

17.2 The Min–Max Theorem

Section 1 ended with a description of $\lambda_k(\mathbf{A})$ that works and cannot be used: a maximum of the Rayleigh quotient over the orthogonal complement of $\mathbf{A}$'s own top eigenvectors. The defect is that the constraint set is built from the answer. This section repairs it by a move that looks like an extravagance and is exactly the right one: keep the maximum over a subspace, stop insisting that the subspace be the right one, and let it range over **all** subspaces of the appropriate dimension. The minimum of what comes out is again $\lambda_k(\mathbf{A})$, and no eigenvector appears anywhere in the statement. That is the Courant–Fischer theorem, and everything in the rest of the chapter is a consequence of it.

Throughout, F is $\mathbb{R}$ or $\mathbb{C}$ and every matrix is Hermitian. The eigenvalues of a Hermitian $\mathbf{A} \in M_n(F)$ are real (Theorem 12.2.1) and are **indexed decreasingly**,

$$\lambda_1(\mathbf{A}) \geq \lambda_2(\mathbf{A}) \geq \dots \geq \lambda_n(\mathbf{A}),$$

each repeated according to multiplicity, so that the list has exactly n entries. This is the book's convention, fixed in Chapter 13 §05, and it is used without exception below. Many sources index the theorem of this section the other way round; the remark after the theorem gives that form, once.

17.2.1 Two subspaces that have to meet

One dimension count does all the work in this chapter, and it is worth a label of its own, because four later proofs cite it.

Two subspaces of an n -dimensional space cannot both be large and disjoint: if their dimensions add to more than n , they share a nonzero vector.

Lemma 17.2.1. Large Subspaces Intersect

Let V be a vector space with $\dim V = n$ over any field, and let U and W be subspaces of V with

$$\dim U + \dim W > n.$$

Then $U \cap W \neq \{0\}$. More precisely, $\dim(U \cap W) \geq \dim U + \dim W - n$.

Proof. By Theorem 1.7.3,

$$\dim(U + W) = \dim U + \dim W - \dim(U \cap W).$$

Both U and W are finite-dimensional, being subspaces of the n -dimensional V (Theorem 1.6.7); $U + W$ is a subspace of V (Theorem 1.7.1), so the same theorem gives $\dim(U + W) \leq n$, and rearranging gives

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

which is > 0 by hypothesis. A subspace of positive dimension contains a nonzero vector. This proves the lemma. ■

Remark. The bound is sharp, and the hypothesis cannot be weakened to $\dim U + \dim W = n$: two distinct lines in $\mathbb{R}^2$ meet only at $\mathbf{0}$, and so do a line and a plane in $\mathbb{R}^3$ chosen in general position. What makes the count useful is that it needs nothing about the two subspaces beyond their dimensions — in every application below, one of them is a span of eigenvectors that we choose and the other is an arbitrary competitor we know nothing about.

17.2.2 The quotient on a subspace

Before quantifying over subspaces, we must know that the inner maximum and minimum exist. They do, and for an algebraic reason: restricting the quadratic form to a subspace produces a smaller Hermitian matrix, to which Proposition 17.1.1 applies verbatim.

Lemma 17.2.2. The Rayleigh Quotient on a Subspace

Let $\mathbf{A} \in M_n(F)$ be Hermitian and let $W \subseteq F^n$ be a subspace with $\dim W = m \geq 1$. Let $\mathbf{Q} \in M_{n \times m}(F)$ be a matrix whose columns form an orthonormal basis of W , and put $\mathbf{B} = \mathbf{Q}^* \mathbf{A} \mathbf{Q} \in M_m(F)$. Then $\mathbf{B}$ is Hermitian and

$$\{R_{\mathbf{A}}(\mathbf{x}) : \mathbf{0} \neq \mathbf{x} \in W\} = [\lambda_m(\mathbf{B}), \lambda_1(\mathbf{B})].$$

In particular $\max_{\mathbf{0} \neq \mathbf{x} \in W} R_{\mathbf{A}}(\mathbf{x})$ and $\min_{\mathbf{0} \neq \mathbf{x} \in W} R_{\mathbf{A}}(\mathbf{x})$ exist and are attained.

Idea. Coordinates on W . Writing a vector of W as $\mathbf{Q}\mathbf{c}$ turns $\langle \mathbf{A}\mathbf{x}, \mathbf{x} \rangle$ into $\langle \mathbf{B}\mathbf{c}, \mathbf{c} \rangle$ and, because the columns of $\mathbf{Q}$ are orthonormal, turns $\langle \mathbf{x}, \mathbf{x} \rangle$ into $\langle \mathbf{c}, \mathbf{c} \rangle$ with no distortion. So the Rayleigh quotient of $\mathbf{A}$ on W is the Rayleigh quotient of $\mathbf{B}$ on F^m , and Proposition 17.1.1 answers the question there.

Proof. Such a $\mathbf{Q}$ exists: W has an orthonormal basis by Theorem 11.2.2, and we take its members as columns. Orthonormality of the columns says exactly $\mathbf{Q}^* \mathbf{Q} = \mathbf{I}_m$. Then $\mathbf{B}^* = \mathbf{Q}^* \mathbf{A}^* \mathbf{Q} = \mathbf{Q}^* \mathbf{A} \mathbf{Q} = \mathbf{B}$, so $\mathbf{B}$ is Hermitian and Proposition 17.1.1 applies to it.

The map $\mathbf{c} \mapsto \mathbf{Q}\mathbf{c}$ sends F^m onto W , since the columns of $\mathbf{Q}$ span W , and it is injective, since $\mathbf{Q}\mathbf{c} = \mathbf{0}$ gives $\mathbf{c} = \mathbf{Q}^* \mathbf{Q}\mathbf{c} = \mathbf{0}$. So it is a bijection $F^m \rightarrow W$ carrying $\mathbf{0}$ to $\mathbf{0}$, hence a bijection from the nonzero vectors of F^m onto those of W . For $\mathbf{c} \neq \mathbf{0}$, writing the standard inner product as $\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{v}^* \mathbf{u}$,

$$\begin{aligned} \langle \mathbf{A}\mathbf{Q}\mathbf{c}, \mathbf{Q}\mathbf{c} \rangle &= \mathbf{c}^* \mathbf{Q}^* \mathbf{A} \mathbf{Q} \mathbf{c} = \langle \mathbf{B}\mathbf{c}, \mathbf{c} \rangle, \\ \langle \mathbf{Q}\mathbf{c}, \mathbf{Q}\mathbf{c} \rangle &= \mathbf{c}^* \mathbf{Q}^* \mathbf{Q} \mathbf{c} = \langle \mathbf{c}, \mathbf{c} \rangle, \end{aligned}$$

so $R_{\mathbf{A}}(\mathbf{Q}\mathbf{c}) = R_{\mathbf{B}}(\mathbf{c})$. Therefore the two value sets $\{R_{\mathbf{A}}(\mathbf{x}) : \mathbf{0} \neq \mathbf{x} \in W\}$ and $\{R_{\mathbf{B}}(\mathbf{c}) : \mathbf{0} \neq \mathbf{c} \in F^m\}$ are equal, and the second is $[\lambda_m(\mathbf{B}), \lambda_1(\mathbf{B})]$ with both ends attained, by Proposition 17.1.1. This proves the lemma. ■

Remark. **Which analysis was invoked: none.** The existence of a maximum of a function over an infinite set is normally an analytic fact — the set of unit vectors of W is compact, by fact (A3) of Chapter 16’s introduction, and $R_{\mathbf{A}}$ is continuous on it, so the extreme value theorem, fact (A4), applies. That route is available and correct. It is not needed. The spectral theorem, already proved in Chapter 12, produces the maximizing vector by name, and [Lemma 17.2.2](#) inherits that. Nothing in this section, or in §01, uses an item of Chapter 16’s imported-analysis list. The analysis has not vanished, only moved: the spectral theorem rests on the fundamental theorem of algebra (Theorem 6.5.1), whose proof in Chapter 6 §05 uses the extreme value theorem on a closed disc.

17.2.3 Every eigenvalue is a min–max

Theorem 17.2.3. Courant–Fischer Min–Max Theorem

Let F be $\mathbb{R}$ or $\mathbb{C}$, let $\mathbf{A} \in M_n(F)$ be Hermitian, and let $1 \leq k \leq n$. Then

$$\begin{aligned}\lambda_k(\mathbf{A}) &= \min_{\dim W = n-k+1} \max_{\mathbf{0} \neq \mathbf{x} \in W} R_{\mathbf{A}}(\mathbf{x}) \\ &= \max_{\dim W = k} \min_{\mathbf{0} \neq \mathbf{x} \in W} R_{\mathbf{A}}(\mathbf{x}),\end{aligned}$$

where in each line W runs over **all** subspaces of F^n of the stated dimension. All four extrema are attained.

Idea. Read the second line first, because it is the one with a slogan: λ_k is the best value you can guarantee on a k -dimensional subspace. On a k -dimensional subspace the quotient has a smallest value, which is what that subspace guarantees; the span of the top k eigenvectors guarantees λ_k , and no other k -dimensional subspace guarantees more.

The proof is three steps, and the first two have entirely different characters. **Step 1** exhibits a **witness** subspace and computes on it, giving one inequality. **Step 2** takes an **arbitrary** competitor subspace and shows it cannot beat the witness; the only thing known about a competitor is its dimension, so the only available tool is [Lemma 17.2.1](#), which forces the competitor to meet a span of eigenvectors from the other end of the list. On that shared nonzero vector the quotient is pinned, and the inequality follows. **Step 3** obtains the second line from the first by applying it to $-\mathbf{A}$, which reverses the order of the eigenvalues and turns every maximum into a minimum.

Proof. By Corollary 12.4.3 over $\mathbb{C}$, or Corollary 12.5.2 over $\mathbb{R}$, fix an orthonormal basis $(\mathbf{q}_1, \dots, \mathbf{q}_n)$ of F^n with $\mathbf{A}\mathbf{q}_i = \lambda_i(\mathbf{A})\mathbf{q}_i$ for every i . Abbreviate $\lambda_i = \lambda_i(\mathbf{A})$ and put

$$U_j = \text{span}(\mathbf{q}_1, \dots, \mathbf{q}_j), \quad L_j = \text{span}(\mathbf{q}_j, \dots, \mathbf{q}_n),$$

so that $\dim U_j = j$ and $\dim L_j = n-j+1$, the lists being orthonormal and hence independent. By [Lemma 17.2.2](#) the inner extrema below all exist.

Claim. For every j : $R_{\mathbf{A}}(\mathbf{x}) \geq \lambda_j$ for all $\mathbf{0} \neq \mathbf{x} \in U_j$, and $R_{\mathbf{A}}(\mathbf{x}) \leq \lambda_j$ for all $\mathbf{0} \neq \mathbf{x} \in L_j$. Both bounds are attained at $\mathbf{q}_j$.

Proof. Let $\mathbf{0} \neq \mathbf{x} \in U_j$ and write $c_i = \langle \mathbf{x}, \mathbf{q}_i \rangle$, so $\mathbf{x} = \sum_{i \leq j} c_i \mathbf{q}_i$ by Theorem 11.2.1 (a) and the definition of U_j . Then $\mathbf{A}\mathbf{x} = \sum_{i \leq j} \lambda_i c_i \mathbf{q}_i$, and Theorem 11.2.1 (b), (c) give

$$R_{\mathbf{A}}(\mathbf{x}) = \frac{\sum_{i \leq j} \lambda_i |c_i|^2}{\sum_{i \leq j} |c_i|^2},$$

a weighted average of $\lambda_1, \dots, \lambda_j$ with nonnegative weights summing to 1. Since the list is indexed decreasingly, the smallest of those eigenvalues is λ_j , so the average is $\geq \lambda_j$. The statement for L_j is Proposition 17.1.3: L_j is precisely the set of vectors orthogonal to $\mathbf{q}_1, \dots, \mathbf{q}_{j-1}$, and the maximum of $R_{\mathbf{A}}$ there is λ_j . In both cases $R_{\mathbf{A}}(\mathbf{q}_j) = \lambda_j$. ■

Step 1. The witness for the first line. For each subspace W of dimension $n - k + 1$ write $f(W) = \max_{\mathbf{0} \neq \mathbf{x} \in W} R_{\mathbf{A}}(\mathbf{x})$, which exists by Lemma 17.2.2. The outer minimum of f over all such W is not yet known to exist — there are infinitely many W — so we show that λ_k is a value of f and a lower bound for all its values; it is then the least value, which is what the minimum means. For the first half take $W = L_k$, of dimension $n - k + 1$: by the claim, $f(L_k) = \lambda_k$.

Step 2. Every competitor is at least as large. Let $W \subseteq F^n$ be any subspace with $\dim W = n - k + 1$. Then

$$\dim W + \dim U_k = (n - k + 1) + k = n + 1 > n,$$

so by Lemma 17.2.1 there is a vector $\mathbf{x} \neq \mathbf{0}$ lying in $W \cap U_k$. Being in U_k , it satisfies $R_{\mathbf{A}}(\mathbf{x}) \geq \lambda_k$ by the claim; being in W , it is one of the vectors the inner maximum ranges over. Therefore

$$\max_{\mathbf{0} \neq \mathbf{y} \in W} R_{\mathbf{A}}(\mathbf{y}) \geq R_{\mathbf{A}}(\mathbf{x}) \geq \lambda_k.$$

This holds for every competitor W , so $f(W) \geq \lambda_k$ for all of them. With Step 1, λ_k is a value of f and a lower bound for every value, so it is the least value: the minimum exists, equals λ_k , and is attained at $W = L_k$. This proves the first line.

Step 3. The second line, from the first applied to $-\mathbf{A}$. The matrix $-\mathbf{A}$ is Hermitian, its eigenvalues are the numbers $-\lambda_i$, and sorting those decreasingly reverses the list:

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

Also $R_{-\mathbf{A}}(\mathbf{x}) = -R_{\mathbf{A}}(\mathbf{x})$ for every $\mathbf{x} \neq \mathbf{0}$, directly from the definition. Since negation turns a maximum of a set of reals into the minimum of the negated set, and a minimum into a maximum,

$$\begin{aligned} \max_{\mathbf{0} \neq \mathbf{x} \in W} R_{-\mathbf{A}}(\mathbf{x}) &= - \min_{\mathbf{0} \neq \mathbf{x} \in W} R_{\mathbf{A}}(\mathbf{x}), \\ \min_{\dim W=m} (-f(W)) &= - \max_{\dim W=m} f(W) \end{aligned}$$

for any real-valued f on subspaces, where each side exists exactly when the other does. Apply the first line, already proved, to $-\mathbf{A}$ with index j :

$$\begin{aligned} -\lambda_{n+1-j}(\mathbf{A}) &= \lambda_j(-\mathbf{A}) = \min_{\dim W=n-j+1} \max_{\mathbf{0} \neq \mathbf{x} \in W} R_{-\mathbf{A}}(\mathbf{x}) \\ &= - \max_{\dim W=n-j+1} \min_{\mathbf{0} \neq \mathbf{x} \in W} R_{\mathbf{A}}(\mathbf{x}). \end{aligned}$$

Now substitute $j = n + 1 - k$, so that $n - j + 1 = k$ and $n + 1 - j = k$. Multiplying by -1 gives

$$\lambda_k(\mathbf{A}) = \max_{\dim W=k} \min_{\mathbf{0} \neq \mathbf{x} \in W} R_{\mathbf{A}}(\mathbf{x}),$$

with the maximum attained, the attainment transported from Step 1 by the same negation; explicitly it occurs at $W = U_k$, which is the subspace L_{n+1-k} of the eigenbasis of $-\mathbf{A}$ taken in its own decreasing order. This proves the theorem. ■

The theorem removes the eigenvectors from the description of $\lambda_k(\mathbf{A})$. What is left mentions only $\mathbf{A}$, through $R_{\mathbf{A}}$, and the subspaces of F^n , which are the same for every matrix. Two Hermitian matrices of the same size can therefore be compared term by term, and that is the whole reason this chapter exists.

Remark. The increasing-index form. If the eigenvalues are listed *increasingly* as $\mu_1 \leq \mu_2 \leq \dots \leq \mu_n$, then $\mu_j = \lambda_{n+1-j}(\mathbf{A})$, and substituting $k = n + 1 - j$ in [Theorem 17.2.3](#) turns it into

$$\begin{aligned} \mu_j &= \min_{\dim W=j} \max_{\mathbf{0} \neq \mathbf{x} \in W} R_{\mathbf{A}}(\mathbf{x}) \\ &= \max_{\dim W=n-j+1} \min_{\mathbf{0} \neq \mathbf{x} \in W} R_{\mathbf{A}}(\mathbf{x}). \end{aligned}$$

The pairing of min max with max min is unchanged; only the dimensions move. Many books state the theorem this way, so a reader comparing sources should check which convention is in force before comparing formulas. This book uses the decreasing one everywhere, and this is the only place the other appears.

Warning. Both the order of the two extrema and the dimension attached to each are load-bearing. Neither

$$\max_{\dim W=n-k+1} \min_{\mathbf{0} \neq \mathbf{x} \in W} R_{\mathbf{A}}(\mathbf{x}) \quad \text{nor} \quad \min_{\dim W=k} \max_{\mathbf{0} \neq \mathbf{x} \in W} R_{\mathbf{A}}(\mathbf{x})$$

equals $\lambda_k(\mathbf{A})$. Each of them is, by [Theorem 17.2.3](#) itself, equal to $\lambda_{n+1-k}(\mathbf{A})$: the first is the second line of the theorem at index $n - k + 1$, and the second is the first line at the index j with $n - j + 1 = k$, namely $j = n - k + 1$. So swapping either the order or the dimension returns the eigenvalue counted from the *other end* of the list.

For a witness take $\mathbf{A} = \text{diag}(2, 0)$ and $k = 1$, so that $\lambda_1(\mathbf{A}) = 2$. The first false expression maximizes over the single subspace $W = F^2$ the minimum of $R_{\mathbf{A}}$, which is 0. The second minimizes over lines the maximum on the line, which is the value of $R_{\mathbf{A}}$ on that line, and the smallest such value is again 0. Both return $\lambda_2(\mathbf{A}) = 0$ instead of 2. The mistake is invisible when $n = 2k - 1$, since then $n + 1 - k = k$, and this is exactly why it survives in the wild.

Quick check. In the first line of [Theorem 17.2.3](#), why is the dimension $n - k + 1$ rather than k ? Check your reason against $k = 1$.

Solution. The competitor W must be forced to meet $U_k = \text{span}(\mathbf{q}_1, \dots, \mathbf{q}_k)$, which has dimension k ; by Lemma 17.2.1 that needs $\dim W > n - k$, and $n - k + 1$ is the smallest such dimension. It must not be larger either: competitors of dimension $n - k + 2$ would be forced to meet U_{k-1} as well, and by the theorem itself the min-max over them is $\lambda_{k-1}(\mathbf{A})$, the wrong eigenvalue.

At $k = 1$ the dimension is $n - 1 + 1 = n$, so the only competitor is $W = F^n$, and the statement reads $\lambda_1(\mathbf{A}) = \max_{\mathbf{x} \neq \mathbf{0}} R_{\mathbf{A}}(\mathbf{x})$, which is Proposition 17.1.1. With the dimension $k = 1$ instead, we would be computing min over lines of the single value of $R_{\mathbf{A}}$ on the line, which is $\lambda_n(\mathbf{A})$ — the wrong end.

Example. Both forms on one matrix

Let

$$\mathbf{A} = \begin{pmatrix} 2 & 2 & 0 \\ 2 & 3 & 2 \\ 0 & 2 & 4 \end{pmatrix} \in M_3(\mathbb{R}),$$

whose eigenvalues are 6, 3, 0, with orthonormal eigenvectors

$$\mathbf{q}_1 = \frac{1}{3}(1, 2, 2), \quad \mathbf{q}_2 = \frac{1}{3}(2, 1, -2), \quad \mathbf{q}_3 = \frac{1}{3}(2, -2, 1).$$

Verify both lines of Theorem 17.2.3 for $k = 2$ by computing the witness and two competitors in each case.

Solution. Throughout we use Lemma 17.2.2: for a subspace W with orthonormal basis the columns of $\mathbf{Q}$, the values of $R_{\mathbf{A}}$ on W fill the interval between the smallest and largest eigenvalues of the 2×2 matrix $\mathbf{B} = \mathbf{Q}^\top \mathbf{A} \mathbf{Q}$.

Max-min form, $\dim W = k = 2$. The witness is $U_2 = \text{span}(\mathbf{q}_1, \mathbf{q}_2)$; since $\mathbf{q}_1, \mathbf{q}_2$ are eigenvectors, $\mathbf{B} = \text{diag}(6, 3)$ and the minimum of $R_{\mathbf{A}}$ on U_2 is $3 = \lambda_2(\mathbf{A})$. Two competitors:

- $W = \text{span}(\mathbf{e}_1, \mathbf{e}_2)$, where $\mathbf{B}$ is the top-left 2×2 block $\begin{pmatrix} 2 & 2 \\ 2 & 3 \end{pmatrix}$, with eigenvalues $\frac{1}{2}(5 \pm \sqrt{17})$. Its minimum is $\frac{1}{2}(5 - \sqrt{17}) \approx 0.438$, well below 3.
- $W = \text{span}(\mathbf{e}_1, \mathbf{e}_3)$, where $\mathbf{B} = \text{diag}(2, 4)$ because $a_{13} = 0$. Its minimum is 2, again below 3.

Min-max form, $\dim W = n - k + 1 = 2$ as well, since $n = 3$. The witness is $L_2 = \text{span}(\mathbf{q}_2, \mathbf{q}_3)$, where $\mathbf{B} = \text{diag}(3, 0)$ and the maximum is $3 = \lambda_2(\mathbf{A})$. The same two competitors now go the other way:

- on $\text{span}(\mathbf{e}_1, \mathbf{e}_2)$ the maximum is $\frac{1}{2}(5 + \sqrt{17}) \approx 4.562 > 3$;
- on $\text{span}(\mathbf{e}_1, \mathbf{e}_3)$ the maximum is $4 > 3$.

So each competitor overshoots in the min-max form and undershoots in the max-min form, and the two witnesses pin $\lambda_2(\mathbf{A}) = 3$ from both sides. Notice that the competitor $\text{span}(\mathbf{e}_1, \mathbf{e}_2)$ is the coordinate plane obtained by deleting the third row and column, and its two eigenvalues $0.438 \dots$ and $4.562 \dots$ are caught between consecutive eigenvalues $0 \leq 0.438 \leq 3 \leq 4.562 \leq 6$ of $\mathbf{A}$. That is no accident; it is the interlacing theorem of §04.

17.2.4 The two debts this pays at once

Corollary 17.2.4. The Loewner Order Compares Every Eigenvalue

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

$$\lambda_i(\mathbf{A}) \geq \lambda_i(\mathbf{B}) \quad \text{for every } i = 1, \dots, n.$$

Idea. $\mathbf{A} \succeq \mathbf{B}$ says the Rayleigh quotient of $\mathbf{A}$ sits above that of $\mathbf{B}$ at every vector. Courant–Fischer builds each λ_i out of the Rayleigh quotient by a minimum and then a maximum, and both operations preserve “everywhere above”. So the domination passes straight through to every eigenvalue.

Proof. By Definition 13.5.1, $\mathbf{A} \succeq \mathbf{B}$ means $\mathbf{A} - \mathbf{B} \succeq \mathbf{O}$, so $\langle (\mathbf{A} - \mathbf{B})\mathbf{x}, \mathbf{x} \rangle \geq 0$ for every $\mathbf{x}$. Dividing by $\langle \mathbf{x}, \mathbf{x} \rangle > 0$ for $\mathbf{x} \neq \mathbf{0}$,

$$R_{\mathbf{A}}(\mathbf{x}) \geq R_{\mathbf{B}}(\mathbf{x}) \quad \text{for every } \mathbf{x} \neq \mathbf{0}.$$

Fix i and let W be any subspace of F^n with $\dim W = i$. Let $\mathbf{x}_1 \in W \setminus \{\mathbf{0}\}$ attain $\min_{\mathbf{0} \neq \mathbf{x} \in W} R_{\mathbf{A}}(\mathbf{x})$, which exists by Lemma 17.2.2. Then

$$\min_{\mathbf{0} \neq \mathbf{x} \in W} R_{\mathbf{A}}(\mathbf{x}) = R_{\mathbf{A}}(\mathbf{x}_1) \geq R_{\mathbf{B}}(\mathbf{x}_1) \geq \min_{\mathbf{0} \neq \mathbf{x} \in W} R_{\mathbf{B}}(\mathbf{x}).$$

So the function $W \mapsto \min_{\mathbf{0} \neq \mathbf{x} \in W} R_{\mathbf{A}}(\mathbf{x})$ dominates the corresponding function for $\mathbf{B}$ at every W , and therefore so do their maxima over the subspaces of dimension i . By Theorem 17.2.3 applied to $\mathbf{A}$ and to $\mathbf{B}$, those maxima are $\lambda_i(\mathbf{A})$ and $\lambda_i(\mathbf{B})$. This proves the corollary. ■

Example. Checking the corollary on a pair

Let $\mathbf{A} = \begin{pmatrix} 4 & 1 \\ 1 & 3 \end{pmatrix}$ and $\mathbf{B} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$. Then $\mathbf{A} - \mathbf{B} = \text{diag}(2, 1)$, which is positive semidefinite, so $\mathbf{A} \succeq \mathbf{B}$. The eigenvalues of $\mathbf{B}$ are 3 and 1, from trace 4 and determinant 3; those of $\mathbf{A}$ are $\frac{7 \pm \sqrt{5}}{2} \approx 4.618, 2.382$, from trace 7 and determinant 11. So $\lambda_1(\mathbf{A}) \approx 4.618 \geq 3 = \lambda_1(\mathbf{B})$ and $\lambda_2(\mathbf{A}) \approx 2.382 \geq 1 = \lambda_2(\mathbf{B})$, as the corollary demands. Note that the two gaps, $\frac{1 + \sqrt{5}}{2} \approx 1.618$ and $\frac{5 - \sqrt{5}}{2} \approx 1.382$, are different: the corollary orders the eigenvalues, and says nothing about moving them by the same amount.

This settles an account opened in Chapter 13 §05. Discussing Theorem 13.5.3 (b), which proved only $\lambda_1(\mathbf{A}) \geq \lambda_1(\mathbf{B})$ and $\lambda_n(\mathbf{A}) \geq \lambda_n(\mathbf{B})$, that section said: “The full statement is that $\lambda_i(\mathbf{A}) \geq \lambda_i(\mathbf{B})$ for **every** i , not only for $i = 1$ and $i = n$; but the intermediate eigenvalues have no description as a plain maximum over the whole unit sphere, and the min–max description that does the job is the Courant–Fischer theorem of Chapter 17.” Corollary 17.2.4 is that full statement, and the diagnosis was exact: what was missing was not an extra trick but a description of λ_i that two different matrices could be measured against. The proof above is short, and all of its weight is carried by Theorem 17.2.3.

Corollary 17.2.5. Courant–Fischer for a Self-Adjoint Operator

Let V be a finite-dimensional inner product space over $F = \mathbb{R}$ or $F = \mathbb{C}$, with $\dim V = n \geq 1$, and let $T \in \mathcal{L}(V)$ be self-adjoint, with eigenvalues $\lambda_1(T) \geq \cdots \geq \lambda_n(T)$ repeated according to multiplicity. Then for every k with $1 \leq k \leq n$,

$$\begin{aligned}\lambda_k(T) &= \min_{\dim U = n-k+1} \max_{0 \neq v \in U} \frac{\langle Tv, v \rangle}{\langle v, v \rangle} \\ &= \max_{\dim U = k} \min_{0 \neq v \in U} \frac{\langle Tv, v \rangle}{\langle v, v \rangle},\end{aligned}$$

where U runs over the subspaces of V of the stated dimension.

Idea. Transport the theorem along coordinates. An orthonormal basis of eigenvectors turns V into F^n without changing any inner product, turns T into a real diagonal matrix, and matches subspaces with subspaces of the same dimension. There is nothing to prove beyond checking that each of those three things is true.

Proof. By Proposition 11.6.7 (c) a self-adjoint operator is normal, so Theorem 12.4.2 over $\mathbb{C}$, and Theorem 12.5.1 over $\mathbb{R}$, give an orthonormal basis $\mathcal{E} = (e_1, \dots, e_n)$ of V consisting of eigenvectors of T ; relabel it so that $Te_i = \lambda_i(T)e_i$ with the eigenvalues in decreasing order, which is possible because they are real by Theorem 12.2.1. In particular the eigenvalue list of T , repeated according to multiplicity, has exactly n entries.

Let $\Phi: V \rightarrow F^n$, $\Phi(v) = [v]_{\mathcal{E}}$, be the coordinate map, an isomorphism by Corollary 2.5.5. By Theorem 11.2.1 (b), $\langle \Phi u, \Phi v \rangle = \langle u, v \rangle$ for all $u, v \in V$, so Φ is an isometric isomorphism in the sense of Corollary 11.2.4. Put $\mathbf{A} = [T]_{\mathcal{E}}^{\mathcal{E}} = \text{diag}(\lambda_1(T), \dots, \lambda_n(T))$, a real diagonal matrix, hence Hermitian, with $\lambda_i(\mathbf{A}) = \lambda_i(T)$ for every i . By Theorem 4.1.1, $\Phi(Tv) = \mathbf{A}\Phi(v)$.

Consequently, for $v \neq 0$,

$$\frac{\langle Tv, v \rangle}{\langle v, v \rangle} = \frac{\langle \mathbf{A}\Phi v, \Phi v \rangle}{\langle \Phi v, \Phi v \rangle} = R_{\mathbf{A}}(\Phi v),$$

and $\Phi v \neq \mathbf{0}$ because Φ is injective. Finally $U \mapsto \Phi(U)$ is a bijection from the subspaces of V to the subspaces of F^n preserving dimension, by Theorem 2.5.2, and it matches the nonzero vectors of U with those of $\Phi(U)$. So each of the two displayed expressions equals the corresponding expression for $\mathbf{A}$ in Theorem 17.2.3, which is $\lambda_k(\mathbf{A}) = \lambda_k(T)$. This proves the corollary. ■

This is the statement Chapter 12 §02 promised when it wrote that “the maximization picture returns in Chapter 17, where the Rayleigh quotient and the Courant–Fischer theorem describe every eigenvalue of a self-adjoint operator by optimization”. The picture has returned and the description is complete: every eigenvalue, not merely the extreme two, and stated for an operator on an abstract inner product space, which is the setting Chapter 12 worked in. The remark in §01 records what it cost — nothing. Chapter 12 declined the maximization route because it would have needed the extreme value theorem to know that a maximum exists; here the spectral theorem supplies the maximizing vectors explicitly, so the min–max description is a *consequence* of Chapter 12 rather than an alternative to it, and no analysis has been imported.

What has been bought is comparison. Two Hermitian matrices are now comparable eigenvalue by eigenvalue whenever their Rayleigh quotients are comparable vector by vector, and [Corollary 17.2.4](#) is the first instance. Section 3 takes the same idea and applies it to $\mathbf{A}$ and $\mathbf{A} + \mathbf{E}$, where the quotients differ by at most $\|\mathbf{E}\|_2$, and gets the perturbation bound that Chapter 16 could not state.

17.2.5 Exercises

17.2.5.1 A. Check your understanding

Exercise. A1

- (a) State both lines of [Theorem 17.2.3](#), being careful about which dimension goes with which line.
- (b) State [Lemma 17.2.1](#), and say at which point of the proof of [Theorem 17.2.3](#) it is used and what it is used on.
- (c) Determine whether the following is correct, and justify your answer:

$$\lambda_k(\mathbf{A}) = \min_{\dim W=k} \max_{\mathbf{0} \neq \mathbf{x} \in W} R_{\mathbf{A}}(\mathbf{x}).$$

- (d) Write the first line of [Theorem 17.2.3](#) in the increasing-index convention $\mu_1 \leq \cdots \leq \mu_n$.
- (e) Explain in one or two sentences why [Theorem 13.5.3](#) (b) could not be extended to intermediate eigenvalues by the argument Chapter 13 used, and what [Theorem 17.2.3](#) supplies instead.

Solution.

- (a) For Hermitian $\mathbf{A} \in M_n(F)$ and $1 \leq k \leq n$: $\lambda_k(\mathbf{A}) = \min \max$ over subspaces of dimension $n - k + 1$ of the Rayleigh quotient, and $= \max \min$ over subspaces of dimension k . The large dimension goes with the minimum of maxima; the small one with the maximum of minima.
- (b) If U, W are subspaces of an n -dimensional space with $\dim U + \dim W > n$, then $U \cap W \neq \{0\}$. It is used in Step 2, applied to an arbitrary competitor W of dimension $n - k + 1$ and the span U_k of the top k eigenvectors, whose dimensions add to $n + 1$.
- (c) Incorrect. By [Theorem 17.2.3](#) itself, that expression is the first line evaluated at the index j with $n - j + 1 = k$, so it equals $\lambda_{n-k+1}(\mathbf{A})$. For $\mathbf{A} = \text{diag}(2, 0)$ and $k = 1$ it returns 0, not $\lambda_1(\mathbf{A}) = 2$.
- (d) $\mu_j = \min_{\dim W=j} \max_{\mathbf{0} \neq \mathbf{x} \in W} R_{\mathbf{A}}(\mathbf{x})$, obtained by setting $k = n + 1 - j$.
- (e) Chapter 13's argument evaluated the quadratic form at a vector that is extremal for the *other* matrix, which works only when the extremal vector is characterized by a plain maximum over the whole sphere — true for λ_1 and λ_n and for no other index, since the range of $R_{\mathbf{A}}$ is an interval ([Proposition 17.1.1](#)). [Theorem 17.2.3](#) supplies a description of every λ_i as an optimization over a family of subspaces that does not depend on the matrix, so the same subspace can be fed to both matrices; that is exactly what the proof of [Corollary 17.2.4](#) does.

17.2.5.2 B. Practice

Exercise. B1: Both forms on a diagonal matrix

Let $\mathbf{A} = \text{diag}(5, 2, -1) \in M_3(\mathbb{R})$ and $k = 2$.

- (a) Give the witness subspace for the max-min form, compute the minimum of $R_{\mathbf{A}}$ on it, and check the value against [Theorem 17.2.3](#).
- (b) Do the same for the min-max form.
- (c) Put $\mathbf{u} = \frac{1}{\sqrt{2}}(0, 1, 1)$. Compute the minimum of $R_{\mathbf{A}}$ over the nonzero vectors of $\text{span}(\mathbf{e}_1, \mathbf{u})$, and over those of $\text{span}(\mathbf{e}_2, \mathbf{e}_3)$, and verify that neither subspace beats the witness in (a).

Solution.

- (a) The eigenvalues are $\lambda_1 = 5$, $\lambda_2 = 2$, $\lambda_3 = -1$ with eigenvectors $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$. The witness is $U_2 = \text{span}(\mathbf{e}_1, \mathbf{e}_2)$, of dimension $k = 2$. There $\mathbf{B} = \text{diag}(5, 2)$, so by [Lemma 17.2.2](#) the values of $R_{\mathbf{A}}$ fill $[2, 5]$ and the minimum is $2 = \lambda_2(\mathbf{A})$.
- (b) Here $n - k + 1 = 2$ as well, and the witness is $L_2 = \text{span}(\mathbf{e}_2, \mathbf{e}_3)$, where $\mathbf{B} = \text{diag}(2, -1)$; the values fill $[-1, 2]$ and the maximum is $2 = \lambda_2(\mathbf{A})$.
- (c) For $W = \text{span}(\mathbf{e}_1, \mathbf{u})$ the two spanning vectors are orthonormal, and

$$\mathbf{B} = \begin{pmatrix} 5 & 0 \\ 0 & \frac{1}{2}(2 - 1) \end{pmatrix} = \text{diag}(5, \frac{1}{2}),$$

the off-diagonal entry vanishing because $\mathbf{A}\mathbf{e}_1 = 5\mathbf{e}_1$ is orthogonal to $(0, 1, 1)$. So the minimum is $\frac{1}{2} < 2$. For $W = \text{span}(\mathbf{e}_2, \mathbf{e}_3)$ we get $\mathbf{B} = \text{diag}(2, -1)$ and minimum $-1 < 2$. Neither competitor reaches $\lambda_2(\mathbf{A}) = 2$, as [Theorem 17.2.3](#) requires.

Exercise. B2: Shifting by a multiple of the identity

Let $\mathbf{A} \in M_n(F)$ be Hermitian and let $c \in \mathbb{R}$. Prove that $\lambda_k(\mathbf{A} + c\mathbf{I}) = \lambda_k(\mathbf{A}) + c$ for every k , using the first line of [Theorem 17.2.3](#) for both matrices.

Solution. The matrix $\mathbf{A} + c\mathbf{I}$ is Hermitian, since c is real. For every $\mathbf{x} \neq \mathbf{0}$,

$$R_{\mathbf{A}+c\mathbf{I}}(\mathbf{x}) = \frac{\langle \mathbf{A}\mathbf{x}, \mathbf{x} \rangle + c\langle \mathbf{x}, \mathbf{x} \rangle}{\langle \mathbf{x}, \mathbf{x} \rangle} = R_{\mathbf{A}}(\mathbf{x}) + c.$$

Adding a constant to a real-valued function adds it to every maximum and every minimum of that function over any nonempty set. So for each subspace W ,

$$\max_{\mathbf{0} \neq \mathbf{x} \in W} R_{\mathbf{A}+c\mathbf{I}}(\mathbf{x}) = \left(\max_{\mathbf{0} \neq \mathbf{x} \in W} R_{\mathbf{A}}(\mathbf{x}) \right) + c,$$

and taking the minimum over all W of dimension $n - k + 1$ adds c again. By [Theorem 17.2.3](#) applied to $\mathbf{A} + c\mathbf{I}$ on the left and to $\mathbf{A}$ on the right, $\lambda_k(\mathbf{A} + c\mathbf{I}) = \lambda_k(\mathbf{A}) + c$.

Exercise. B3: Comparing two matrices

Let

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

Verify that $\mathbf{A} \succeq \mathbf{B}$, compute both eigenvalue lists, and check [Corollary 17.2.4](#). Check whether $\lambda_1(\mathbf{A}) - \lambda_1(\mathbf{B}) \leq \lambda_1(\mathbf{A} - \mathbf{B})$ holds for this pair. (Whether it holds for every pair is a question for Section 3.)

Solution. $\mathbf{A} - \mathbf{B} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$ is symmetric with leading principal minors $1 > 0$ and $\det = 2 - 1 = 1 > 0$, so $\mathbf{A} - \mathbf{B} \succ \mathbf{O}$ by Theorem 13.1.5; in particular $\mathbf{A} \succeq \mathbf{B}$.

The eigenvalues: $\mathbf{A}$ has $p_{\mathbf{A}}(x) = (x - 3)^2 - 1$, so $\lambda_1(\mathbf{A}) = 4$, $\lambda_2(\mathbf{A}) = 2$; and $\mathbf{B}$ is diagonal with $\lambda_1(\mathbf{B}) = 2$, $\lambda_2(\mathbf{B}) = 1$. Indeed $4 \geq 2$ and $2 \geq 1$, as [Corollary 17.2.4](#) requires.

For the last question, $\mathbf{A} - \mathbf{B}$ has trace 3 and determinant 1, so its eigenvalues are $\frac{1}{2}(3 \pm \sqrt{5})$ and $\lambda_1(\mathbf{A} - \mathbf{B}) = \frac{1}{2}(3 + \sqrt{5}) \approx 2.618$. The inequality asks whether $4 - 2 \leq 2.618$, which holds. (It is not an accident: it is the simplest of the perturbation inequalities of §03, which carry Weyl's name. Note that [Corollary 17.2.4](#) alone gives no upper bound on the gap $\lambda_i(\mathbf{A}) - \lambda_i(\mathbf{B})$.)

17.2.5.3 C. Going deeper
Exercise. C1: The definite generalized eigenvalue problem

Let $\mathbf{A}, \mathbf{B} \in M_n(F)$ be Hermitian with $\mathbf{A} \succ \mathbf{O}$, and recall Definition 13.4.1: λ is a generalized eigenvalue of the pair $(\mathbf{A}, \mathbf{B})$ when $\mathbf{B}\mathbf{x} = \lambda\mathbf{A}\mathbf{x}$ for some $\mathbf{x} \neq \mathbf{0}$. Define the **generalized Rayleigh quotient**

$$R_{\mathbf{A}, \mathbf{B}}(\mathbf{x}) = \frac{\langle \mathbf{B}\mathbf{x}, \mathbf{x} \rangle}{\langle \mathbf{A}\mathbf{x}, \mathbf{x} \rangle} \quad (\mathbf{x} \neq \mathbf{0}).$$

- Prove that $R_{\mathbf{A}, \mathbf{B}}$ is well defined and real-valued.
- Put $\mathbf{C} = \mathbf{A}^{-1/2}\mathbf{B}\mathbf{A}^{-1/2}$. Prove that $\mathbf{C}$ is Hermitian and that $\det(\mathbf{B} - \lambda\mathbf{A}) = \det(\mathbf{A})\det(\mathbf{C} - \lambda\mathbf{I})$, so that the generalized eigenvalues of $(\mathbf{A}, \mathbf{B})$, with multiplicity, are exactly the eigenvalues of $\mathbf{C}$. Write them $d_1 \geq \dots \geq d_n$.
- Prove that for every k ,

$$d_k = \min_{\dim W = n-k+1} \max_{\mathbf{0} \neq \mathbf{x} \in W} R_{\mathbf{A}, \mathbf{B}}(\mathbf{x}).$$

Hint for (c): the substitution $\mathbf{x} = \mathbf{A}^{-1/2}\mathbf{y}$.

Solution.

- By Theorem 13.2.1 and Theorem 13.1.5, $\mathbf{A} \succ \mathbf{O}$ has a positive definite square root $\mathbf{A}^{1/2}$, which is invertible; write $\mathbf{A}^{-1/2}$ for its inverse, again Hermitian and positive definite. For $\mathbf{x} \neq \mathbf{0}$, $\langle \mathbf{A}\mathbf{x}, \mathbf{x} \rangle > 0$ by definiteness, so the denominator is a nonzero real; the numerator is real by Proposition 11.6.7 (a). So the quotient is a well-defined real number.

(b) $\mathbf{C}^* = \mathbf{A}^{-1/2} \mathbf{B}^* \mathbf{A}^{-1/2} = \mathbf{C}$. Since $\mathbf{A}^{1/2} \mathbf{C} \mathbf{A}^{1/2} = \mathbf{B}$ and $\mathbf{A}^{1/2} \mathbf{A}^{1/2} = \mathbf{A}$,

$$\mathbf{B} - \lambda \mathbf{A} = \mathbf{A}^{1/2} (\mathbf{C} - \lambda \mathbf{I}) \mathbf{A}^{1/2},$$

so taking determinants and using Theorem 7.5.4 gives

$$\det(\mathbf{B} - \lambda \mathbf{A}) = \det(\mathbf{A}^{1/2})^2 \det(\mathbf{C} - \lambda \mathbf{I}) = \det(\mathbf{A}) \det(\mathbf{C} - \lambda \mathbf{I}).$$

Since $\det \mathbf{A} \neq 0$, the two polynomials in λ have the same roots with the same multiplicities. By Theorem 13.4.2 (b) the generalized eigenvalues of $(\mathbf{A}, \mathbf{B})$ are the roots of $\det(\mathbf{B} - \lambda \mathbf{A})$, so they are the eigenvalues of $\mathbf{C}$, all real because $\mathbf{C}$ is Hermitian.

(c) Put $\mathbf{S} = \mathbf{A}^{-1/2}$, which is Hermitian and invertible. For $\mathbf{y} \neq \mathbf{0}$ set $\mathbf{x} = \mathbf{S}\mathbf{y} \neq \mathbf{0}$. Then

$$\langle \mathbf{B}\mathbf{x}, \mathbf{x} \rangle = \langle \mathbf{SBSy}, \mathbf{y} \rangle = \langle \mathbf{C}\mathbf{y}, \mathbf{y} \rangle, \quad \langle \mathbf{A}\mathbf{x}, \mathbf{x} \rangle = \langle \mathbf{SASy}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{y} \rangle,$$

using $\mathbf{SAS} = \mathbf{A}^{-1/2} \mathbf{A} \mathbf{A}^{-1/2} = \mathbf{I}$. Hence $R_{\mathbf{A}, \mathbf{B}}(\mathbf{S}\mathbf{y}) = R_{\mathbf{C}}(\mathbf{y})$. The map $\mathbf{y} \mapsto \mathbf{S}\mathbf{y}$ is an isomorphism of F^n , so $W \mapsto \mathbf{S}^{-1}W$ is a bijection of the subspaces of dimension $n - k + 1$ onto themselves, matching nonzero vectors with nonzero vectors. Therefore

$$\max_{\mathbf{0} \neq \mathbf{x} \in W} R_{\mathbf{A}, \mathbf{B}}(\mathbf{x}) = \max_{\mathbf{0} \neq \mathbf{y} \in \mathbf{S}^{-1}W} R_{\mathbf{C}}(\mathbf{y}),$$

and minimizing both sides over all W of dimension $n - k + 1$ gives, by Theorem 17.2.3 applied to $\mathbf{C}$, the value $\lambda_k(\mathbf{C}) = d_k$.

Exercise. C2: How much weaker the eigenvalue comparison is

- Give Hermitian $\mathbf{A}, \mathbf{B} \in M_2(\mathbb{R})$ with $\lambda_i(\mathbf{A}) \geq \lambda_i(\mathbf{B})$ for $i = 1, 2$ but $\mathbf{A} \not\geq \mathbf{B}$. So the converse of Corollary 17.2.4 is false.
- Prove that if $\mathbf{A}$ and $\mathbf{B}$ are Hermitian with $\lambda_i(\mathbf{A}) \geq \lambda_i(\mathbf{B})$ for every i and $\text{tr } \mathbf{A} = \text{tr } \mathbf{B}$, then $\lambda_i(\mathbf{A}) = \lambda_i(\mathbf{B})$ for every i .
- Prove that if $\mathbf{A} \succeq \mathbf{B}$ and $\text{tr } \mathbf{A} = \text{tr } \mathbf{B}$, then $\mathbf{A} = \mathbf{B}$.

Solution.

- Take $\mathbf{A} = \text{diag}(2, 0)$ and $\mathbf{B} = \text{diag}(0, 1)$. Then $\lambda_1(\mathbf{A}) = 2 \geq 1 = \lambda_1(\mathbf{B})$ and $\lambda_2(\mathbf{A}) = 0 \geq 0 = \lambda_2(\mathbf{B})$. But $\mathbf{A} - \mathbf{B} = \text{diag}(2, -1)$ has the eigenvalue $-1 < 0$, so $\mathbf{A} - \mathbf{B} \not\geq \mathbf{0}$ by Theorem 13.1.4 and $\mathbf{A} \not\geq \mathbf{B}$. Ordering the two spectra says nothing about how the eigenvectors are placed relative to each other, and the Loewner order sees both.
- By Theorem 9.3.3 the trace is the sum of the eigenvalues with multiplicity, so

$$0 = \text{tr } \mathbf{A} - \text{tr } \mathbf{B} = \sum_{i=1}^n (\lambda_i(\mathbf{A}) - \lambda_i(\mathbf{B})).$$

Every summand is ≥ 0 by hypothesis, and a sum of nonnegative reals is zero only if each is zero. Hence $\lambda_i(\mathbf{A}) = \lambda_i(\mathbf{B})$ for every i .

- Put $\mathbf{C} = \mathbf{A} - \mathbf{B}$, Hermitian with $\mathbf{C} \succeq \mathbf{0}$, so every eigenvalue of $\mathbf{C}$ is ≥ 0 by Theorem 13.1.4. Also $\text{tr } \mathbf{C} = \text{tr } \mathbf{A} - \text{tr } \mathbf{B} = 0$, which by Theorem 9.3.3 is the sum of those eigenvalues; a sum of nonnegative reals vanishing forces each to vanish, so every eigenvalue of $\mathbf{C}$ is 0. By Corollary 12.4.3 (or Corollary 12.5.2), $\mathbf{C} = \mathbf{U}\mathbf{D}\mathbf{U}^*$ with $\mathbf{D}$ the

diagonal matrix of eigenvalues, which is $\mathbf{O}$; hence $\mathbf{C} = \mathbf{O}$ and $\mathbf{A} = \mathbf{B}$. (One may also finish with (b) and [Corollary 17.2.4](#), which give the equality of the two spectra but not yet of the matrices; the trace argument on $\mathbf{C}$ is what closes the gap.)

Exercise. C3: Does enlarging the field change the answer?

Let $\mathbf{A} \in M_n(\mathbb{R})$ be symmetric. It may be read as an element of $M_n(\mathbb{C})$, where it is Hermitian, and then [Theorem 17.2.3](#) minimizes over a much larger collection of subspaces.

- (a) Prove that the eigenvalue lists of $\mathbf{A}$ over $\mathbb{R}$ and over $\mathbb{C}$, with multiplicity, are the same.
- (b) Deduce that for every k ,

$$\min_{\substack{W \subseteq \mathbb{R}^n \\ \dim W = n-k+1}} \max_{\mathbf{0} \neq \mathbf{x} \in W} R_{\mathbf{A}}(\mathbf{x}) = \min_{\substack{W \subseteq \mathbb{C}^n \\ \dim W = n-k+1}} \max_{\mathbf{0} \neq \mathbf{x} \in W} R_{\mathbf{A}}(\mathbf{x}).$$

- (c) Show that the second minimum really does range over strictly more subspaces, by exhibiting a one-dimensional subspace of $\mathbb{C}^2$ that contains no nonzero real vector.

Solution.

- (a) By [Corollary 12.5.2](#) there is an orthogonal $\mathbf{Q} \in M_n(\mathbb{R})$ and a real diagonal $\mathbf{D}$ with $\mathbf{A} = \mathbf{Q}\mathbf{D}\mathbf{Q}^\top$, and the diagonal of $\mathbf{D}$ is the eigenvalue list of $\mathbf{A}$ over $\mathbb{R}$ with multiplicity. Read over $\mathbb{C}$, the same $\mathbf{Q}$ is unitary, since $\mathbf{Q}^* = \overline{\mathbf{Q}}^\top = \mathbf{Q}^\top = \mathbf{Q}^{-1}$, and the same equation $\mathbf{A} = \mathbf{Q}\mathbf{D}\mathbf{Q}^*$ holds. So $\mathbf{A}$ is unitarily similar over $\mathbb{C}$ to the same $\mathbf{D}$, and its complex eigenvalue list with multiplicity is again the diagonal of $\mathbf{D}$.
- (b) Apply [Theorem 17.2.3](#) twice: once with $F = \mathbb{R}$, where the left-hand side equals $\lambda_k(\mathbf{A})$ computed over $\mathbb{R}$, and once with $F = \mathbb{C}$, where the right-hand side equals $\lambda_k(\mathbf{A})$ computed over $\mathbb{C}$. By (a) these are the same number.
- (c) Take $W = \text{span}_{\mathbb{C}}\{(1, i)\} \subseteq \mathbb{C}^2$. A nonzero element of W is $c(1, i) = (c, ci)$ with $c \neq 0$; for it to be real we would need $c \in \mathbb{R}$ and $ci \in \mathbb{R}$, and $c \neq 0$ makes the second impossible. So W is not the complex span of any real subspace.

The lesson is worth stating. Enlarging the field adds competitors to the minimization, which can only push the minimum **down**, and adds vectors to each inner maximization, which can only push those maxima **up**. The two effects cancel exactly, because both the witness subspace and the bound of Step 2 in the proof of [Theorem 17.2.3](#) survive the enlargement unchanged: the witness is spanned by real eigenvectors, and the dimension count of [Lemma 17.2.1](#) holds over any field.