

17.4 Deleting a Row and a Column

Take a Hermitian matrix, cross out the last row and the last column, and look at what is left. It is again Hermitian, one size smaller, and it has its own eigenvalues. Two lists of numbers are now in play, n of them and $n - 1$ of them, and they cannot be independent of each other: the smaller matrix is the larger one restricted to a hyperplane, and Theorem 17.2.3 describes every eigenvalue by what happens on subspaces. Restricting the supply of subspaces can only make the description worse in one direction and better in the other, and the exact bookkeeping is this section's subject.

The answer is as tight as it could be. The $n - 1$ eigenvalues of the corner fall into the $n - 1$ gaps of the eigenvalues of the whole, one to a gap. Nothing stronger holds: every interlacing pattern actually occurs (Exercise C2 proves the case $n = 2$).

Throughout this section the field is $\mathbb{R}$ or $\mathbb{C}$, and every matrix whose eigenvalues are written down is Hermitian, with eigenvalues real and indexed decreasingly. For an index set $I \subseteq \{1, \dots, n\}$ we write $\mathbf{A}_{I,I}$ for the **principal submatrix** of $\mathbf{A}$ on the rows and columns indexed by I , as in Definition 7.7.1. A principal submatrix of a Hermitian matrix is Hermitian, since $(\mathbf{A}_{I,I})^* = (\mathbf{A}^*)_{I,I} = \mathbf{A}_{I,I}$.

17.4.1 A matrix restricted to a coordinate hyperplane

Deleting row m and column m is not an arbitrary mutilation of $\mathbf{A}$. It is what $\mathbf{A}$ looks like to vectors whose m -th coordinate is zero, and that is the only observation the proof needs.

Lemma 17.4.1. Deleting a Row and a Column Is Restricting to a Hyperplane

Let $\mathbf{A} \in M_n(F)$ be Hermitian with $n \geq 2$, fix m with $1 \leq m \leq n$, put $I = \{1, \dots, n\} \setminus \{m\}$ and $\mathbf{B} = \mathbf{A}_{I,I} \in M_{n-1}(F)$. Let

$$H = \{\mathbf{x} \in F^n : x_m = 0\},$$

and let $\varphi: H \rightarrow F^{n-1}$ delete the m -th coordinate. Then φ is a linear isomorphism, and for every $\mathbf{x} \in H$,

$$\langle \mathbf{A}\mathbf{x}, \mathbf{x} \rangle = \langle \mathbf{B}\varphi(\mathbf{x}), \varphi(\mathbf{x}) \rangle, \quad \|\mathbf{x}\| = \|\varphi(\mathbf{x})\|.$$

In particular $R_{\mathbf{A}}(\mathbf{x}) = R_{\mathbf{B}}(\varphi(\mathbf{x}))$ for every $\mathbf{x} \in H$ with $\mathbf{x} \neq \mathbf{0}$.

Proof. The map φ is linear, being given by deleting a coordinate, and it is bijective from H to F^{n-1} because inserting 0 in position m inverts it. Write $I = \{i_1 < \dots < i_{n-1}\}$, so that $\varphi(\mathbf{x})_r = x_{i_r}$ and $(\mathbf{B})_{rs} = a_{i_r i_s}$ by Definition 7.7.1.

Let $\mathbf{x} \in H$. Expanding the inner product of F^n , which is linear in the first slot and conjugate-linear in the second,

$$\langle \mathbf{A}\mathbf{x}, \mathbf{x} \rangle = \sum_{p=1}^n \sum_{q=1}^n a_{pq} x_q \overline{x_p}.$$

Every term with $p = m$ or $q = m$ vanishes, because $x_m = 0$. The remaining terms are indexed by $p, q \in I$, say $p = i_r$ and $q = i_s$, and there the summand is $a_{i_r i_s} \varphi(\mathbf{x})_s \overline{\varphi(\mathbf{x})_r} = (\mathbf{B})_{rs} \varphi(\mathbf{x})_s \overline{\varphi(\mathbf{x})_r}$. Summing over r, s gives $\langle \mathbf{B}\varphi(\mathbf{x}), \varphi(\mathbf{x}) \rangle$. The same deletion of a zero coor-

dinate gives $\|\mathbf{x}\|^2 = \sum_{p \neq m} |x_p|^2 = \|\varphi(\mathbf{x})\|^2$. Dividing the first identity by the second, which is positive when $\mathbf{x} \neq \mathbf{0}$, gives the statement about the Rayleigh quotients (Definition 17.1.1). This proves the lemma. ■

Because φ is an isomorphism, it matches the k -dimensional subspaces of H with the k -dimensional subspaces of F^{n-1} , for every k . So a statement about $R_{\mathbf{B}}$ on subspaces of F^{n-1} is a statement about $R_{\mathbf{A}}$ on subspaces of H , and the only thing lost in passing from $\mathbf{A}$ to $\mathbf{B}$ is the subspaces of F^n that are not inside H .

17.4.2 Interlacing

Theorem 17.4.2. Cauchy's Interlacing Theorem

Let $\mathbf{A} \in M_n(F)$ be Hermitian with $n \geq 2$, let $I \subseteq \{1, \dots, n\}$ with $|I| = n - 1$, and let $\mathbf{B} = \mathbf{A}_{I,I}$. Then

$$\lambda_k(\mathbf{A}) \geq \lambda_k(\mathbf{B}) \geq \lambda_{k+1}(\mathbf{A}) \quad (1 \leq k \leq n - 1),$$

that is,

$$\lambda_1(\mathbf{A}) \geq \lambda_1(\mathbf{B}) \geq \lambda_2(\mathbf{A}) \geq \lambda_2(\mathbf{B}) \geq \dots \geq \lambda_{n-1}(\mathbf{B}) \geq \lambda_n(\mathbf{A}).$$

Idea. Both halves of Theorem 17.2.3 are used, one for each inequality, and each time the move is the same: a subspace of H is also a subspace of F^n .

① For $\lambda_k(\mathbf{B}) \leq \lambda_k(\mathbf{A})$ use the **max–min** form. The best k -dimensional subspace for $\mathbf{B}$ lives inside H ; transplanted to F^n it is one competitor among more, so the maximum over all of them is at least as large.

② For $\lambda_{k+1}(\mathbf{A}) \leq \lambda_k(\mathbf{B})$ use the **min–max** form, and count dimensions. The best subspace for $\mathbf{B}$ there has dimension $(n - 1) - k + 1 = n - k$, and $n - k = n - (k + 1) + 1$ is exactly the dimension the min–max form for $\lambda_{k+1}(\mathbf{A})$ asks for. So the same subspace, viewed in F^n , is admissible for $\mathbf{A}$ at the index $k + 1$.

The dimension count that makes min–max work at all is Lemma 17.2.1, and it has already been spent inside Theorem 17.2.3; nothing more is needed here than the coincidence of the two numbers in ②.

Proof. Let m be the index not in I , and let H and φ be as in Lemma 17.4.1, so that $R_{\mathbf{A}}(\mathbf{x}) = R_{\mathbf{B}}(\varphi(\mathbf{x}))$ for every nonzero $\mathbf{x} \in H$. Both $\mathbf{A}$ and $\mathbf{B}$ are Hermitian, so Theorem 17.2.3 applies to each, to $\mathbf{A}$ with n and to $\mathbf{B}$ with $n - 1$ in the role of the size. Fix k with $1 \leq k \leq n - 1$.

($\lambda_k(\mathbf{B}) \leq \lambda_k(\mathbf{A})$) By the max–min form of Theorem 17.2.3 for $\mathbf{B}$, there is a subspace $W' \leq F^{n-1}$ with $\dim W' = k$ and

$$\lambda_k(\mathbf{B}) = \min_{\mathbf{0} \neq \mathbf{y} \in W'} R_{\mathbf{B}}(\mathbf{y}).$$

Put $W = \varphi^{-1}(W') \leq H \leq F^n$. Since φ is an isomorphism, $\dim W = k$, and φ carries the

nonzero vectors of W onto the nonzero vectors of W' , so

$$\min_{\mathbf{0} \neq \mathbf{x} \in W} R_{\mathbf{A}}(\mathbf{x}) = \min_{\mathbf{0} \neq \mathbf{y} \in W'} R_{\mathbf{B}}(\mathbf{y}) = \lambda_k(\mathbf{B}).$$

Now W is one of the k -dimensional subspaces of F^n over which the max-min form for $\mathbf{A}$ takes its maximum, so $\lambda_k(\mathbf{A}) \geq \lambda_k(\mathbf{B})$.

($\lambda_{k+1}(\mathbf{A}) \leq \lambda_k(\mathbf{B})$) By the min-max form of Theorem 17.2.3 for $\mathbf{B}$, whose matrices have size $n-1$, there is a subspace $W' \leq F^{n-1}$ with $\dim W' = (n-1) - k + 1 = n - k$ and

$$\lambda_k(\mathbf{B}) = \max_{\mathbf{0} \neq \mathbf{y} \in W'} R_{\mathbf{B}}(\mathbf{y}).$$

Put $W = \varphi^{-1}(W') \leq F^n$, of dimension $n - k$. Since $1 \leq k + 1 \leq n$ and $n - (k + 1) + 1 = n - k = \dim W$, the subspace W is admissible in the min-max form for $\lambda_{k+1}(\mathbf{A})$, and that form takes the minimum over all such subspaces. Hence

$$\lambda_{k+1}(\mathbf{A}) \leq \max_{\mathbf{0} \neq \mathbf{x} \in W} R_{\mathbf{A}}(\mathbf{x}) = \max_{\mathbf{0} \neq \mathbf{y} \in W'} R_{\mathbf{B}}(\mathbf{y}) = \lambda_k(\mathbf{B}).$$

This proves the theorem. ■

The picture to keep is one of nesting. The interval $[\lambda_n(\mathbf{A}), \lambda_1(\mathbf{A})]$ contains the interval $[\lambda_{n-1}(\mathbf{B}), \lambda_1(\mathbf{B})]$, and inside it the two lists alternate, an eigenvalue of $\mathbf{B}$ between each consecutive pair of eigenvalues of $\mathbf{A}$. Deleting a row and a column can only shrink the spectrum inward, never spread it out.

Quick check. Both inequalities in Theorem 17.4.2 can be equalities. Give a matrix $\mathbf{A}$ for which $\lambda_1(\mathbf{A}) = \lambda_1(\mathbf{B})$, and one for which $\lambda_1(\mathbf{B}) = \lambda_2(\mathbf{A})$.

Solution. For the first, $\mathbf{A} = \mathbf{I}_2$ with $I = \{1\}$: then $\mathbf{B} = (1)$ and $\lambda_1(\mathbf{A}) = \lambda_1(\mathbf{B}) = 1$. For the second, $\mathbf{A} = \text{diag}(1, 0)$ with $I = \{2\}$: then $\mathbf{B} = \mathbf{O}_1$, and $\lambda_1(\mathbf{B}) = 0 = \lambda_2(\mathbf{A})$. Nothing in the theorem forces a gap; when a gap is forced is exactly the question settled later in this section.

Deleting one index at a time and iterating gives the general statement.

Corollary 17.4.3. Deleting m Rows and the Same m Columns

Let $\mathbf{A} \in M_n(F)$ be Hermitian, let $0 \leq m \leq n - 1$, let $I \subseteq \{1, \dots, n\}$ with $|I| = n - m$, and let $\mathbf{B} = \mathbf{A}_{I,I}$. Then

$$\lambda_i(\mathbf{A}) \geq \lambda_i(\mathbf{B}) \geq \lambda_{i+m}(\mathbf{A}) \quad (1 \leq i \leq n - m).$$

Proof. Induction on m . For $m = 0$ we have $\mathbf{B} = \mathbf{A}$ and both inequalities are equalities.

Let $m \geq 1$ and suppose the statement holds for $m - 1$, for every Hermitian matrix. Choose an index $j \notin I$ and put $I' = I \cup \{j\}$, so $|I'| = n - m + 1$, and $\mathbf{B}' = \mathbf{A}_{I',I'} \in M_{n-m+1}(F)$,

which is Hermitian. Since $I \subseteq I'$, taking the principal submatrix on I of $\mathbf{B}'$ means keeping the rows and columns of $\mathbf{A}$ indexed by I , so $(\mathbf{B}')_{I,I} = \mathbf{B}$ after relabeling the indices of I inside I' ; and $|I| = |I'| - 1$.

Fix i with $1 \leq i \leq n - m$. By the inductive hypothesis applied to $\mathbf{A}$ and $\mathbf{B}'$,

$$\lambda_i(\mathbf{A}) \geq \lambda_i(\mathbf{B}') \geq \lambda_{i+m-1}(\mathbf{A}),$$

which is available because $i \leq n - m < n - m + 1$. By [Theorem 17.4.2](#) applied to $\mathbf{B}'$ and its principal submatrix $\mathbf{B}$, which is one size smaller,

$$\lambda_k(\mathbf{B}') \geq \lambda_k(\mathbf{B}) \geq \lambda_{k+1}(\mathbf{B}') \quad (1 \leq k \leq n - m).$$

Combining, $\lambda_i(\mathbf{B}) \leq \lambda_i(\mathbf{B}') \leq \lambda_i(\mathbf{A})$, which is the first inequality. For the second, take $k = i$ in the display above and then use the inductive hypothesis at the index $i + 1$, legitimate since $i + 1 \leq n - m + 1$:

$$\lambda_i(\mathbf{B}) \geq \lambda_{i+1}(\mathbf{B}') \geq \lambda_{(i+1)+m-1}(\mathbf{A}) = \lambda_{i+m}(\mathbf{A}).$$

This proves the corollary. ■

The extreme case $m = n - 1$ is worth writing down, because it recovers a fact about diagonal entries that also follows from Chapter 13's description of the extreme eigenvalues.

Example. The diagonal lies inside the spectrum

Let $\mathbf{A} \in M_n(F)$ be Hermitian. Show that $\lambda_n(\mathbf{A}) \leq a_{ii} \leq \lambda_1(\mathbf{A})$ for every i .

Solution. Take $I = \{i\}$, so $m = n - 1$ and $\mathbf{B} = \mathbf{A}_{I,I} = (a_{ii})$, a 1×1 Hermitian matrix whose single eigenvalue is a_{ii} — real, as a diagonal entry of a Hermitian matrix must be. [Corollary 17.4.3](#) at $i = 1$ gives

$$\lambda_1(\mathbf{A}) \geq \lambda_1(\mathbf{B}) = a_{ii} \geq \lambda_{1+(n-1)}(\mathbf{A}) = \lambda_n(\mathbf{A}).$$

The same conclusion follows from Lemma 13.5.2 applied to the unit vector $\mathbf{e}_i$, since $\langle \mathbf{A}\mathbf{e}_i, \mathbf{e}_i \rangle = a_{ii}$. What interlacing adds is the version for larger index sets: any k diagonal entries are the diagonal of a $k \times k$ principal submatrix, whose i -th eigenvalue is trapped between $\lambda_{i+n-k}(\mathbf{A})$ and $\lambda_i(\mathbf{A})$. How much more can be said about the diagonal of a Hermitian matrix, given its spectrum, is the subject of Section 8.

17.4.3 When is the interlacing strict?

[Theorem 17.4.2](#) allows equalities, and the Quick check produced them. Deciding when they occur is the delicate part of the subject, and it is easiest to see when the deleted index is the last one, so that $\mathbf{A}$ is presented as $\mathbf{B}$ with a border added.

Theorem 17.4.4. Bordering a Hermitian Matrix

Let $\mathbf{B} \in M_{n-1}(F)$ be Hermitian with $n \geq 2$, let $\mathbf{b} \in F^{n-1}$, let $c \in \mathbb{R}$, and let

$$\mathbf{A} = \begin{pmatrix} \mathbf{B} & \mathbf{b} \\ \mathbf{b}^* & c \end{pmatrix} \in M_n(F),$$

which is Hermitian and has $\mathbf{A}_{I,I} = \mathbf{B}$ for $I = \{1, \dots, n-1\}$. For an eigenvalue ν of $\mathbf{B}$, write P_ν for the orthogonal projection of F^{n-1} onto the eigenspace $E_\nu(\mathbf{B})$, and $m_\nu = \dim E_\nu(\mathbf{B})$.

- (a) If $P_\nu \mathbf{b} \neq \mathbf{0}$, then ν is a root of $p_{\mathbf{A}}$ of multiplicity **exactly** $m_\nu - 1$; in particular, if $m_\nu = 1$ then ν is not an eigenvalue of $\mathbf{A}$.
- (b) If $P_\nu \mathbf{b} = \mathbf{0}$, then ν is a root of $p_{\mathbf{A}}$ of multiplicity **at least** m_ν .
- (c) The interlacing $\lambda_1(\mathbf{A}) \geq \lambda_1(\mathbf{B}) \geq \dots \geq \lambda_{n-1}(\mathbf{B}) \geq \lambda_n(\mathbf{A})$ is strict at every one of its $2(n-1)$ inequalities if and only if $\mathbf{B}$ has $n-1$ **distinct** eigenvalues **and** $P_\nu \mathbf{b} \neq \mathbf{0}$ for every eigenvalue ν of $\mathbf{B}$.

Idea. Write the characteristic polynomial of $\mathbf{A}$ down. With $\mathbf{B}$ diagonalized, a Schur complement turns $\det(x\mathbf{I} - \mathbf{A})$ into

$$\left(\prod_k (x - \mu_k) \right) \left((x - c) - \sum_k \frac{|\tilde{b}_k|^2}{x - \mu_k} \right),$$

and clearing the denominators leaves a polynomial identity in which each μ_k appears with one factor fewer than one might expect — provided $P_\nu \mathbf{b} \neq \mathbf{0}$, that is, provided the $|\tilde{b}_k|^2$ over the group belonging to that eigenvalue do not all vanish (for a repeated eigenvalue the individual $\tilde{b}_k$ depend on the choice of basis, but their sum does not). Part (c) then costs almost nothing: strict interlacing means the combined list of $2n-1$ numbers is strictly decreasing, hence has no repetitions at all, and (a) and (b) say exactly which repetitions are possible.

Proof. Step 1: diagonalize the corner. By the spectral theorem (Corollary 12.4.3 over $\mathbb{C}$, Corollary 12.5.2 over $\mathbb{R}$) write $\mathbf{B} = \mathbf{U}\mathbf{D}\mathbf{U}^*$ with $\mathbf{U} \in M_{n-1}(F)$ unitary and $\mathbf{D} = \text{diag}(\mu_1, \dots, \mu_{n-1})$ carrying the eigenvalues of $\mathbf{B}$ in decreasing order, the columns $\mathbf{u}_1, \dots, \mathbf{u}_{n-1}$ of $\mathbf{U}$ being an orthonormal basis of eigenvectors. Let $\mathbf{V} = \begin{pmatrix} \mathbf{U} & \mathbf{0} \\ \mathbf{0}^* & 1 \end{pmatrix} \in M_n(F)$, which is unitary because $\mathbf{V}^* \mathbf{V} = \mathbf{I}_n$. Multiplying in blocks (Theorem 8.1.1) gives

$$\mathbf{V}^* \mathbf{A} \mathbf{V} = \begin{pmatrix} \mathbf{D} & \tilde{\mathbf{b}} \\ \tilde{\mathbf{b}}^* & c \end{pmatrix}, \quad \tilde{\mathbf{b}} = \mathbf{U}^* \mathbf{b}.$$

Similar matrices have the same characteristic polynomial (Theorem 7.10.3), so we may compute with the right-hand side. Moreover $\tilde{b}_k = (\mathbf{U}^* \mathbf{b})_k = \langle \mathbf{b}, \mathbf{u}_k \rangle$, and the $\mathbf{u}_k$ with $\mu_k = \nu$ form an orthonormal basis of $E_\nu(\mathbf{B})$: they lie in it, and a vector $\mathbf{x} = \sum_k c_k \mathbf{u}_k$ with $\mathbf{B}\mathbf{x} = \nu\mathbf{x}$ has $(\mu_k - \nu)c_k = 0$ for every k , so $c_k = 0$ unless $\mu_k = \nu$. Hence exactly m_ν of the μ_k equal ν , and

$$P_\nu \mathbf{b} = \sum_{k: \mu_k = \nu} \langle \mathbf{b}, \mathbf{u}_k \rangle \mathbf{u}_k \quad \text{and} \quad \|P_\nu \mathbf{b}\|^2 = \sum_{k: \mu_k = \nu} |\tilde{b}_k|^2.$$

In particular $P_\nu \mathbf{b} = \mathbf{0}$ exactly when $\tilde{b}_k = 0$ for every k with $\mu_k = \nu$.

Step 2: the characteristic polynomial. We claim that, as polynomials in x ,

$$\begin{aligned}
p_{\mathbf{A}}(x) &= (x - c) \prod_{k=1}^{n-1} (x - \mu_k) \\
&\quad - \sum_{k=1}^{n-1} |\tilde{b}_k|^2 \prod_{l \neq k} (x - \mu_l).
\end{aligned} \tag{17.4.1}$$

Claim. Eq. 17.4.1 holds.

Proof. Fix a scalar x_0 that is not an eigenvalue of $\mathbf{B}$. Then $x_0 \mathbf{I}_{n-1} - \mathbf{D}$ is invertible, being diagonal with nonzero diagonal entries, so Theorem 8.3.2 (a), applied to

$$\mathbf{M} = x_0 \mathbf{I}_n - \mathbf{V}^* \mathbf{A} \mathbf{V} = \begin{pmatrix} x_0 \mathbf{I}_{n-1} - \mathbf{D} & -\tilde{\mathbf{b}} \\ -\tilde{\mathbf{b}}^* & x_0 - c \end{pmatrix}$$

with the invertible block in the top left, gives $\det \mathbf{M} = \det(x_0 \mathbf{I} - \mathbf{D}) \cdot \det(\mathbf{M}/(x_0 \mathbf{I} - \mathbf{D}))$, where the Schur complement is the 1×1 matrix

$$(x_0 - c) - (-\tilde{\mathbf{b}}^*)(x_0 \mathbf{I} - \mathbf{D})^{-1}(-\tilde{\mathbf{b}}) = (x_0 - c) - \sum_k \frac{|\tilde{b}_k|^2}{x_0 - \mu_k}.$$

Multiplying out $\det(x_0 \mathbf{I} - \mathbf{D}) = \prod_k (x_0 - \mu_k)$ against this scalar, and canceling one factor $(x_0 - \mu_k)$ in the k -th summand, gives exactly the right-hand side of Eq. 17.4.1 evaluated at x_0 . So the two sides of Eq. 17.4.1, which are polynomials of degree n , agree at every scalar outside the finite set $\text{spec}(\mathbf{B})$. Their difference is therefore a polynomial of degree at most n with infinitely many roots in F , and a nonzero polynomial of degree at most n has at most n roots (Corollary 6.1.4); so the difference is zero. ■

Step 3: multiplicities. Let $\nu_1 > \dots > \nu_p$ be the distinct eigenvalues of $\mathbf{B}$, with multiplicities $m_1, \dots, m_p$ summing to $n - 1$, and put $\beta_t = \|P_{\nu_t} \mathbf{b}\|^2$. Grouping equal μ_k in Eq. 17.4.1 — the product $\prod_{l \neq k} (x - \mu_l)$ is the same for all k in one group, and the coefficients $|\tilde{b}_k|^2$ add up to β_t by Step 1 — turns it into

$$\begin{aligned}
p_{\mathbf{A}}(x) &= (x - c) \prod_{t=1}^p (x - \nu_t)^{m_t} \\
&\quad - \sum_{t=1}^p \beta_t (x - \nu_t)^{m_t-1} \prod_{s \neq t} (x - \nu_s)^{m_s}.
\end{aligned}$$

Fix t . Every term on the right is divisible by $(x - \nu_t)^{m_t-1}$, so $p_{\mathbf{A}}(x) = (x - \nu_t)^{m_t-1} g(x)$ with

$$\begin{aligned}
g(x) &= (x - c)(x - \nu_t) \prod_{s \neq t} (x - \nu_s)^{m_s} - \beta_t \prod_{s \neq t} (x - \nu_s)^{m_s} \\
&\quad - \sum_{u \neq t} \beta_u (x - \nu_t)(x - \nu_u)^{m_u-1} \prod_{s \neq u, t} (x - \nu_s)^{m_s}.
\end{aligned}$$

Evaluate at $x = \nu_t$. The first term vanishes because of its factor $(x - \nu_t)$, and so does every summand of the third, for the same reason. What is left is

$$g(\nu_t) = -\beta_t \prod_{s \neq t} (\nu_t - \nu_s)^{m_s},$$

and the product is nonzero because the ν_s are distinct. So $g(\nu_t) \neq 0$ exactly when $\beta_t \neq 0$, that is, exactly when $P_{\nu_t} \mathbf{b} \neq \mathbf{0}$. In that case ν_t is a root of $p_{\mathbf{A}}$ of multiplicity exactly $m_t - 1$, which proves (a); in the other case $(x - \nu_t)$ divides g , so the multiplicity is at least m_t , which proves (b).

Step 4: strictness. ($\Leftarrow$) Suppose $\mathbf{B}$ has $n - 1$ distinct eigenvalues, so every $m_t = 1$, and $P_{\nu} \mathbf{b} \neq \mathbf{0}$ for each of them. By (a), no eigenvalue of $\mathbf{B}$ is an eigenvalue of $\mathbf{A}$. [Theorem 17.4.2](#) gives $\lambda_k(\mathbf{A}) \geq \lambda_k(\mathbf{B}) \geq \lambda_{k+1}(\mathbf{A})$ for each k ; since $\lambda_k(\mathbf{B})$ is an eigenvalue of $\mathbf{B}$ and $\lambda_k(\mathbf{A})$, $\lambda_{k+1}(\mathbf{A})$ are eigenvalues of $\mathbf{A}$, neither inequality can be an equality. So all $2(n - 1)$ of them are strict.

($\Rightarrow$) Suppose all of them are strict, that is,

$$\lambda_1(\mathbf{A}) > \lambda_1(\mathbf{B}) > \lambda_2(\mathbf{A}) > \cdots > \lambda_{n-1}(\mathbf{B}) > \lambda_n(\mathbf{A}).$$

Reading two steps at a time gives $\lambda_k(\mathbf{B}) > \lambda_{k+1}(\mathbf{A}) > \lambda_{k+1}(\mathbf{B})$ for $1 \leq k \leq n - 2$, so the eigenvalues of $\mathbf{B}$ are distinct. And the displayed chain is strictly decreasing, so its $2n - 1$ entries are pairwise distinct; in particular no eigenvalue of $\mathbf{B}$ is an eigenvalue of $\mathbf{A}$. By (b), read as a contrapositive, $P_{\nu} \mathbf{b} \neq \mathbf{0}$ for every eigenvalue ν of $\mathbf{B}$. This proves the theorem. ■

Remark. Part (b) is deliberately one-sided. When $P_{\nu} \mathbf{b} = \mathbf{0}$ the multiplicity of ν in $p_{\mathbf{A}}$ may be m_{ν} or $m_{\nu} + 1$ — never more, since by Cauchy interlacing q equal eigenvalues ν of $\mathbf{A}$ squeeze $q - 1$ eigenvalues of $\mathbf{B}$ to ν — and both occur: $\mathbf{B} = \text{diag}(3, 1)$, $\mathbf{b} = (0, 1)$, $c = 0$ gives $p_{\mathbf{A}}(x) = (x - 3)(x^2 - x - 1)$, where 3 keeps multiplicity $1 = m_3$; while $\mathbf{B} = (\mu)$, $\mathbf{b} = \mathbf{0}$, $c = \mu$ gives $\mathbf{A} = \mu \mathbf{I}_2$, where μ has multiplicity $2 = m_{\mu} + 1$. Deciding between the two needs the value of c as well, and [Theorem 17.4.4](#) does not attempt it; what it does settle completely is the strictness question (c), which is what interlacing asks.

Warning. **Strictness needs both clauses, and a nonzero border is not enough.** It is tempting to say that the interlacing is strict as soon as $\mathbf{b} \neq \mathbf{0}$, or as soon as $\mathbf{b}$ meets every eigenspace of $\mathbf{B}$. The first is wrong even when the eigenvalues of $\mathbf{B}$ are distinct (Exercise B3(a)); the second fails exactly when $\mathbf{B}$ has a repeated eigenvalue. Take $\mathbf{B} = 2\mathbf{I}_2$, $\mathbf{b} = (1, 1)$, and $c = 2$. Here $\mathbf{b}$ has a nonzero component in the one eigenspace $E_2(\mathbf{B}) = F^2$, and yet $\lambda_1(\mathbf{B}) = \lambda_2(\mathbf{B}) = 2$, so the middle of the interlacing chain reads $2 \geq \lambda_2(\mathbf{A}) \geq 2$ and $\lambda_2(\mathbf{A}) = 2$ with no freedom at all. A repeated eigenvalue of $\mathbf{B}$ is passed on to $\mathbf{A}$, by [Theorem 17.4.4](#) (a) with $m_{\nu} = 2$.

To apply [Theorem 17.4.4](#) when the deleted index is not the last one, permute. Let $I = \{i_1 < \cdots < i_{n-1}\} = \{1, \dots, n\} \setminus \{m\}$, and let $\mathbf{P}$ be the matrix with columns $\mathbf{e}_{i_1}, \dots, \mathbf{e}_{i_{n-1}}, \mathbf{e}_m$, a permutation matrix. Its columns are orthonormal, so $\mathbf{P}^{\top} \mathbf{P} = \mathbf{I}$ and $\mathbf{P}$ is unitary (Lemma 3.4.7 (b)); hence $\mathbf{P}^{\top} \mathbf{A} \mathbf{P}$ is Hermitian and similar to $\mathbf{A}$, with the same eigenvalues. Its (r, s) entry is $\mathbf{e}_{i_r}^{\top} \mathbf{A} \mathbf{e}_{i_s} = a_{i_r i_s}$ for $r, s \leq n - 1$, so its leading $(n - 1) \times (n - 1)$ corner is $\mathbf{A}_{I, I}$, and its border is the column $(a_{i_1 m}, \dots, a_{i_{n-1} m})$ with corner entry a_{mm} . So $\mathbf{P}^{\top} \mathbf{A} \mathbf{P}$ is a bordered matrix with corner $\mathbf{A}_{I, I}$, and [Theorem 17.4.4](#) applies to it verbatim.

The strictness criterion pays a debt from Chapter 11. The solution of Exercise (C1: Interlacing roots) proved by hand that the roots of consecutive orthogonal polynomials interlace for degrees

1, 2, 3, and then said: “The general case is not proved this way. The coefficients a_k and b_k assemble into a symmetric tridiagonal matrix whose characteristic polynomial is p_k , and the interlacing then follows from the interlacing theorem for the eigenvalues of a symmetric matrix and its leading submatrices, which is Chapter 17’s business.” Two things in that sentence were asserted, not proved: that the characteristic polynomial is p_k , and that the interlacing is **strict**, which is what “exactly one root strictly between” demands. Plain [Theorem 17.4.2](#) gives only $\geq$. Strictness needs [Theorem 17.4.4](#) (c), and (c) needs one fact about tridiagonal matrices.

Lemma 17.4.5. Eigenvectors of an Unreduced Tridiagonal Matrix

Let $\mathbf{T} \in M_k(\mathbb{R})$ be symmetric and tridiagonal, with every entry $t_{i,i+1} = t_{i+1,i}$ ($1 \leq i \leq k-1$) **nonzero**. Then every eigenvector $\mathbf{u}$ of $\mathbf{T}$ has $u_k \neq 0$. Consequently every eigenspace of $\mathbf{T}$ has dimension 1, so $\mathbf{T}$ has k distinct eigenvalues.

Proof. Let $\mathbf{T}\mathbf{u} = \nu\mathbf{u}$ and suppose $u_k = 0$; we show $\mathbf{u} = \mathbf{0}$. If $k = 1$ this is immediate. Otherwise row k of $\mathbf{T}\mathbf{u} = \nu\mathbf{u}$ reads $t_{k,k-1}u_{k-1} + t_{kk}u_k = \nu u_k$, so $t_{k,k-1}u_{k-1} = 0$ and $u_{k-1} = 0$, since $t_{k,k-1} \neq 0$. Now suppose $u_i = u_{i+1} = 0$ for some $2 \leq i \leq k-1$. Row i reads $t_{i,i-1}u_{i-1} + t_{ii}u_i + t_{i,i+1}u_{i+1} = \nu u_i$, so $t_{i,i-1}u_{i-1} = 0$ and $u_{i-1} = 0$. Walking up from $i = k-1$ to $i = 2$ gives $u_{k-2} = \cdots = u_1 = 0$, so $\mathbf{u} = \mathbf{0}$, and $\mathbf{u}$ is not an eigenvector.

If some eigenspace $E_\nu(\mathbf{T})$ contained two independent vectors $\mathbf{u}, \mathbf{u}'$, then $u'_k\mathbf{u} - u_k\mathbf{u}'$ would be a nonzero vector of $E_\nu(\mathbf{T})$ — nonzero by independence, since $u_k \neq 0$ — with last coordinate 0, which we have just excluded. So each eigenspace is a line. Since $\mathbf{T}$ is real symmetric, its eigenspaces have dimensions adding up to k (Corollary 12.5.2), so there are k of them, that is, k distinct eigenvalues. This proves the lemma. ■

Corollary 17.4.6. Zeros of Orthogonal Polynomials Interlace

Optional: nothing later depends on this.

Let (p_k) be the monic orthogonal polynomial sequence of an inner product on $\mathbb{R}[x]$ to which Theorem 11.10.3 applies, with coefficients a_k and $b_k > 0$; for the node inner product on $\mathbb{R}[x]_{\leq n}$ read everything below for $k+1 \leq n$. For $k \geq 1$ let $\mathbf{T}_k$ (the **Jacobi matrix** of the sequence) be the $k \times k$ real symmetric tridiagonal matrix with diagonal $a_0, a_1, \dots, a_{k-1}$ and entries $\sqrt{b_1}, \dots, \sqrt{b_{k-1}}$ just above and just below it.

- (a) $p_{\mathbf{T}_k} = p_k$; so the roots of p_k are the eigenvalues of $\mathbf{T}_k$, and they are real and distinct.
- (b) If $r_1 > \cdots > r_{k+1}$ are the roots of p_{k+1} and $s_1 > \cdots > s_k$ those of p_k , then

$$r_1 > s_1 > r_2 > s_2 > \cdots > s_k > r_{k+1}.$$

Idea. The polynomials p_k are the characteristic polynomials of a nested family of symmetric tridiagonal matrices, each the leading corner of the next. So their zeros are eigenvalues of a matrix and of its corner, and Cauchy interlacing applies at once — nonstrictly. Strictness is the bordered-matrix criterion, and both of its clauses come from one fact about tridiagonal matrices with nonzero off-diagonal entries: an eigenvector can never end in a zero.

Proof.

- (a) Put $D_k(x) = \det(x\mathbf{I}_k - \mathbf{T}_k)$ and $D_0 = 1$. For each real x_0 , the matrix $x_0\mathbf{I}_k - \mathbf{T}_k$ is tridiagonal in the sense of Chapter 7, with diagonal entries $x_0 - a_{i-1}$ and off-diagonal entries $-\sqrt{b_i}$ in both positions $(i, i+1)$ and $(i+1, i)$, so Theorem 7.8.3 gives $D_1(x_0) = x_0 - a_0$ and

$$\begin{aligned} D_{k+1}(x_0) &= (x_0 - a_k)D_k(x_0) - (-\sqrt{b_k})^2 D_{k-1}(x_0) \\ &= (x_0 - a_k)D_k(x_0) - b_k D_{k-1}(x_0) \end{aligned}$$

for $k \geq 1$. Here $D_k(x_0) = \det(x_0\mathbf{I} - \mathbf{T}_k)$ is the characteristic polynomial evaluated at x_0 (Lemma 7.10.1), and two real polynomials that agree at every real number are equal (Theorem 6.1.5 (b)), so these identities hold as polynomials. They are the recurrence of Theorem 11.10.3 with the same starting values $p_0 = 1$, $p_1 = x - a_0$, so $D_k = p_k$ for every k by induction. The roots of $p_{\mathbf{T}_k}$ are the eigenvalues of $\mathbf{T}_k$ (Theorem 9.2.1), which are real because $\mathbf{T}_k$ is real symmetric (Theorem 12.2.1) and distinct by Lemma 17.4.5, whose hypothesis holds because every $\sqrt{b_i} > 0$.

- (b) Write $\mathbf{T}_{k+1}$ in bordered form:

$$\mathbf{T}_{k+1} = \begin{pmatrix} \mathbf{T}_k & \sqrt{b_k} \mathbf{e}_k \\ \sqrt{b_k} \mathbf{e}_k^\top & a_k \end{pmatrix}, \quad \mathbf{e}_k \in \mathbb{R}^k.$$

The border of $\mathbf{T}_{k+1}$ is $\mathbf{w} = \sqrt{b_k} \mathbf{e}_k$, whose k -th entry is $\sqrt{b_k}$. We call it $\mathbf{w}$ rather than $\mathbf{b}$ because b_k already names a recurrence coefficient. We check the two clauses of Theorem 17.4.4 (c) for $\mathbf{B} = \mathbf{T}_k$, with $\mathbf{w}$ in place of $\mathbf{b}$. By Lemma 17.4.5, $\mathbf{T}_k$ has k distinct eigenvalues. For each eigenvalue ν , the eigenspace $E_\nu(\mathbf{T}_k)$ is spanned by one unit vector $\mathbf{u}$, so $P_\nu \mathbf{w} = \langle \mathbf{w}, \mathbf{u} \rangle \mathbf{u} = \sqrt{b_k} u_k \mathbf{u}$, which is nonzero because $b_k > 0$ and $u_k \neq 0$ by the same lemma. So Theorem 17.4.4 (c) applies, and the eigenvalues of $\mathbf{T}_{k+1}$ and of $\mathbf{T}_k$ interlace strictly. By (a) these are the roots of p_{k+1} and of p_k . This proves the corollary. ■

Part (a) gives a second proof, with no integrals, that the roots of p_k are real and simple — half of Theorem 11.10.5. The other half, that they lie inside the interval of the weight, is not visible in the matrix and still needs the argument of Chapter 11. For the monic Legendre polynomials of Example (The monic Legendre polynomials, continued), $a_k = 0$, $b_1 = \frac{1}{3}$, and $b_2 = \frac{4}{15}$, so

$$\mathbf{T}_3 = \begin{pmatrix} 0 & 1/\sqrt{3} & 0 \\ 1/\sqrt{3} & 0 & 2/\sqrt{15} \\ 0 & 2/\sqrt{15} & 0 \end{pmatrix},$$

and expanding along the first row,

$$p_{\mathbf{T}_3}(x) = x^3 - \left(\frac{1}{3} + \frac{4}{15}\right)x = x^3 - \frac{3}{5}x = p_3.$$

Its leading corner $\mathbf{T}_2$ has characteristic polynomial $x^2 - \frac{1}{3} = p_2$ and unit eigenvectors $\frac{1}{\sqrt{2}}(1, \pm 1)$, both with nonzero last coordinate, so the border $\frac{2}{\sqrt{15}}\mathbf{e}_2$ meets both eigenspaces. The prediction is strict interlacing, and indeed

$$\sqrt{3/5} > 1/\sqrt{3} > 0 > -1/\sqrt{3} > -\sqrt{3/5},$$

that is $0.7746 > 0.5774 > 0 > -0.5774 > -0.7746$, where the strict comparisons $3/5 > 1/3$ and $1/\sqrt{3} > 0$ need no decimals.

17.4.4 A worked example

Example. One matrix and all three of its corners

Let

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

Compute the eigenvalues of $\mathbf{A}$ and of each of its three 2×2 principal submatrices, exhibit the interlacing in each case, and explain the pattern of strict and nonstrict inequalities using [Theorem 17.4.4](#).

Solution. The eigenvalues of $\mathbf{A}$. Expanding $\det(x\mathbf{I} - \mathbf{A})$ along the first row,

$$\begin{aligned} p_{\mathbf{A}}(x) &= (x - 2)[(x - 2)^2 - 1] + 1 \cdot [-(x - 2)] \\ &= (x - 2)[(x - 2)^2 - 2], \end{aligned}$$

so

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

The three corners. With I the index set kept,

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

A matrix $\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$ has trace 4 and determinant 3, so its eigenvalues are 3 and 1; and $2\mathbf{I}_2$ has the eigenvalue 2 twice.

The interlacing. In all three cases [Theorem 17.4.2](#) is confirmed:

$$\begin{aligned} I = \{1, 2\} : \quad & 3.41421 > 3 > 2 > 1 > 0.58579, \\ I = \{1, 3\} : \quad & 3.41421 > 2 = 2 = 2 > 0.58579, \\ I = \{2, 3\} : \quad & 3.41421 > 3 > 2 > 1 > 0.58579. \end{aligned}$$

The first and third are strict everywhere; the second has two equalities in the middle. The exact comparisons need no decimals: $2 + \sqrt{2} > 3$ is $\sqrt{2} > 1$, and $1 > 2 - \sqrt{2}$ is $\sqrt{2} > 1$ again.

Why. Read $\mathbf{A}$ as a bordered matrix in each case, using the permutation remark above.

For $I = \{1, 2\}$ the deleted index is already the last one, so $\mathbf{B} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$, $\mathbf{b} = (0, 1)$, and $c = 2$. The eigenvalues 3 and 1 of $\mathbf{B}$ are distinct, with unit eigenvectors $\frac{1}{\sqrt{2}}(1, 1)$ and $\frac{1}{\sqrt{2}}(1, -1)$. Both inner products with $\mathbf{b}$ are nonzero:

$$\langle \mathbf{b}, \frac{1}{\sqrt{2}}(1, 1) \rangle = \frac{1}{\sqrt{2}}, \quad \langle \mathbf{b}, \frac{1}{\sqrt{2}}(1, -1) \rangle = -\frac{1}{\sqrt{2}}.$$

So both clauses of [Theorem 17.4.4](#) (c) hold and the interlacing is strict, as computed. The case $I = \{2, 3\}$ is the same: moving index 1 to the end gives the corner $\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$ again, with border $(a_{21}, a_{31}) = (1, 0)$, whose inner products with the two unit eigenvectors $\frac{1}{\sqrt{2}}(1, \pm 1)$ are both $\frac{1}{\sqrt{2}}$, nonzero.

For $I = \{1, 3\}$ the deleted index is 2. Conjugating by the permutation matrix that sends the coordinate order $(1, 2, 3)$ to $(1, 3, 2)$ turns $\mathbf{A}$ into

$$\begin{pmatrix} 2 & 0 & 1 \\ 0 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}, \quad \mathbf{B} = 2\mathbf{I}_2, \quad \mathbf{b} = (1, 1), \quad c = 2.$$

Here $\mathbf{b} \neq \mathbf{0}$ and $P_2\mathbf{b} = \mathbf{b} \neq \mathbf{0}$, so the first clause of [Theorem 17.4.4](#) (c) is the one that fails: $\mathbf{B}$ has the repeated eigenvalue 2. Part (a) with $m_2 = 2$ then says that 2 is an eigenvalue of $\mathbf{A}$ of multiplicity exactly 1, which is visible in $p_{\mathbf{A}}$, and the interlacing collapses at the middle to $2 = 2 = 2$. This is exactly the situation of the warning above, occurring inside a perfectly ordinary integer matrix.

17.4.5 Inertia under compression

Interlacing controls not only where the eigenvalues are but how many of them have each sign, and that is a statement about inertia.

Corollary 17.4.7. A Principal Submatrix Keeps Most of the Inertia

Let $\mathbf{A} \in M_n(F)$ be Hermitian, let $0 \leq m \leq n - 1$, let $|I| = n - m$ and $\mathbf{B} = \mathbf{A}_{I,I}$. Write $n_+(\mathbf{M})$ and $n_-(\mathbf{M})$ for the number of positive and of negative eigenvalues of a Hermitian $\mathbf{M}$, counted with multiplicity. Then

$$n_+(\mathbf{B}) \leq n_+(\mathbf{A}) \leq n_+(\mathbf{B}) + m, \quad n_-(\mathbf{B}) \leq n_-(\mathbf{A}) \leq n_-(\mathbf{B}) + m.$$

In particular $\mathbf{A} \succeq \mathbf{O}$ implies $\mathbf{B} \succeq \mathbf{O}$, and $\mathbf{A} \succ \mathbf{O}$ implies $\mathbf{B} \succ \mathbf{O}$.

Proof. Write $r = n_+(\mathbf{B})$ and $s = n_+(\mathbf{A})$, and abbreviate $\lambda_i = \lambda_i(\mathbf{A})$, $\mu_i = \lambda_i(\mathbf{B})$. By [Corollary 17.4.3](#), $\lambda_i \geq \mu_i \geq \lambda_{i+m}$ for $1 \leq i \leq n - m$.

($n_+(\mathbf{B}) \leq n_+(\mathbf{A})$) For $i \leq r$ we have $\mu_i > 0$, hence $\lambda_i \geq \mu_i > 0$. So $\mathbf{A}$ has at least r positive eigenvalues.

($n_+(\mathbf{A}) \leq n_+(\mathbf{B}) + m$) If $r + m \geq n$ there is nothing to prove, since $s \leq n$. Otherwise $r + 1 \leq n - m$, so the index $i = r + 1$ is admissible and $\mu_{r+1} \leq 0$ by the definition of r ; hence $\lambda_{r+1+m} \leq \mu_{r+1} \leq 0$, and the eigenvalues $\lambda_{r+m+1}, \dots, \lambda_n$ are all ≤ 0 by the decreasing indexing. So $s \leq r + m$.

For the two statements about n_- , apply what has just been proved to $-\mathbf{A}$, whose principal submatrix on I is $-\mathbf{B}$. By Lemma 17.3.2 the eigenvalue lists are negated and reversed, so $n_+(-\mathbf{A}) = n_-(\mathbf{A})$ and $n_+(-\mathbf{B}) = n_-(\mathbf{B})$, and the two displayed chains are the same statement.

Finally, $\mathbf{A} \succeq \mathbf{O}$ means every eigenvalue of $\mathbf{A}$ is ≥ 0 (Theorem 13.1.4 (b)), that is, $n_-(\mathbf{A}) = 0$; then $n_-(\mathbf{B}) \leq n_-(\mathbf{A}) = 0$, so $\mathbf{B} \succeq \mathbf{O}$. And $\mathbf{A} \succ \mathbf{O}$ means $n_+(\mathbf{A}) = n$ (Theorem 13.1.5 (b)); then $n_+(\mathbf{B}) \geq n_+(\mathbf{A}) - m = n - m$, which is the size of $\mathbf{B}$, so every eigenvalue of $\mathbf{B}$ is positive and $\mathbf{B} \succ \mathbf{O}$. This proves the corollary. ■

The last two sentences are a second proof of something Chapter 13 already knew. That a principal submatrix of a positive semidefinite matrix is positive semidefinite is the step (a) $\Rightarrow$ (e) in the proof of Theorem 13.1.4, where it comes from the observation that a principal submatrix is what the quadratic form does to vectors supported on the corresponding coordinates — one line, needing no interlacing. What is new here is the quantitative form. Sylvester’s law of inertia (Theorem 14.5.1) says that the triple of counts is an invariant of a real symmetric matrix up to congruence, and Theorem 14.5.3 identifies Chapter 14’s counts with the eigenvalue counts n_+ and n_- used here; the corollary says that deleting m rows and the matching m columns can destroy at most m positive and at most m negative eigenvalues, and create none of either. Section 10 gives a second proof of Sylvester’s law from Courant–Fischer, and bounds how far a low-rank perturbation can move the inertia.

17.4.6 Interlacing is a Hermitian phenomenon

Everything above rests on Theorem 17.2.3, which rests on the quadratic form being real. Without the Hermitian hypothesis, nothing survives — not a weakened interlacing, not even a bound.

Example. A corner outside the spectrum

Let

$$\mathbf{A} = \begin{pmatrix} 5 & 1 \\ -8 & -1 \end{pmatrix}.$$

Show that $\mathbf{A}$ has real eigenvalues, so that the conclusion of Theorem 17.4.2 can be stated for it, and that the conclusion is false.

Solution. $\text{tr } \mathbf{A} = 4$ and $\det \mathbf{A} = 5(-1) - 1(-8) = 3$, so $p_{\mathbf{A}}(x) = x^2 - 4x + 3 = (x-3)(x-1)$ and the eigenvalues are 3 and 1, both real. The two 1×1 principal submatrices are $\mathbf{A}_{\{1\},\{1\}} = (5)$ and $\mathbf{A}_{\{2\},\{2\}} = (-1)$, with eigenvalues 5 and -1 .

Interlacing would require $3 \geq 5 \geq 1$ in the first case and $3 \geq -1 \geq 1$ in the second. Both fail, and they fail on opposite sides: the diagonal entries 5 and -1 lie outside the interval $[1, 3]$, which is the convex hull of the spectrum. For a Hermitian matrix that is impossible, by Example (The diagonal lies inside the spectrum).

Remark. The failure is total, not marginal. For any $a \in F$, the matrix

$$\mathbf{A}_a = \begin{pmatrix} a & 1 \\ -(a-1)(a-3) & 4-a \end{pmatrix}$$

has trace 4 and determinant $a(4-a) + (a-1)(a-3) = 4a - a^2 + a^2 - 4a + 3 = 3$, hence the same characteristic polynomial $(x-3)(x-1)$ as above. Its $(1, 1)$ principal submatrix is (a) . So a general matrix with spectrum $\{1, 3\}$ can have any prescribed number whatsoever as its $(1, 1)$ entry: the eigenvalue of a single principal submatrix is subject to no constraint at all. (The two diagonal entries together still add up to the trace, 4; that constraint holds for every matrix and has nothing to do with interlacing.) The choice $a = 5$ recovers $\mathbf{A}$, and $a = 0$ gives $\begin{pmatrix} 0 & 1 \\ -3 & 4 \end{pmatrix}$.

Warning. **Principal means the same index set on both sides.** Interlacing says nothing about a submatrix obtained by deleting row i and column j with $i \neq j$, even for a Hermitian $\mathbf{A}$. Take $\mathbf{A} = \begin{pmatrix} -10 & 1 \\ 1 & -10 \end{pmatrix}$, whose eigenvalues are -9 and -11 . Deleting row 1 and column 2 leaves the 1×1 matrix (1) , and 1 is nowhere near $[-11, -9]$. The submatrix in Theorem 17.4.2 is $\mathbf{A}_{I,I}$, and the two I 's are not decoration.

17.4.7 Exercises

17.4.7.1 A. Check your understanding

Exercise. A1

- State Theorem 17.4.2, and say which hypothesis on $\mathbf{A}$ and which hypothesis on the submatrix are used.
- Let $\mathbf{A} \in M_6(\mathbb{C})$ be Hermitian and let $\mathbf{B} = \mathbf{A}_{I,I}$ with $|I| = 4$. Which eigenvalues of $\mathbf{A}$ bound $\lambda_3(\mathbf{B})$ above and below?
- Determine whether the following is correct, and justify your answer: if $\mathbf{b} \neq \mathbf{0}$ then the interlacing between $\begin{pmatrix} \mathbf{B} & \mathbf{b} \\ \mathbf{b}^* & c \end{pmatrix}$ and $\mathbf{B}$ is strict.
- Explain why a principal submatrix of a positive semidefinite matrix is positive semidefinite, in one sentence, without using interlacing.
- Give a 2×2 matrix with real eigenvalues whose $(1, 1)$ entry is larger than both of them.

Solution.

- See Theorem 17.4.2. The matrix $\mathbf{A}$ must be **Hermitian**, so that Theorem 17.2.3 applies and the eigenvalues are real and can be put in order; and $\mathbf{B}$ must be a **principal** submatrix, $\mathbf{B} = \mathbf{A}_{I,I}$ with the same index set on rows and columns, so that Lemma 17.4.1 identifies $R_{\mathbf{B}}$ with $R_{\mathbf{A}}$ restricted to a coordinate hyperplane.
- Here $n = 6$ and $m = 2$, so Corollary 17.4.3 gives $\lambda_3(\mathbf{A}) \geq \lambda_3(\mathbf{B}) \geq \lambda_5(\mathbf{A})$.
- Incorrect. The warning after Theorem 17.4.4 gives $\mathbf{B} = 2\mathbf{I}_2$, $\mathbf{b} = (1, 1)$, $c = 2$, where $\mathbf{b} \neq \mathbf{0}$ and the interlacing is not strict, because $\mathbf{B}$ has a repeated eigenvalue. The correct criterion is Theorem 17.4.4 (c), which needs distinct eigenvalues of $\mathbf{B}$ as well as $P_{\nu}\mathbf{b} \neq \mathbf{0}$ for each ν .
- Because $\langle \mathbf{B}\mathbf{y}, \mathbf{y} \rangle = \langle \mathbf{A}\mathbf{x}, \mathbf{x} \rangle \geq 0$, where $\mathbf{x} \in F^n$ is $\mathbf{y}$ placed in the coordinates of I and 0 elsewhere (Lemma 17.4.1); this is how Theorem 13.1.4 obtained its clause (e).
- $\begin{pmatrix} 5 & 1 \\ -8 & -1 \end{pmatrix}$, with eigenvalues 3 and 1 and $(1, 1)$ entry 5 (Example (A corner outside the spectrum)). It is necessarily non-Hermitian, by Example (The diagonal lies inside the spectrum).

17.4.7.2 B. Practice

Exercise. B1: Interlacing in integers

Let

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

Compute the eigenvalues of $\mathbf{A}$ and of all three of its 2×2 principal submatrices, and exhibit the interlacing in each case. Hence say, for each corner, whether the interlacing is strict.

Solution. The eigenvalues of $\mathbf{A}$. The vector $(1, -1, 0)$ satisfies $\mathbf{A}(1, -1, 0) = (-1, 1, 0)$, so -1 is an eigenvalue. Dividing, $p_{\mathbf{A}}(x) = x^3 - 3x^2 - 9x - 5 = (x - 5)(x + 1)^2$, so

$$\lambda(\mathbf{A}) = (5, -1, -1).$$

(As a check, $\text{tr } \mathbf{A} = 3 = 5 - 1 - 1$, and expanding along the first row $\det \mathbf{A} = 0 - 1(3 - 4) + 2(2 - 0) = 5 = 5 \cdot (-1) \cdot (-1)$.)

The corners.

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

The first has trace 0 and determinant -1 , so its eigenvalues are 1 and -1 . The second has trace 3 and determinant -4 , so its characteristic polynomial is $x^2 - 3x - 4 = (x - 4)(x + 1)$ and its eigenvalues are 4 and -1 .

Interlacing.

$$I = \{1, 2\} : \quad 5 > 1 > -1 = -1 = -1,$$

$$I = \{1, 3\} : \quad 5 > 4 > -1 = -1 = -1,$$

$$I = \{2, 3\} : \quad 5 > 4 > -1 = -1 = -1.$$

No corner gives strict interlacing, and the reason is on the side of $\mathbf{A}$: whichever index is deleted, $\lambda_2(\mathbf{A}) = \lambda_3(\mathbf{A}) = -1$, so the chain $\lambda_2(\mathbf{A}) \geq \lambda_2(\mathbf{B}) \geq \lambda_3(\mathbf{A})$ squeezes $\lambda_2(\mathbf{B})$ to -1 . Strict interlacing would make the combined chain strictly decreasing, and so would force $\mathbf{A}$ to have three distinct eigenvalues. [Theorem 17.4.4](#) sees the same thing from the side of the corner. For $I = \{1, 2\}$ the border is $\mathbf{b} = (2, 2)$, orthogonal to $(1, -1)$, which spans $E_{-1}(\mathbf{B})$; so $P_{-1}\mathbf{b} = \mathbf{0}$, and part (b) predicts that -1 is an eigenvalue of $\mathbf{A}$. For $I = \{1, 3\}$ the border is $(a_{12}, a_{32}) = (1, 2)$, orthogonal to $(2, -1)$, which spans E_{-1} of $\begin{pmatrix} 0 & 2 \\ 2 & 3 \end{pmatrix}$, with the same conclusion.

Exercise. B2: Bounds without a characteristic polynomial

Let $\mathbf{A} \in M_4(\mathbb{R})$ be symmetric with

$$\mathbf{A} = \begin{pmatrix} 7 & * & * & * \\ * & 2 & * & * \\ * & * & -3 & * \\ * & * & * & 1 \end{pmatrix},$$

the entries marked $*$ being unknown.

- (a) What do [Corollary 17.4.3](#) and [Example \(The diagonal lies inside the spectrum\)](#) say about $\lambda_1(\mathbf{A})$ and $\lambda_4(\mathbf{A})$?

- (b) Suppose in addition that $\mathbf{A}_{\{1,2\},\{1,2\}} = \begin{pmatrix} 7 & 4 \\ 4 & 2 \end{pmatrix}$. Improve the lower bound on $\lambda_1(\mathbf{A})$, and give a lower bound on $\lambda_2(\mathbf{A})$.

Solution.

- (a) Each diagonal entry is a 1×1 principal submatrix, so $\lambda_1(\mathbf{A}) \geq \max\{7, 2, -3, 1\} = 7$ and $\lambda_4(\mathbf{A}) \leq \min\{7, 2, -3, 1\} = -3$. In particular $\mathbf{A}$ is indefinite whatever the unknown entries are.
- (b) The submatrix $\mathbf{B} = \begin{pmatrix} 7 & 4 \\ 4 & 2 \end{pmatrix}$ has trace 9 and determinant $14 - 16 = -2$, so its characteristic polynomial is $x^2 - 9x - 2$ and its eigenvalues are

$$\lambda_1(\mathbf{B}) = \frac{9+\sqrt{89}}{2} \approx 9.2170, \quad \lambda_2(\mathbf{B}) = \frac{9-\sqrt{89}}{2} \approx -0.2170.$$

Here $n = 4$ and $m = 2$, and the left-hand inequality $\lambda_i(\mathbf{A}) \geq \lambda_i(\mathbf{B})$ of [Corollary 17.4.3](#) at $i = 1, 2$ gives

$$\lambda_1(\mathbf{A}) \geq \frac{9+\sqrt{89}}{2} \approx 9.2170, \quad \lambda_2(\mathbf{A}) \geq \frac{9-\sqrt{89}}{2} \approx -0.2170.$$

The first improves the bound 7 from (a). The second is new: the diagonal entries alone say nothing about $\lambda_2(\mathbf{A})$. The right-hand inequalities $\lambda_i(\mathbf{B}) \geq \lambda_{i+2}(\mathbf{A})$ give $\lambda_3(\mathbf{A}) \leq 9.2170$, which is true but of little use, and $\lambda_4(\mathbf{A}) \leq -0.2170$, which is weaker than the bound $\lambda_4(\mathbf{A}) \leq -3$ of (a). Each principal submatrix contributes what it can, and one keeps the best of the bounds.

Exercise. B3: Deciding strictness

For each of the following bordered matrices $\mathbf{A} = \begin{pmatrix} \mathbf{B} & \mathbf{b} \\ \mathbf{b}^\top & c \end{pmatrix}$, decide without computing $p_{\mathbf{A}}$ whether the interlacing between $\mathbf{A}$ and $\mathbf{B}$ is strict, and justify with [Theorem 17.4.4](#).

- (a) $\mathbf{B} = \text{diag}(3, 1)$, $\mathbf{b} = (0, 1)$, $c = 0$.
- (b) $\mathbf{B} = \text{diag}(3, 1)$, $\mathbf{b} = (2, -1)$, $c = 5$.
- (c) $\mathbf{B} = \begin{pmatrix} 4 & 0 \\ 0 & 4 \end{pmatrix}$, $\mathbf{b} = (1, 2)$, $c = 0$.

Solution. In each case $\mathbf{B}$ is diagonal, so the standard basis vectors are an orthonormal eigenbasis and $P_\nu \mathbf{b}$ is just the part of $\mathbf{b}$ in the coordinates where the diagonal entry equals ν .

- (a) Not strict. The eigenvalues 3, 1 of $\mathbf{B}$ are distinct, but $P_3 \mathbf{b} = (0, 0) = \mathbf{0}$. By [Theorem 17.4.4](#) (b), 3 is an eigenvalue of $\mathbf{A}$, so one of the inequalities $\lambda_1(\mathbf{A}) \geq 3 \geq \lambda_2(\mathbf{A})$ is an equality. (Indeed $p_{\mathbf{A}}(x) = (x - 3)(x^2 - x - 1)$.)
- (b) Strict. The eigenvalues 3, 1 are distinct and $P_3 \mathbf{b} = (2, 0) \neq \mathbf{0}$, $P_1 \mathbf{b} = (0, -1) \neq \mathbf{0}$. Both clauses of [Theorem 17.4.4](#) (c) hold. (Indeed $p_{\mathbf{A}}(x) = (x - 2)(x^2 - 7x + 4)$, with roots $\frac{7+\sqrt{33}}{2} \approx 6.372$, 2, and $\frac{7-\sqrt{33}}{2} \approx 0.628$, and $6.372 > 3 > 2 > 1 > 0.628$.)
- (c) Not strict. $P_4 \mathbf{b} = \mathbf{b} = (1, 2) \neq \mathbf{0}$, so the second clause holds, but $\mathbf{B}$ has the repeated eigenvalue 4 and the first clause fails. By [Theorem 17.4.4](#) (a) with $m_4 = 2$, the number 4 is an eigenvalue of $\mathbf{A}$ of multiplicity exactly 1, and the middle of the chain reads $4 \geq \lambda_2(\mathbf{A}) \geq 4$. (Indeed $p_{\mathbf{A}}(x) = (x - 5)(x - 4)(x + 1)$.)

17.4.7.3 C. Going deeper

Exercise. C1: Rank under compression

Let $\mathbf{A} \in M_n(F)$ be Hermitian of rank r , and let $\mathbf{B} = \mathbf{A}_{I,I}$ with $|I| = n - m$.

- Prove that $\text{rank } \mathbf{B} \geq r - 2m$.
- Show that the bound is attained, by exhibiting for each n and each $m \leq n/2$ a Hermitian $\mathbf{A}$ of rank $2m$ with a principal submatrix of size $n - m$ and rank 0.

Solution.

- For a Hermitian matrix $\mathbf{M}$, diagonalizing by the spectral theorem gives $\text{rank } \mathbf{M} = n_+(\mathbf{M}) + n_-(\mathbf{M})$, the total number of nonzero eigenvalues. By [Corollary 17.4.7](#), $n_+(\mathbf{B}) \geq n_+(\mathbf{A}) - m$ and $n_-(\mathbf{B}) \geq n_-(\mathbf{A}) - m$. Adding,

$$\text{rank } \mathbf{B} = n_+(\mathbf{B}) + n_-(\mathbf{B}) \geq (n_+(\mathbf{A}) + n_-(\mathbf{A})) - 2m = r - 2m.$$

- Let $m \leq n/2$ and let $I = \{m + 1, m + 2, \dots, n\}$, of size $n - m$. Define $\mathbf{A}$ by $a_{i,m+i} = a_{m+i,i} = 1$ for $1 \leq i \leq m$, and all other entries 0. Then $\mathbf{A}$ is real symmetric. It is a direct sum of m copies of $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, acting on the coordinate pairs $\{i, m + i\}$, together with zeros on the remaining coordinates, so its eigenvalues are ± 1 each m times and 0 otherwise; hence $\text{rank } \mathbf{A} = 2m$ and $n_+(\mathbf{A}) = n_-(\mathbf{A}) = m$. Every nonzero entry of $\mathbf{A}$ has one index in $\{1, \dots, m\}$, which I omits, so $\mathbf{A}_{I,I} = \mathbf{O}$ and $\text{rank } \mathbf{A}_{I,I} = 0 = \text{rank } \mathbf{A} - 2m$. The bound in (a) is therefore sharp for every m .

Exercise. C2: The converse, in the smallest case

[Theorem 17.4.2](#) says that the spectrum of a corner interlaces the spectrum of the whole. This exercise asks whether every interlacing pattern occurs.

- Let $\lambda_1 > \mu_1 > \lambda_2$ be real numbers. Find $b \in \mathbb{R}$ and $c \in \mathbb{R}$ such that $\mathbf{A} = \begin{pmatrix} \mu_1 & b \\ b & c \end{pmatrix}$ has eigenvalues λ_1 and λ_2 .
- Deduce that for $n = 2$ the interlacing condition of [Theorem 17.4.2](#) is not merely necessary but also sufficient.
- Explain why the strict hypothesis $\lambda_1 > \mu_1 > \lambda_2$ may be relaxed to $\lambda_1 \geq \mu_1 \geq \lambda_2$.

Hint: match the trace and the determinant.

Solution.

- A real symmetric 2×2 matrix with eigenvalues λ_1, λ_2 must have $\text{tr } \mathbf{A} = \lambda_1 + \lambda_2$ and $\det \mathbf{A} = \lambda_1 \lambda_2$. The first forces

$$c = \lambda_1 + \lambda_2 - \mu_1,$$

and the second then forces $\mu_1 c - b^2 = \lambda_1 \lambda_2$, that is,

$$b^2 = \mu_1(\lambda_1 + \lambda_2 - \mu_1) - \lambda_1 \lambda_2 = -(\mu_1 - \lambda_1)(\mu_1 - \lambda_2) = (\lambda_1 - \mu_1)(\mu_1 - \lambda_2),$$

where the middle equality is the expansion $-(\mu_1^2 - (\lambda_1 + \lambda_2)\mu_1 + \lambda_1\lambda_2)$. The hypothesis $\lambda_1 > \mu_1 > \lambda_2$ makes both factors positive, so $b = \sqrt{(\lambda_1 - \mu_1)(\mu_1 - \lambda_2)}$ is a real number. With this b and c , the matrix $\mathbf{A}$ is real symmetric with the prescribed trace and determinant, hence with characteristic polynomial $x^2 - (\lambda_1 + \lambda_2)x + \lambda_1\lambda_2 = (x - \lambda_1)(x - \lambda_2)$, hence with eigenvalues λ_1 and λ_2 .

For example $\lambda_1 = 4$, $\mu_1 = 2$, $\lambda_2 = 1$ gives $c = 3$ and $b^2 = 2$, that is $\mathbf{A} = \begin{pmatrix} 2 & \sqrt{2} \\ \sqrt{2} & 3 \end{pmatrix}$, whose trace is 5 and determinant $6 - 2 = 4$, as required.

- (b) Necessity is [Theorem 17.4.2](#) with $n = 2$ and $I = \{1\}$. Sufficiency is part (a): given any $\lambda_1 > \mu_1 > \lambda_2$, a Hermitian $\mathbf{A}$ with spectrum $\{\lambda_1, \lambda_2\}$ and $\mathbf{A}_{\{1\},\{1\}} = (\mu_1)$ exists.
- (c) If $\mu_1 = \lambda_1$ or $\mu_1 = \lambda_2$ then the formula gives $b = 0$, and $\mathbf{A} = \text{diag}(\mu_1, \lambda_1 + \lambda_2 - \mu_1)$ is diagonal with the two entries λ_1 and λ_2 in one order or the other. So the construction still works, with a zero border; consistently with [Theorem 17.4.4](#), a zero border is exactly what makes the interlacing nonstrict.

Exercise. C3: How far apart can the extremes be

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

- (a) Prove that $\lambda_1(\mathbf{A}) - \lambda_n(\mathbf{A}) \geq \max_{i,j} |a_{ii} - a_{jj}|$.
- (b) Prove that $\lambda_1(\mathbf{A}) - \lambda_n(\mathbf{A}) \geq 2 \max_{i \neq j} |a_{ij}|$.

Hint for (b): look at a 2×2 principal submatrix.

Solution.

- (a) By [Example \(The diagonal lies inside the spectrum\)](#), $\lambda_n(\mathbf{A}) \leq a_{kk} \leq \lambda_1(\mathbf{A})$ for every k . So for any i, j , both $a_{ii} - a_{jj}$ and $a_{jj} - a_{ii}$ are at most $\lambda_1(\mathbf{A}) - \lambda_n(\mathbf{A})$, which gives the bound after taking the maximum over i, j .
- (b) Fix $i \neq j$ and let $I = \{i, j\}$, so that

$$\mathbf{B} = \mathbf{A}_{I,I} = \begin{pmatrix} a_{ii} & a_{ij} \\ \overline{a_{ij}} & a_{jj} \end{pmatrix}$$

is Hermitian of size 2. Its trace is $a_{ii} + a_{jj}$ and its determinant is $a_{ii}a_{jj} - |a_{ij}|^2$, so its eigenvalues are

$$\frac{a_{ii} + a_{jj}}{2} \pm \sqrt{\left(\frac{a_{ii} - a_{jj}}{2}\right)^2 + |a_{ij}|^2},$$

and their difference is $2\sqrt{((a_{ii} - a_{jj})/2)^2 + |a_{ij}|^2} \geq 2|a_{ij}|$. Now apply [Corollary 17.4.3](#) with $m = n - 2$: it gives $\lambda_1(\mathbf{A}) \geq \lambda_1(\mathbf{B})$ and $\lambda_2(\mathbf{B}) \geq \lambda_{2+(n-2)}(\mathbf{A}) = \lambda_n(\mathbf{A})$, so

$$\lambda_1(\mathbf{A}) - \lambda_n(\mathbf{A}) \geq \lambda_1(\mathbf{B}) - \lambda_2(\mathbf{B}) \geq 2|a_{ij}|.$$

Taking the maximum over the pairs $i \neq j$ finishes the proof. (For $n = 2$ the step with $m = 0$ is the identity $\mathbf{B} = \mathbf{A}$, and the conclusion is the computation itself.)