\ \varepsilon & 0 \end{pmatrix}$ has the real eigenvalues $\pm\sqrt{\varepsilon(1+\varepsilon)}$, which have moved from 0 by more than ε . If $\mathbf{E}$ is not Hermitian, the conclusion usually cannot even be stated: $\mathbf{A} = \mathbf{O}$ is Hermitian and $\mathbf{E} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ has $\|\mathbf{E}\|_2 = 1$, but $\mathbf{A} + \mathbf{E}$ has

eigenvalues $\pm i$, so there is no decreasing list $\lambda_i(\mathbf{A} + \mathbf{E})$ of real numbers to compare with anything. A bound for a diagonalizable $\mathbf{A}$ and an arbitrary $\mathbf{E}$ does exist — the theorem of Bauer and Fike, in Chapter 20 — and it carries an extra factor that measures how far the eigenvector basis of $\mathbf{A}$ is from orthonormal.

Example. Locating eigenvalues without recomputing them

Let

$$\mathbf{A} = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix}, \quad \mathbf{E} = \frac{1}{10}(\mathbf{E}_{13} + \mathbf{E}_{31}).$$

Use [Corollary 17.3.6](#) to locate the eigenvalues of $\mathbf{A} + \mathbf{E}$, then compute them exactly and compare.

Solution. The unperturbed spectrum. Expanding along the first row, $p_{\mathbf{A}}(x) = (x - 2)((x - 2)^2 - 2)$, so

$$\lambda(\mathbf{A}) = (2 + \sqrt{2}, 2, 2 - \sqrt{2}) \approx (3.41421, 2, 0.58579).$$

The size of the perturbation. $\mathbf{E}$ is real symmetric with $\mathbf{E}^2 = \frac{1}{100}(\mathbf{E}_{11} + \mathbf{E}_{33})$, whose eigenvalues are $\frac{1}{100}, \frac{1}{100}, 0$; so the singular values of $\mathbf{E}$ are $\frac{1}{10}, \frac{1}{10}, 0$ and $\|\mathbf{E}\|_2 = \frac{1}{10}$ by Theorem 16.3.3 (c). (One can also read this off [Lemma 17.3.5](#): the eigenvalues of $\mathbf{E}$ are $\frac{1}{10}, 0, -\frac{1}{10}$.)

The prediction. By [Corollary 17.3.6](#) each eigenvalue of $\mathbf{A} + \mathbf{E}$ lies within $\frac{1}{10}$ of the corresponding eigenvalue of $\mathbf{A}$:

$$\lambda_1 \in [3.31421, 3.51421], \quad \lambda_2 \in [1.9, 2.1], \quad \lambda_3 \in [0.48579, 0.68579].$$

In particular $\mathbf{A} + \mathbf{E}$ is positive definite, since $\lambda_3 > 0$, and its three eigenvalues are distinct, since the three intervals are disjoint.

The check. The characteristic polynomial of $\mathbf{A} + \mathbf{E}$ factors as

$$p_{\mathbf{A}+\mathbf{E}}(x) = \left(x - \frac{19}{10}\right)\left(x^2 - \frac{41}{10}x + \frac{11}{5}\right),$$

so

$$\begin{aligned} \lambda(\mathbf{A} + \mathbf{E}) &= \left(\frac{41+3\sqrt{89}}{20}, \frac{19}{10}, \frac{41-3\sqrt{89}}{20}\right) \\ &\approx (3.46510, 1.9, 0.63490). \end{aligned}$$

All three lie in the predicted intervals. The middle one lies at the very end of its interval: λ_2 has moved by exactly $\frac{1}{10} = \|\mathbf{E}\|_2$. That is not luck. The vector $(1, 0, -1)$ is an eigenvector of $\mathbf{A}$ for 2 and of $\mathbf{E}$ for $-\frac{1}{10}$, hence an eigenvector of $\mathbf{A} + \mathbf{E}$ for $2 - \frac{1}{10}$; a shared eigenvector is exactly what makes the bound tight.

17.3.4 Perturbations of small rank

The last corollary trades size for rank. A perturbation of rank r need not be small in norm at all, but it can only disturb r directions, and Weyl converts that into a shift of the index by r .

Corollary 17.3.8. A Rank- r Perturbation Shifts the Index by r

Let $\mathbf{A}, \mathbf{B} \in M_n(F)$ be Hermitian, let $r \geq 0$ be an integer, and suppose $\text{rank } \mathbf{B} \leq r$. Then

$$\begin{aligned}\lambda_{i+r}(\mathbf{A} + \mathbf{B}) &\leq \lambda_i(\mathbf{A}) & (1 \leq i, i+r \leq n), \\ \lambda_i(\mathbf{A}) &\leq \lambda_{i-r}(\mathbf{A} + \mathbf{B}) & (i \leq n, i-r \geq 1).\end{aligned}$$

Idea. A Hermitian matrix of rank at most r has at most r nonzero eigenvalues, so once one walks $r+1$ places down its decreasing list one is certainly at a nonpositive number, and once one stops r places short of the bottom one is certainly at a nonnegative number. Feed those two facts into the two Weyl families at $j = r+1$ and $j = n-r$.

Proof. By the spectral theorem $\mathbf{B} = \mathbf{U}\mathbf{D}\mathbf{U}^*$ with $\mathbf{D} = \text{diag}(\lambda_1(\mathbf{B}), \dots, \lambda_n(\mathbf{B}))$ and $\mathbf{U}$ unitary, so $\text{rank } \mathbf{B} = \text{rank } \mathbf{D}$ is the number of nonzero $\lambda_q(\mathbf{B})$, which is at most r by hypothesis.

Two facts about the list of $\mathbf{B}$. If $\lambda_{r+1}(\mathbf{B}) > 0$, then $\lambda_1(\mathbf{B}), \dots, \lambda_{r+1}(\mathbf{B})$ are all positive by the decreasing indexing, giving $r+1$ nonzero eigenvalues, a contradiction; hence

$$\lambda_{r+1}(\mathbf{B}) \leq 0 \quad \text{whenever } r+1 \leq n.$$

If $\lambda_{n-r}(\mathbf{B}) < 0$, then $\lambda_{n-r}(\mathbf{B}), \dots, \lambda_n(\mathbf{B})$ are all negative, again $r+1$ nonzero eigenvalues and a contradiction; hence

$$\lambda_{n-r}(\mathbf{B}) \geq 0 \quad \text{whenever } n-r \geq 1.$$

First inequality. Assume $i \geq 1$ and $i+r \leq n$; then $r+1 \leq n$, so the index $j = r+1$ is legitimate and $i+j-1 = i+r \leq n$. By [Theorem 17.3.3](#) (a) and the first fact,

$$\lambda_{i+r}(\mathbf{A} + \mathbf{B}) \leq \lambda_i(\mathbf{A}) + \lambda_{r+1}(\mathbf{B}) \leq \lambda_i(\mathbf{A}).$$

Second inequality. Assume $i \leq n$ and $i-r \geq 1$; then $n-r \geq n-i+1 \geq 1$, so the index $j = n-r$ is legitimate, and $i+j-n = i-r \geq 1$. By [Theorem 17.3.3](#) (b) and the second fact,

$$\lambda_{i-r}(\mathbf{A} + \mathbf{B}) \geq \lambda_i(\mathbf{A}) + \lambda_{n-r}(\mathbf{B}) \geq \lambda_i(\mathbf{A}).$$

This proves the corollary. ■

Read together, the two inequalities say that the two eigenvalue lists interlace with a lag of r :

$$\lambda_{i+r}(\mathbf{A} + \mathbf{B}) \leq \lambda_i(\mathbf{A}) \leq \lambda_{i-r}(\mathbf{A} + \mathbf{B}),$$

whenever all three indices make sense. For $r = 0$ this collapses to $\lambda_i(\mathbf{A} + \mathbf{B}) = \lambda_i(\mathbf{A})$, correctly, since a Hermitian matrix of rank 0 is $\mathbf{O}$. For $r = 1$ it is the statement that adding a Hermitian rank-one matrix moves each eigenvalue by at most one place in the list. That case is the subject of Section 5, and Section 9 returns to low-rank perturbations when it finishes the approximation theorem that Chapter 13 §11 could only state in the Frobenius norm.

Remark. The bound is about the *index*, not the *value*. A rank-one $\mathbf{B}$ can be enormous: $\mathbf{B} = 10^6 \mathbf{e}_1 \mathbf{e}_1^*$ has rank 1, and $\lambda_1(\mathbf{A} + \mathbf{B})$ lies within $\|\mathbf{A}\|_2$ of 10^6 , by [Corollary 17.3.6](#) with the roles of the two matrices exchanged. What [Corollary 17.3.8](#) guarantees is that the others are trapped all the same: $\lambda_{k+1}(\mathbf{A}) \leq \lambda_k(\mathbf{A} + \mathbf{B}) \leq \lambda_{k-1}(\mathbf{A})$ for $2 \leq k \leq n-1$, however large the one that escapes.

17.3.5 Exercises

17.3.5.1 A. Check your understanding

Exercise. A1

- State both families of [Theorem 17.3.3](#), with their index conditions, and say why the condition $i + j - 1 \leq n$ cannot be dropped from the first.
- Determine whether the following is correct, and justify your answer: for Hermitian $\mathbf{A}, \mathbf{B}$, $\lambda_2(\mathbf{A} + \mathbf{B}) \leq \lambda_2(\mathbf{A}) + \lambda_2(\mathbf{B})$.
- [Corollary 16.7.4](#) already said that eigenvalues move continuously. What does [Corollary 17.3.6](#) add?
- Explain why [Corollary 17.3.6](#) cannot be applied to $\mathbf{A} = \mathbf{J}_2(0)$ and $\mathbf{E} = \varepsilon \mathbf{E}_{21}$, and what goes wrong if one applies it anyway.
- If $\mathbf{B}$ is Hermitian of rank 3 and $\mathbf{A}$ is Hermitian of size 10, which eigenvalue of $\mathbf{A} + \mathbf{B}$ is guaranteed to be at most $\lambda_4(\mathbf{A})$?

Solution.

- See [Theorem 17.3.3](#). Family (a) needs $i + j - 1 \leq n$ and family (b) needs $i + j - n \geq 1$. Without the first condition the symbol $\lambda_{i+j-1}(\mathbf{A} + \mathbf{B})$ names nothing: a matrix of size n has only n eigenvalues, and there is no λ_{n+1} .
- Incorrect. With $\mathbf{A} = \text{diag}(1, 0)$ and $\mathbf{B} = \text{diag}(0, 1)$ it reads $\lambda_2(\mathbf{I}_2) = 1 \leq 0 + 0$. The correct instance of [Theorem 17.3.3](#) (a) at $i = j = 2$ needs $i + j - 1 = 3 \leq n$, so it says nothing when $n = 2$; when $n \geq 3$ it bounds λ_3 , not λ_2 .
- A constant. [Corollary 16.7.4](#) gives, for each $\mathbf{A}$ and ε , some unspecified δ depending on $\mathbf{A}$; [Corollary 17.3.6](#) gives $\delta = \varepsilon$ for every $\mathbf{A}$ simultaneously, that is, a Lipschitz constant of 1, independent of n , of the conditioning of $\mathbf{A}$, and of the gaps between its eigenvalues. It also matches the eigenvalues by index rather than as a multiset.
- Neither matrix is Hermitian, and the corollary's hypothesis is that both are. Applying it anyway predicts movement at most $\|\mathbf{E}\|_2 = \varepsilon$, whereas the eigenvalues are $\pm\sqrt{\varepsilon}$ and have moved a distance $\sqrt{\varepsilon}$, which exceeds ε for every $\varepsilon \in (0, 1)$ — see Example (A Jordan block's eigenvalues move like $\varepsilon^{1/k}$).
- $\lambda_7(\mathbf{A} + \mathbf{B})$, and hence every later one, by [Corollary 17.3.8](#) with $i = 4$ and $r = 3$.

17.3.5.2 B. Practice

Exercise. B1: A small table

Let $\mathbf{A} = \begin{pmatrix} 0 & 4 \\ 4 & 0 \end{pmatrix}$ and $\mathbf{B} = \begin{pmatrix} 3 & 0 \\ 0 & -3 \end{pmatrix}$. Compute the eigenvalues of $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{A} + \mathbf{B}$, write out every inequality of [Theorem 17.3.3](#) for this pair, and verify [Corollary 17.3.6](#) with $\mathbf{E} = \mathbf{B}$.

Solution. $p_{\mathbf{A}}(x) = x^2 - 16$, so $\lambda(\mathbf{A}) = (4, -4)$; $\mathbf{B}$ is diagonal, so $\lambda(\mathbf{B}) = (3, -3)$; and $\mathbf{A} + \mathbf{B} = \begin{pmatrix} 3 & 4 \\ 4 & -3 \end{pmatrix}$ has $\text{tr} = 0$ and $\det = -25$, so $p_{\mathbf{A}+\mathbf{B}}(x) = x^2 - 25$ and $\lambda(\mathbf{A} + \mathbf{B}) = (5, -5)$.

Family (a), for $i + j \leq 3$:

$$\begin{aligned} (1, 1) : 5 &\leq 4 + 3 = 7, \\ (1, 2) : -5 &\leq 4 - 3 = 1, \\ (2, 1) : -5 &\leq -4 + 3 = -1. \end{aligned}$$

Family (b), for $i + j \geq 3$:

$$\begin{aligned} (1, 2) : 5 &\geq 4 - 3 = 1, \\ (2, 1) : 5 &\geq -4 + 3 = -1, \\ (2, 2) : -5 &\geq -4 - 3 = -7. \end{aligned}$$

All six hold. For the last part, $\mathbf{B}$ is Hermitian with eigenvalues ± 3 , so $\|\mathbf{B}\|_2 = 3$ by [Lemma 17.3.5](#), and

$$\begin{aligned} |\lambda_1(\mathbf{A} + \mathbf{B}) - \lambda_1(\mathbf{A})| &= |5 - 4| = 1 \leq 3, \\ |\lambda_2(\mathbf{A} + \mathbf{B}) - \lambda_2(\mathbf{A})| &= |-5 + 4| = 1 \leq 3, \end{aligned}$$

as [Corollary 17.3.6](#) requires.

Exercise. B2: An interval for each eigenvalue

Let $\mathbf{A} = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix}$, whose eigenvalues are $2 \pm \sqrt{2}$ and 2, and let $\mathbf{E} = \frac{1}{100}\mathbf{J}$, where $\mathbf{J} \in M_3(\mathbb{R})$ is the all-ones matrix.

- Compute $\|\mathbf{E}\|_2$ and give an interval containing each eigenvalue of $\mathbf{A} + \mathbf{E}$.
- Deduce that $\mathbf{A} + \mathbf{E}$ is positive definite.

Solution.

- $\mathbf{J}$ is real symmetric with $\mathbf{J}^2 = 3\mathbf{J}$, so its eigenvalues satisfy $\lambda^2 = 3\lambda$ and are 3 and 0; since $\text{rank } \mathbf{J} = 1$ the list is $(3, 0, 0)$. Hence $\mathbf{E}$ has eigenvalues $(\frac{3}{100}, 0, 0)$ and $\|\mathbf{E}\|_2 = \frac{3}{100}$ by [Lemma 17.3.5](#). By [Corollary 17.3.6](#),

$$\begin{aligned} \lambda_1(\mathbf{A} + \mathbf{E}) &\in [3.38421, 3.44421], \\ \lambda_2(\mathbf{A} + \mathbf{E}) &\in [1.97, 2.03], \\ \lambda_3(\mathbf{A} + \mathbf{E}) &\in [0.55579, 0.61579], \end{aligned}$$

the centers being $2 + \sqrt{2} \approx 3.41421$, 2, and $2 - \sqrt{2} \approx 0.58579$.

- Every interval lies strictly to the right of 0, so all three eigenvalues of the Hermitian matrix $\mathbf{A} + \mathbf{E}$ are positive and $\mathbf{A} + \mathbf{E} \succ \mathbf{O}$ by [Theorem 13.1.5 \(b\)](#). No characteristic polynomial was computed. (For the record, the exact eigenvalues are 2 and $\frac{403 \pm \sqrt{81609}}{200}$,

that is 2, 3.44336..., and 0.58664...; the middle one is unmoved because $(1, 0, -1)$ is an eigenvector of both $\mathbf{A}$ and $\mathbf{J}$.)

Exercise. B3: What a low-rank update can do

Let $\mathbf{A} \in M_5(\mathbb{C})$ be Hermitian with $\lambda(\mathbf{A}) = (5, 4, 3, 2, 1)$, and let $\mathbf{B} \in M_5(\mathbb{C})$ be Hermitian with $\text{rank } \mathbf{B} \leq 2$. Find the best bounds that [Corollary 17.3.8](#) gives for $\lambda_3(\mathbf{A} + \mathbf{B})$, and give examples showing that each is attained.

Solution. Take $r = 2$. The first inequality reads $\lambda_{i+2}(\mathbf{A} + \mathbf{B}) \leq \lambda_i(\mathbf{A})$, and it mentions $\lambda_3(\mathbf{A} + \mathbf{B})$ only at $i = 1$, giving $\lambda_3(\mathbf{A} + \mathbf{B}) \leq \lambda_1(\mathbf{A}) = 5$. The second reads $\lambda_i(\mathbf{A}) \leq \lambda_{i-2}(\mathbf{A} + \mathbf{B})$, and it mentions $\lambda_3(\mathbf{A} + \mathbf{B})$ only at $i = 5$, giving $\lambda_3(\mathbf{A} + \mathbf{B}) \geq \lambda_5(\mathbf{A}) = 1$. So the corollary gives exactly $\lambda_3(\mathbf{A} + \mathbf{B}) \in [1, 5]$.

Both ends are attained. Write $\mathbf{A} = \mathbf{U} \text{diag}(5, 4, 3, 2, 1) \mathbf{U}^*$ with $\mathbf{U}$ unitary (Corollary 12.4.3) and let $\mathbf{q}_1, \dots, \mathbf{q}_5$ be its columns, an orthonormal eigenbasis.

For the upper bound take $\mathbf{B} = 4(\mathbf{q}_4 \mathbf{q}_4^* + \mathbf{q}_5 \mathbf{q}_5^*)$, which is Hermitian of rank 2. It adds 4 to the fourth and fifth eigenvalues and leaves the rest alone, so $\mathbf{A} + \mathbf{B} = \mathbf{U} \text{diag}(5, 4, 3, 6, 5) \mathbf{U}^*$, whose eigenvalue list sorted decreasingly is $(6, 5, 5, 4, 3)$. Hence $\lambda_3(\mathbf{A} + \mathbf{B}) = 5 = \lambda_1(\mathbf{A})$.

For the lower bound take $\mathbf{B} = -4(\mathbf{q}_1 \mathbf{q}_1^* + \mathbf{q}_2 \mathbf{q}_2^*)$, again Hermitian of rank 2. Then $\mathbf{A} + \mathbf{B} = \mathbf{U} \text{diag}(1, 0, 3, 2, 1) \mathbf{U}^*$, sorted $(3, 2, 1, 1, 0)$, so $\lambda_3(\mathbf{A} + \mathbf{B}) = 1 = \lambda_5(\mathbf{A})$.

17.3.5.3 C. Going deeper

Exercise. C1: The Loewner order through Weyl

Let $\mathbf{A}, \mathbf{B} \in M_n(F)$ be Hermitian.

- Prove that if $\mathbf{A} \succeq \mathbf{B}$ then $\lambda_i(\mathbf{A}) \geq \lambda_i(\mathbf{B})$ for every i , as a consequence of [Corollary 17.3.4](#).
- Hence deduce that $|\lambda_i(\mathbf{A}) - \lambda_i(\mathbf{B})| \leq \|\mathbf{A} - \mathbf{B}\|_2$ for every i , without invoking [Corollary 17.3.6](#).

Solution.

- Put $\mathbf{C} = \mathbf{A} - \mathbf{B}$, which is Hermitian, and $\mathbf{C} \succeq \mathbf{O}$ by hypothesis and Definition 13.5.1. By Theorem 13.1.4 (b) every eigenvalue of $\mathbf{C}$ is ≥ 0 , in particular $\lambda_n(\mathbf{C}) \geq 0$. Applying [Corollary 17.3.4](#) to $\mathbf{B}$ and $\mathbf{C}$,

$$\lambda_i(\mathbf{A}) = \lambda_i(\mathbf{B} + \mathbf{C}) \geq \lambda_i(\mathbf{B}) + \lambda_n(\mathbf{C}) \geq \lambda_i(\mathbf{B}).$$

This is Corollary 17.2.4.

- Put $\mathbf{C} = \mathbf{A} - \mathbf{B}$ again, now with no sign hypothesis, and let $t = \|\mathbf{C}\|_2$. By [Lemma 17.3.5](#), $-t \leq \lambda_n(\mathbf{C})$ and $\lambda_1(\mathbf{C}) \leq t$, so $t\mathbf{I} - \mathbf{C} \succeq \mathbf{O}$ and $\mathbf{C} + t\mathbf{I} \succeq \mathbf{O}$ by Theorem 13.1.4 (b) applied to the Hermitian matrices $t\mathbf{I} - \mathbf{C}$ and $\mathbf{C} + t\mathbf{I}$, whose eigenvalues are $t - \lambda_q(\mathbf{C})$ and $\lambda_q(\mathbf{C}) + t$. Hence $\mathbf{B} + t\mathbf{I} \succeq \mathbf{A} \succeq \mathbf{B} - t\mathbf{I}$, and part (a)

gives

$$\begin{aligned}\lambda_i(\mathbf{B}) + t &= \lambda_i(\mathbf{B} + t\mathbf{I}) \geq \lambda_i(\mathbf{A}), \\ \lambda_i(\mathbf{A}) &\geq \lambda_i(\mathbf{B} - t\mathbf{I}) = \lambda_i(\mathbf{B}) - t,\end{aligned}$$

using $\lambda_i(\mathbf{M} + s\mathbf{I}) = \lambda_i(\mathbf{M}) + s$ for real s , which holds because $\mathbf{M}$ and $\mathbf{M} + s\mathbf{I}$ have the same eigenvectors with eigenvalues shifted by s , and shifting by a constant preserves the decreasing order. Rearranged, this is $|\lambda_i(\mathbf{A}) - \lambda_i(\mathbf{B})| \leq t$, as claimed.

Exercise. C2: The largest eigenvalue is a convex function

Let $\mathbf{A}, \mathbf{B} \in M_n(F)$ be Hermitian and let $t \in [0, 1]$.

- (a) Prove that $\lambda_1(t\mathbf{A} + (1-t)\mathbf{B}) \leq t\lambda_1(\mathbf{A}) + (1-t)\lambda_1(\mathbf{B})$.
- (b) Deduce that for every $c \in \mathbb{R}$ the set $\{\mathbf{M} \in M_n(F) : \mathbf{M}^* = \mathbf{M}, \lambda_1(\mathbf{M}) \leq c\}$ is convex.
- (c) Show that λ_2 is **not** convex in this sense, by producing two Hermitian 2×2 matrices and a value of t where the inequality fails.

Solution.

- (a) For $s \geq 0$ and Hermitian $\mathbf{M}$ we have $\lambda_1(s\mathbf{M}) = s\lambda_1(\mathbf{M})$, since $s\mathbf{M} = \mathbf{U}(s\mathbf{D})\mathbf{U}^*$ and multiplying a decreasing list of reals by $s \geq 0$ keeps it decreasing. By [Theorem 17.3.3](#) (a) at $i = j = 1$, applied to the Hermitian matrices $t\mathbf{A}$ and $(1-t)\mathbf{B}$,

$$\lambda_1(t\mathbf{A} + (1-t)\mathbf{B}) \leq \lambda_1(t\mathbf{A}) + \lambda_1((1-t)\mathbf{B}) = t\lambda_1(\mathbf{A}) + (1-t)\lambda_1(\mathbf{B}).$$

- (b) Let $\mathbf{A}, \mathbf{B}$ lie in the set and let $t \in [0, 1]$. The matrix $t\mathbf{A} + (1-t)\mathbf{B}$ is Hermitian, and by (a) its largest eigenvalue is at most $tc + (1-t)c = c$. So the set contains the whole segment between any two of its points.
- (c) Take $\mathbf{A} = \text{diag}(1, 0)$, $\mathbf{B} = \text{diag}(0, 1)$, and $t = \frac{1}{2}$. Then $\lambda_2(\mathbf{A}) = \lambda_2(\mathbf{B}) = 0$, while $t\mathbf{A} + (1-t)\mathbf{B} = \frac{1}{2}\mathbf{I}_2$ has $\lambda_2 = \frac{1}{2}$. So $\frac{1}{2} \leq \frac{1}{2} \cdot 0 + \frac{1}{2} \cdot 0 = 0$ is false. The reason is visible in [Theorem 17.3.3](#): at $i = j = 2$ the index on the left is 3, not 2.

Exercise. C3: Sharpness of the whole family

Fix n and indices i, j with $i + j - 1 \leq n$. Construct Hermitian $\mathbf{A}, \mathbf{B} \in M_n(\mathbb{R})$, both diagonal, for which

$$\lambda_{i+j-1}(\mathbf{A} + \mathbf{B}) = \lambda_i(\mathbf{A}) + \lambda_j(\mathbf{B}).$$

Hint: with diagonal matrices every quantity in sight is a sorting problem; arrange for the $(i+j-1)$ -st largest entry of the sum to sit in a coordinate where $\mathbf{A}$ contributes its i -th largest and $\mathbf{B}$ its j -th largest.

Solution. Let $k = i + j - 1$ and let $\mathbf{A} = \text{diag}(a_1, \dots, a_n)$ and $\mathbf{B} = \text{diag}(b_1, \dots, b_n)$ with

$$a_p = \begin{cases} 1 & p \leq i-1, \\ 0 & p \geq i, \end{cases} \quad b_p = \begin{cases} 0 & p \leq i-1, \\ 1 & i \leq p \leq k, \\ 0 & p > k. \end{cases}$$

Both are diagonal with entries 0 and 1, hence Hermitian, and their diagonals are already their eigenvalue lists up to sorting. The matrix $\mathbf{A}$ has $i - 1$ entries equal to 1, so $\lambda_i(\mathbf{A}) = 0$. The matrix $\mathbf{B}$ has $k - i + 1 = j$ entries equal to 1, so $\lambda_j(\mathbf{B}) = 1$. The sum is $\mathbf{A} + \mathbf{B} = \text{diag}(1, \dots, 1, 0, \dots, 0)$ with exactly k ones, since the ones of $\mathbf{A}$ occupy coordinates $1, \dots, i - 1$ and those of $\mathbf{B}$ occupy $i, \dots, k$, and these two sets of coordinates do not overlap. Hence $\lambda_k(\mathbf{A} + \mathbf{B}) = 1 = \lambda_i(\mathbf{A}) + \lambda_j(\mathbf{B})$.

Every inequality of the first family is therefore attained, for every admissible pair (i, j) , already by 0–1 diagonal matrices. Applying [Lemma 17.3.2](#) to $-\mathbf{A}$ and $-\mathbf{B}$, as in the proof of [Theorem 17.3.3](#) (b), turns this family of equalities into one showing that the second family is attained as well.