

17.5 Adding a Rank-One Matrix

Section 4 deleted a row and the matching column and found that the spectrum of the smaller matrix is pinned between the eigenvalues of the larger one. This section makes the opposite kind of change: it leaves the size alone and adds the simplest nonzero Hermitian matrix there is, a positive multiple of $\mathbf{v}\mathbf{v}^*$. The answer has the same shape — the new eigenvalues interlace the old ones — but here we can say more than that they interlace. Under one extra hypothesis the new eigenvalues are the roots of a single explicit scalar equation in one variable, and the graph of that function shows at a glance where every one of them sits.

Throughout this section $F = \mathbb{R}$ **or** $F = \mathbb{C}$, **the matrix** $\mathbf{A} \in M_n(F)$ **is Hermitian, and its eigenvalues are indexed decreasingly**, $\lambda_1(\mathbf{A}) \geq \dots \geq \lambda_n(\mathbf{A})$, as everywhere in this chapter. We fix $\mathbf{v} \in F^n$ and a real number τ , and study

$$\tilde{\mathbf{A}} := \mathbf{A} + \tau\mathbf{v}\mathbf{v}^*.$$

The added term is Hermitian, since $(\mathbf{v}\mathbf{v}^*)^* = \mathbf{v}\mathbf{v}^*$ and τ is real, so $\tilde{\mathbf{A}}$ is Hermitian too and its eigenvalues are indexed the same way. The added term has rank 1 when $\tau \neq 0$ and $\mathbf{v} \neq \mathbf{0}$, and is the zero matrix otherwise. For $\tau > 0$ it is positive semidefinite, because

$$\langle \tau\mathbf{v}\mathbf{v}^*\mathbf{x}, \mathbf{x} \rangle = \tau\mathbf{x}^*\mathbf{v}\mathbf{v}^*\mathbf{x} = \tau|\mathbf{v}^*\mathbf{x}|^2 \geq 0$$

for every $\mathbf{x}$, which is clause (P2) of Definition 13.1.1.

17.5.1 One bump, every eigenvalue up

Adding a positive semidefinite matrix can only raise the quadratic form, so no eigenvalue can go down. The content of the theorem is the ceiling: however large τ is, the i -th new eigenvalue cannot climb past the $(i-1)$ -st old one. All the room created by the bump is spent at the top.

Theorem 17.5.1. Interlacing for a Rank-One Update

Let $\mathbf{A} \in M_n(F)$ be Hermitian, let $\mathbf{v} \in F^n$, let $\tau \geq 0$, and put $\tilde{\mathbf{A}} = \mathbf{A} + \tau\mathbf{v}\mathbf{v}^*$. Then

$$\lambda_i(\tilde{\mathbf{A}}) \geq \lambda_i(\mathbf{A}) \geq \lambda_{i+1}(\tilde{\mathbf{A}}) \quad (1 \leq i \leq n-1),$$

and $\lambda_n(\tilde{\mathbf{A}}) \geq \lambda_n(\mathbf{A})$. Written as one chain,

$$\begin{aligned} \lambda_1(\tilde{\mathbf{A}}) &\geq \lambda_1(\mathbf{A}) \geq \lambda_2(\tilde{\mathbf{A}}) \geq \lambda_2(\mathbf{A}) \\ &\geq \dots \geq \lambda_n(\tilde{\mathbf{A}}) \geq \lambda_n(\mathbf{A}). \end{aligned}$$

Idea. The two halves have different characters and different sources. The left-hand inequalities say the eigenvalues go **up**, and that is a statement about the quadratic form: on every vector, $\tilde{\mathbf{A}}$ scores at least what $\mathbf{A}$ scores, so every max–min it competes in can only improve. The right-hand inequalities say the eigenvalues go up **by at most one index**, and that has nothing to do with the sign of τ : it is the rank of the perturbation doing the work, and Corollary 17.3.8 has already isolated exactly that.

Proof. If $\tau = 0$ or $\mathbf{v} = \mathbf{0}$ then $\tilde{\mathbf{A}} = \mathbf{A}$ and every inequality is an equality, so assume $\tau > 0$ and $\mathbf{v} \neq \mathbf{0}$.

For the inequalities $\lambda_i(\tilde{\mathbf{A}}) \geq \lambda_i(\mathbf{A})$, fix i and let $\mathbf{x} \neq \mathbf{0}$. By the computation above,

$$R_{\tilde{\mathbf{A}}}(\mathbf{x}) = R_{\mathbf{A}}(\mathbf{x}) + \frac{\tau |\mathbf{v}^* \mathbf{x}|^2}{\|\mathbf{x}\|^2} \geq R_{\mathbf{A}}(\mathbf{x}),$$

using Definition 17.1.1 and $\tau > 0$. Multiplying through by $\|\mathbf{x}\|^2$, this says $\langle \tilde{\mathbf{A}} \mathbf{x}, \mathbf{x} \rangle \geq \langle \mathbf{A} \mathbf{x}, \mathbf{x} \rangle$ for every $\mathbf{x}$, that is, $\tilde{\mathbf{A}} \succeq \mathbf{A}$ in the Loewner order of Definition 13.5.1. So Corollary 17.2.4 gives $\lambda_i(\tilde{\mathbf{A}}) \geq \lambda_i(\mathbf{A})$ for every i .

For the inequalities $\lambda_i(\mathbf{A}) \geq \lambda_{i+1}(\tilde{\mathbf{A}})$, note that $\text{rank}(\tau \mathbf{v} \mathbf{v}^*) = 1$, because every column of $\mathbf{v} \mathbf{v}^*$ is a multiple of $\mathbf{v} \neq \mathbf{0}$ and at least one of them is nonzero. So Corollary 17.3.8 applies with $\mathbf{B} = \tau \mathbf{v} \mathbf{v}^*$ and $r = 1$, and its first inequality reads

$$\lambda_{i+1}(\mathbf{A} + \tau \mathbf{v} \mathbf{v}^*) \leq \lambda_i(\mathbf{A})$$

for $1 \leq i \leq n - 1$. This proves the theorem. ■

Two remarks on the proof. The first half used nothing about the rank: it is Corollary 17.2.4, the general fact that $\tilde{\mathbf{A}} \succeq \mathbf{A}$ forces $\lambda_i(\tilde{\mathbf{A}}) \geq \lambda_i(\mathbf{A})$ for every i , specialized to a rank-one increment. The second half used nothing about the sign: the eigenvalues of $\mathbf{A} + \mathbf{B}$ and of $\mathbf{A}$ are never more than $\text{rank } \mathbf{B}$ indices apart. Interlacing is what happens when both restrictions apply at once.

The secular equation of the next subsection gives a sharper, **strict** version of the theorem — it locates each new eigenvalue exactly, strictly between consecutive old ones — but only under extra hypotheses (a diagonal $\mathbf{A}$ with distinct entries, and a $\mathbf{v}$ with no zero coordinate), so it is not a second proof of the general statement. The other short route to the general theorem is Cauchy interlacing (Theorem 17.4.2) applied to a suitable bordered matrix.

Example. A bump on a diagonal matrix

Let $\mathbf{A} = \text{diag}(4, 1)$, $\mathbf{v} = (1, 2)$, and $\tau = 1$. Compute $\tilde{\mathbf{A}}$ and its eigenvalues, and check Theorem 17.5.1.

Solution. Here $\mathbf{v} \mathbf{v}^T = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$, so

$$\tilde{\mathbf{A}} = \begin{pmatrix} 4 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} = \begin{pmatrix} 5 & 2 \\ 2 & 5 \end{pmatrix}.$$

Its trace is 10 and its determinant is $25 - 4 = 21$, so its characteristic polynomial is $x^2 - 10x + 21 = (x - 7)(x - 3)$ and $\lambda_1(\tilde{\mathbf{A}}) = 7$, $\lambda_2(\tilde{\mathbf{A}}) = 3$. The chain of the theorem reads

$$7 \geq 4 \geq 3 \geq 1,$$

and all four inequalities are strict. Note the accounting: the two eigenvalues rose by 3 and by 2, a total of $5 = \|\mathbf{v}\|^2$. That is no coincidence, and the last result of this section says why.

Quick check. The matrix $\mathbf{A} = \text{diag}(9, 5, 2)$ is bumped by $\tau \mathbf{v} \mathbf{v}^*$ with $\tau > 0$. Without knowing $\mathbf{v}$ or τ , what is the largest possible value of $\lambda_3(\mathbf{A} + \tau \mathbf{v} \mathbf{v}^*)$, and what is the smallest?

Solution. By Theorem 17.5.1, $\lambda_2(\mathbf{A}) \geq \lambda_3(\tilde{\mathbf{A}}) \geq \lambda_3(\mathbf{A})$, that is, $5 \geq \lambda_3(\tilde{\mathbf{A}}) \geq 2$. Both ends are reached: $\mathbf{v} = \mathbf{e}_3$ with τ large gives $\tilde{\mathbf{A}} = \text{diag}(9, 5, 2 + \tau)$, whose third eigenvalue is 5 as soon as $\tau \geq 3$; and $\mathbf{v} = \mathbf{e}_1$ gives $\tilde{\mathbf{A}} = \text{diag}(9 + \tau, 5, 2)$, whose third eigenvalue is 2. So the smallest eigenvalue can be pushed up as far as the *second* smallest of $\mathbf{A}$ and no further, no matter how big τ is.

17.5.2 Where the new eigenvalues actually are

Interlacing brackets each new eigenvalue in an interval. To pin it down we need an equation, and there is a cheap one: an eigenvalue of $\tilde{\mathbf{A}}$ is a number λ making $\tilde{\mathbf{A}} - \lambda \mathbf{I}$ singular, and $\tilde{\mathbf{A}} - \lambda \mathbf{I} = (\mathbf{A} - \lambda \mathbf{I}) + \tau \mathbf{v} \mathbf{v}^*$ is a rank-one update of a matrix whose determinant we know. Chapter 8 computed the determinant of exactly such an update.

Divide the characteristic polynomial of the bumped matrix by that of the original one; what is left is a sum of n simple fractions.

Definition 17.5.1. Secular function

Let $\mathbf{A} = \text{diag}(d_1, \dots, d_n)$ with $d_1, \dots, d_n \in \mathbb{R}$, let $\mathbf{v} \in F^n$, and let $\tau \in \mathbb{R}$. The **secular function** of the update $\mathbf{A} + \tau \mathbf{v} \mathbf{v}^*$ is

$$f(\lambda) := 1 + \tau \sum_{i=1}^n \frac{|v_i|^2}{d_i - \lambda},$$

defined for every real λ **other than** $d_1, \dots, d_n$. The equation $f(\lambda) = 0$ is the **secular equation**.

In words: f is the constant 1 plus one simple pole at each diagonal entry of $\mathbf{A}$, weighted by τ times the squared modulus of the matching coordinate of $\mathbf{v}$. The name is inherited from celestial mechanics, where equations of this shape governed slow (“secular”) drifts in planetary orbits.

Theorem 17.5.2. Eigenvalues from the Secular Equation

Let $\mathbf{A} = \text{diag}(d_1, \dots, d_n)$ with

$$d_1 > d_2 > \dots > d_n$$

strictly decreasing and real, let $\tau > 0$, and let $\mathbf{v} \in F^n$ have **every** coordinate nonzero. Write $\tilde{\mathbf{A}} = \mathbf{A} + \tau \mathbf{v} \mathbf{v}^*$ and let f be its secular function. Then:

- (a) no d_i is an eigenvalue of $\tilde{\mathbf{A}}$;
- (b) a real number $\lambda \notin \{d_1, \dots, d_n\}$ is an eigenvalue of $\tilde{\mathbf{A}}$ if and only if $f(\lambda) = 0$;
- (c) f has exactly n roots $\mu_1 > \dots > \mu_n$, one in each of the intervals

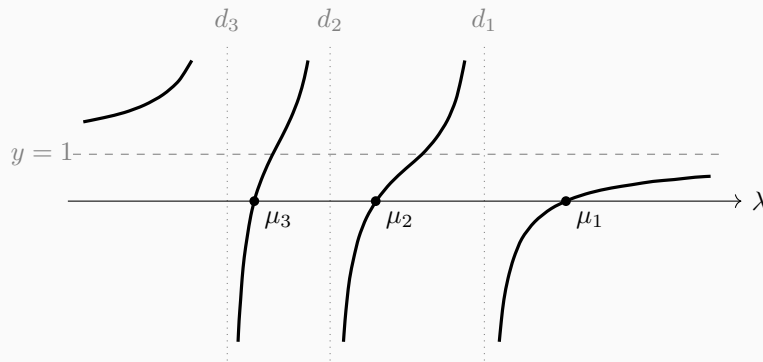
$$(d_1, \infty), (d_2, d_1), \dots, (d_n, d_{n-1}),$$

and none below d_n ; these are the eigenvalues of $\tilde{\mathbf{A}}$, each simple, and

$$\mu_1 > d_1 > \mu_2 > d_2 > \cdots > \mu_n > d_n;$$

(d) for each j , the vector $(\mathbf{A} - \mu_j \mathbf{I})^{-1} \mathbf{v}$ is an eigenvector of $\tilde{\mathbf{A}}$ for μ_j .

Idea. Part (b) is one line of Chapter 8 once (a) has cleared the denominators out of the way. Part (c) is a picture. Each term $|v_i|^2/(d_i - \lambda)$ increases with λ on any interval that avoids d_i , so on each of the $n + 1$ open intervals into which the poles cut the line f climbs steadily from left to right. Just to the left of each pole f shoots up to $+\infty$; just to the right it starts again from $-\infty$; and far out in either direction it settles at 1. Draw that, and the roots count themselves. The proof turns the picture into a count that needs no limits at all: the n eigenvalues turn out to be distinct, there are exactly n intervals where a root can live, and each holds at most one.



The dotted verticals are the poles, at the old eigenvalues $d_3 < d_2 < d_1$; the dashed horizontal is $y = 1$, the level f approaches at both ends. On the branch left of d_3 every term of the sum is positive, so the curve stays above 1 and never meets the axis. Each of the other three branches runs from $-\infty$ to $+\infty$ or from $-\infty$ to 1, increasing throughout, so each crosses the axis exactly once. Three branches, three roots, one strictly inside each gap and one above the top pole.

Proof. (a) Suppose $\tilde{\mathbf{A}}\mathbf{x} = d_i\mathbf{x}$ for some $\mathbf{x} \neq \mathbf{0}$. Rearranging, $(\mathbf{A} - d_i\mathbf{I})\mathbf{x} = -\tau(\mathbf{v}^*\mathbf{x})\mathbf{v}$. The matrix $\mathbf{A} - d_i\mathbf{I}$ is diagonal with i -th diagonal entry 0, so the i -th coordinate of the left-hand side is 0; the i -th coordinate of the right-hand side is $-\tau(\mathbf{v}^*\mathbf{x})v_i$. Since $\tau \neq 0$ and $v_i \neq 0$, this forces $\mathbf{v}^*\mathbf{x} = 0$, and then $(\mathbf{A} - d_i\mathbf{I})\mathbf{x} = \mathbf{0}$. The diagonal entries d_j are distinct, so the only diagonal entry of $\mathbf{A} - d_i\mathbf{I}$ that vanishes is the i -th, and therefore $\mathbf{x} = c\mathbf{e}_i$ for some scalar c . But then $0 = \mathbf{v}^*\mathbf{x} = c\bar{v}_i$, and $v_i \neq 0$ gives $c = 0$, so $\mathbf{x} = \mathbf{0}$ — a contradiction.

(b) Let $\lambda \notin \{d_1, \dots, d_n\}$. Then $\mathbf{A} - \lambda\mathbf{I}$ is diagonal with nonzero diagonal entries, hence invertible, with

$$(\mathbf{A} - \lambda\mathbf{I})^{-1} = \text{diag}\left(\frac{1}{d_1 - \lambda}, \dots, \frac{1}{d_n - \lambda}\right).$$

Apply the determinant statement of Corollary 8.3.6 — itself a 1×1 case of Theorem 8.3.2 — to the invertible matrix $\mathbf{A} - \lambda\mathbf{I}$, with the column $\tau\mathbf{v}$ and the row $\mathbf{v}^*$ — that is, with $\tau\mathbf{v}$

as its $\mathbf{u}$ and $\bar{\mathbf{v}}$ in place of its $\mathbf{v}$, so that its $\mathbf{v}^\top$ is our $\mathbf{v}^*$:

$$\begin{aligned}\det(\tilde{\mathbf{A}} - \lambda\mathbf{I}) &= \det((\mathbf{A} - \lambda\mathbf{I}) + (\tau\mathbf{v})\mathbf{v}^*) \\ &= \det(\mathbf{A} - \lambda\mathbf{I}) (1 + \tau\mathbf{v}^*(\mathbf{A} - \lambda\mathbf{I})^{-1}\mathbf{v}).\end{aligned}$$

The scalar in the second factor is $\sum_i |v_i|^2 / (d_i - \lambda)$, since $\mathbf{v}^*\mathbf{D}\mathbf{v} = \sum_i \bar{v}_i d_{ii} v_i$ for diagonal $\mathbf{D}$. So

$$\det(\tilde{\mathbf{A}} - \lambda\mathbf{I}) = \left(\prod_{i=1}^n (d_i - \lambda) \right) f(\lambda), \quad (*)$$

and the product is nonzero. Hence $\det(\tilde{\mathbf{A}} - \lambda\mathbf{I}) = 0$ if and only if $f(\lambda) = 0$. Finally $\det(\lambda\mathbf{I} - \tilde{\mathbf{A}}) = (-1)^n \det(\tilde{\mathbf{A}} - \lambda\mathbf{I})$, since negating all n rows of a matrix multiplies its determinant by $(-1)^n$, by n applications of Theorem 7.5.1 (b); so the two determinants vanish together, and by Theorem 7.10.4 (a) this happens exactly when λ is an eigenvalue of $\tilde{\mathbf{A}}$.

(c) Let J be one of the $n+1$ open intervals into which $d_1, \dots, d_n$ cut the real line, and let $\lambda < \lambda'$ lie in J . For each i , the numbers $d_i - \lambda$ and $d_i - \lambda'$ are nonzero and have the same sign, because $d_i \notin J$; hence

$$\frac{1}{d_i - \lambda'} - \frac{1}{d_i - \lambda} = \frac{\lambda' - \lambda}{(d_i - \lambda)(d_i - \lambda')} > 0.$$

Multiplying by $\tau|v_i|^2 > 0$ and summing over i gives $f(\lambda') > f(\lambda)$. So f is **strictly increasing** on each such interval, and in particular has at most one root in each.

Next, every eigenspace of $\tilde{\mathbf{A}}$ is one-dimensional. Let $\tilde{\mathbf{A}}\mathbf{x} = \mu\mathbf{x}$. By (a), $\mu \notin \{d_1, \dots, d_n\}$, so $\mathbf{A} - \mu\mathbf{I}$ is invertible, and rearranging as in (a) gives $(\mathbf{A} - \mu\mathbf{I})\mathbf{x} = -\tau(\mathbf{v}^*\mathbf{x})\mathbf{v}$, that is,

$$\mathbf{x} = -\tau(\mathbf{v}^*\mathbf{x})(\mathbf{A} - \mu\mathbf{I})^{-1}\mathbf{v}.$$

So $\ker(\tilde{\mathbf{A}} - \mu\mathbf{I}) \subseteq \text{span}((\mathbf{A} - \mu\mathbf{I})^{-1}\mathbf{v})$, a space of dimension 1. Now write $\tilde{\mathbf{A}} = \mathbf{U}\mathbf{D}\mathbf{U}^*$ with $\mathbf{U}$ unitary and $\mathbf{D}$ diagonal carrying the eigenvalue list of $\tilde{\mathbf{A}}$ with multiplicity (Corollary 12.4.3, or Corollary 12.5.2 over $\mathbb{R}$). Since $\mathbf{U}$ is invertible, $\ker(\tilde{\mathbf{A}} - \mu\mathbf{I}) = \mathbf{U}\ker(\mathbf{D} - \mu\mathbf{I})$, and $\dim \ker(\mathbf{D} - \mu\mathbf{I})$ is the number of diagonal entries of $\mathbf{D}$ equal to μ . That number is therefore at most 1: the n eigenvalues of $\tilde{\mathbf{A}}$ are **distinct**, and each is simple.

Now count. Call the eigenvalues $\mu_1 > \dots > \mu_n$. By (b) each is a root of f , and by (a) none equals a d_i , so each lies in one of the $n+1$ open intervals. None lies in $(-\infty, d_n)$: there every $d_i - \lambda$ is positive, so every term of the sum is positive and $f(\lambda) > 1$. That leaves n intervals, each containing at most one root of f , to accommodate n distinct roots — so each of the intervals $(d_1, \infty), (d_2, d_1), \dots, (d_n, d_{n-1})$ contains exactly one of them, and since they are listed in decreasing order, $\mu_1 \in (d_1, \infty)$ and $\mu_{i+1} \in (d_{i+1}, d_i)$. Conversely every root of f is an eigenvalue by (b), so f has no roots besides these n . Reading the intervals in decreasing order gives the displayed chain.

(d) Fix j and put $\mathbf{x} = (\mathbf{A} - \mu_j\mathbf{I})^{-1}\mathbf{v}$, which is defined because $\mu_j \notin \{d_1, \dots, d_n\}$ and nonzero because $\mathbf{v} \neq \mathbf{0}$. As in (b),

$$\tau\mathbf{v}^*(\mathbf{A} - \mu_j\mathbf{I})^{-1}\mathbf{v} = \tau \sum_i \frac{|v_i|^2}{d_i - \mu_j} = f(\mu_j) - 1 = -1.$$

Therefore

$$\begin{aligned}(\tilde{\mathbf{A}} - \mu_j\mathbf{I})\mathbf{x} &= (\mathbf{A} - \mu_j\mathbf{I})\mathbf{x} + \tau\mathbf{v}(\mathbf{v}^*\mathbf{x}) \\ &= \mathbf{v} + \mathbf{v}(\tau\mathbf{v}^*(\mathbf{A} - \mu_j\mathbf{I})^{-1}\mathbf{v}) = \mathbf{v} - \mathbf{v} = \mathbf{0}.\end{aligned}$$

This proves the theorem. ■

Identity (*) is worth keeping on its own: it says that the characteristic polynomial of the bumped matrix is the characteristic polynomial of the original one multiplied by the secular function. The poles of f cancel the roots of the product exactly, and what survives is a polynomial of the same degree with its roots shifted one gap to the right.

Example. A three-by-three update, found from its secular equation

Let

$$\mathbf{A} = \text{diag}(7, 1, -3), \quad \mathbf{v} = (2, 1, 2), \quad \tau = 1.$$

Write down $\tilde{\mathbf{A}} = \mathbf{A} + \mathbf{v}\mathbf{v}^\top$ and its secular function, find all three eigenvalues from the secular equation, and confirm them independently.

Solution. The matrix $\mathbf{v}\mathbf{v}^\top$ has (i, j) entry $v_i v_j$, so

$$\tilde{\mathbf{A}} = \begin{pmatrix} 7 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -3 \end{pmatrix} + \begin{pmatrix} 4 & 2 & 4 \\ 2 & 1 & 2 \\ 4 & 2 & 4 \end{pmatrix} = \begin{pmatrix} 11 & 2 & 4 \\ 2 & 2 & 2 \\ 4 & 2 & 1 \end{pmatrix}.$$

The hypotheses of [Theorem 17.5.2](#) hold: $7 > 1 > -3$ are distinct, $\tau = 1 > 0$, and no coordinate of $\mathbf{v}$ vanishes.

The secular function. With $|v_i|^2 = 4, 1, 4$,

$$f(\lambda) = 1 + \frac{4}{7-\lambda} + \frac{1}{1-\lambda} + \frac{4}{-3-\lambda}.$$

The roots. Multiply through by $(7-\lambda)(1-\lambda)(-3-\lambda)$, which by (*) turns f into $\det(\tilde{\mathbf{A}} - \lambda\mathbf{I})$:

$$\begin{aligned} & (7-\lambda)(1-\lambda)(-3-\lambda) + 4(1-\lambda)(-3-\lambda) \\ & + (7-\lambda)(-3-\lambda) + 4(7-\lambda)(1-\lambda) \\ & = -(\lambda-13)(\lambda-2)(\lambda+1). \end{aligned}$$

So the eigenvalues are $\mu_1 = 13$, $\mu_2 = 2$, $\mu_3 = -1$. They sit where the theorem promised:

$$13 > 7 > 2 > 1 > -1 > -3.$$

It is quicker to check the three roots than to expand the product. At $\lambda = 13$,

$$f(13) = 1 - \frac{4}{6} - \frac{1}{12} - \frac{4}{16} = \frac{12-8-1-3}{12} = 0;$$

at $\lambda = 2$, $f(2) = 1 + \frac{4}{5} - 1 - \frac{4}{5} = 0$; at $\lambda = -1$, $f(-1) = 1 + \frac{1}{2} + \frac{1}{2} - 2 = 0$.

An independent confirmation. Part (d) of the theorem produces eigenvectors without any further elimination: $(\mathbf{A} - \mu\mathbf{I})^{-1}\mathbf{v}$ is diagonal applied to $\mathbf{v}$. For $\mu = 13$, $\mathbf{A} - 13\mathbf{I} = \text{diag}(-6, -12, -16)$ and

$$(\mathbf{A} - 13\mathbf{I})^{-1}\mathbf{v} = \left(-\frac{1}{3}, -\frac{1}{12}, -\frac{1}{8}\right) \sim (8, 2, 3)$$

after scaling by -24 . Multiplying out, $\tilde{\mathbf{A}}(8, 2, 3) = (88 + 4 + 12, 16 + 4 + 6, 32 + 4 + 3) = (104, 26, 39) = 13(8, 2, 3)$. The same recipe gives $(2, -5, -2)$ for $\mu = 2$ and $(1, 2, -4)$ for $\mu = -1$, and both check by one matrix–vector product. Finally the trace: $11 + 2 + 1 = 14 = 13 + 2 - 1$, as Corollary 12.1.4 requires.

Warning. Both extra hypotheses of Theorem 17.5.2 earn their place, and the sign of τ reverses the picture.

A zero coordinate deflates. Take $\mathbf{A} = \text{diag}(2, 1, 0)$ and $\mathbf{v} = (1, 0, 1)$, so $v_2 = 0$. Then $\mathbf{e}_2$ is an eigenvector of $\tilde{\mathbf{A}} = \mathbf{A} + \mathbf{v}\mathbf{v}^\top$ with eigenvalue $1 = d_2$, because $\mathbf{v}^*\mathbf{e}_2 = 0$. So conclusion (a) fails. The secular function $1 + 1/(2 - \lambda) - 1/\lambda$ has only two poles and only two roots, $2 \pm \sqrt{2}$; the third eigenvalue, 1, is invisible to it. The eigenvalues still interlace, $2 + \sqrt{2} \geq 2 \geq 1 \geq 1 \geq 2 - \sqrt{2} \geq 0$, but not strictly.

A repeated d_i deflates too. If $d_1 = d_2$, any vector in $\text{span}(\mathbf{e}_1, \mathbf{e}_2)$ orthogonal to $\mathbf{v}$ is an eigenvector of $\tilde{\mathbf{A}}$ for d_1 , for the same reason.

Negative τ reverses the interlacing. For $\tau < 0$ the matrix $\tau\mathbf{v}\mathbf{v}^*$ is negative semidefinite and every eigenvalue goes **down**: applying Theorem 17.5.1 to $\tilde{\mathbf{A}}$ and $-\tau > 0$ gives $\lambda_i(\mathbf{A}) \geq \lambda_i(\tilde{\mathbf{A}}) \geq \lambda_{i+1}(\mathbf{A})$. Reading the chain of the theorem in the wrong direction is the standard way to get this backwards.

17.5.3 The shifts add up

Interlacing says where each eigenvalue may go. One number says how far they all go together, and it costs a line.

Corollary 17.5.3. Total Eigenvalue Shift

Let $\mathbf{A} \in M_n(F)$ be Hermitian, $\mathbf{v} \in F^n$, $\tau \in \mathbb{R}$, and $\tilde{\mathbf{A}} = \mathbf{A} + \tau\mathbf{v}\mathbf{v}^*$. Then

$$\sum_{i=1}^n (\lambda_i(\tilde{\mathbf{A}}) - \lambda_i(\mathbf{A})) = \tau \|\mathbf{v}\|^2.$$

Proof. By Theorem 0.9.6 (3) applied to the $n \times 1$ matrix $\mathbf{v}$ and the $1 \times n$ matrix $\mathbf{v}^*$,

$$\text{tr}(\mathbf{v}\mathbf{v}^*) = \text{tr}(\mathbf{v}^*\mathbf{v}) = \|\mathbf{v}\|^2.$$

Hence $\text{tr } \tilde{\mathbf{A}} = \text{tr } \mathbf{A} + \tau \|\mathbf{v}\|^2$, by part (1) of the same theorem. Both $\mathbf{A}$ and $\tilde{\mathbf{A}}$ are Hermitian, so Corollary 12.1.4 expresses each trace as the sum of the eigenvalues with multiplicity, and subtracting the two expressions gives the claim. ■

So the whole budget for the movement is $\tau \|\mathbf{v}\|^2$, and for $\tau \geq 0$ Theorem 17.5.1 says how it may be spent: each eigenvalue takes a nonnegative share, and no share except the top one is large enough to carry an eigenvalue past its old upper neighbor. In Example (A bump on a diagonal matrix) the budget 5 was split as $3 + 2$; in Example (A three-by-three update, found from its

secular equation) the budget $\|\mathbf{v}\|^2 = 9$ was split as $6 + 1 + 2$.

Remark. The secular equation of [Theorem 17.5.2](#) is not only a proof device. Because f is increasing between consecutive poles and the bracketing interval is known in advance, each root can be located by bisection on an interval that is guaranteed to contain exactly one of them — a far better-behaved problem than finding the roots of a characteristic polynomial. The secular equation is the engine of the divide-and-conquer eigenvalue algorithms for Hermitian matrices, which split the matrix into two halves joined by a rank-one term and solve one secular equation to merge their spectra. The deflation observed in the warning above, where a zero coordinate of $\mathbf{v}$ hands over an eigenvalue for free, is the same phenomenon that makes the algorithm fast in practice.

17.5.4 Exercises

17.5.4.1 A. Check your understanding

Exercise. A1

- State [Theorem 17.5.1](#), including the hypothesis on τ .
- Write down the secular function of $\mathbf{A} + 2\mathbf{v}\mathbf{v}^\top$ for $\mathbf{A} = \text{diag}(6, 2, -1)$ and $\mathbf{v} = (1, 3, 1)$.
- Explain in one sentence why no d_i is an eigenvalue of $\mathbf{A} + \tau\mathbf{v}\mathbf{v}^*$ under the hypotheses of [Theorem 17.5.2](#).
- Determine whether the following is correct, and justify your answer: if $\tau < 0$ then $\lambda_i(\mathbf{A} + \tau\mathbf{v}\mathbf{v}^*) \geq \lambda_i(\mathbf{A})$ for every i .
- A Hermitian 4×4 matrix $\mathbf{A}$ has eigenvalues 8, 5, 5, 0. What does [Corollary 17.5.3](#) say about $\mathbf{A} + 3\mathbf{v}\mathbf{v}^*$ when $\|\mathbf{v}\| = 2$?

Solution.

- For Hermitian $\mathbf{A} \in M_n(F)$, $\mathbf{v} \in F^n$, and $\tau \geq 0$, the eigenvalues of $\tilde{\mathbf{A}} = \mathbf{A} + \tau\mathbf{v}\mathbf{v}^*$ satisfy $\lambda_i(\tilde{\mathbf{A}}) \geq \lambda_i(\mathbf{A}) \geq \lambda_{i+1}(\tilde{\mathbf{A}})$ for $1 \leq i \leq n-1$, together with $\lambda_n(\tilde{\mathbf{A}}) \geq \lambda_n(\mathbf{A})$. The hypothesis $\tau \geq 0$ is what makes the perturbation positive semidefinite; without it only the second family of inequalities survives.
- $f(\lambda) = 1 + 2\left(\frac{1}{6-\lambda} + \frac{9}{2-\lambda} + \frac{1}{-1-\lambda}\right)$.
- If $\tilde{\mathbf{A}}\mathbf{x} = d_i\mathbf{x}$, comparing i -th coordinates in $(\mathbf{A} - d_i\mathbf{I})\mathbf{x} = -\tau(\mathbf{v}^*\mathbf{x})\mathbf{v}$ forces $\mathbf{v}^*\mathbf{x} = 0$ because $v_i \neq 0$; then $\mathbf{x}$ is a multiple of $\mathbf{e}_i$, and $\mathbf{v}^*\mathbf{x} = 0$ forces that multiple to be 0.
- Incorrect; the inequality goes the other way. For $\tau < 0$ the added matrix is negative semidefinite, and $\lambda_i(\mathbf{A}) \geq \lambda_i(\mathbf{A} + \tau\mathbf{v}\mathbf{v}^*)$ for every i . Concretely, $\mathbf{A} = \text{diag}(1, 0)$, $\tau = -1$, $\mathbf{v} = \mathbf{e}_1$ gives $\tilde{\mathbf{A}} = \text{diag}(0, 0)$ and λ_1 drops from 1 to 0.
- It says the four eigenvalues of $\mathbf{A} + 3\mathbf{v}\mathbf{v}^*$ sum to $8 + 5 + 5 + 0 + 3 \cdot 4 = 30$, that is, 12 more than those of $\mathbf{A}$. Combined with [Theorem 17.5.1](#), which caps λ_2 at 8, λ_3 at 5, and λ_4 at 5, it forces $\lambda_1(\mathbf{A} + 3\mathbf{v}\mathbf{v}^*) \geq 30 - 18 = 12$.

17.5.4.2 B. Practice

Exercise. B1: A secular equation, end to end

Let $\mathbf{A} = \text{diag}(7, 0, -3)$, $\mathbf{v} = (3, 2, 2)$, and $\tau = 1$.

- Write down $\tilde{\mathbf{A}} = \mathbf{A} + \mathbf{v}\mathbf{v}^\top$ and the secular function f .
- Verify that 21, 2, and -2 are roots of f , and explain why these are all the eigenvalues of $\tilde{\mathbf{A}}$.
- Exhibit the interlacing chain of [Theorem 17.5.1](#) for this pair.

Solution.

- $\mathbf{v}\mathbf{v}^\top$ has entries $v_i v_j$, so

$$\tilde{\mathbf{A}} = \begin{pmatrix} 7 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -3 \end{pmatrix} + \begin{pmatrix} 9 & 6 & 6 \\ 6 & 4 & 4 \\ 6 & 4 & 4 \end{pmatrix} = \begin{pmatrix} 16 & 6 & 6 \\ 6 & 4 & 4 \\ 6 & 4 & 1 \end{pmatrix},$$

$$\text{and } f(\lambda) = 1 + \frac{9}{7-\lambda} + \frac{4}{0-\lambda} + \frac{4}{-3-\lambda}.$$

- At $\lambda = 21$: $1 - \frac{9}{14} - \frac{4}{21} - \frac{4}{24} = 1 - \frac{9}{14} - \frac{4}{21} - \frac{1}{6}$. Over 42 this is $(42 - 27 - 8 - 7)/42 = 0$. At $\lambda = 2$: $1 + \frac{9}{5} - 2 - \frac{4}{5} = (5 + 9 - 10 - 4)/5 = 0$. At $\lambda = -2$: $1 + 1 + 2 - 4 = 0$. The hypotheses of [Theorem 17.5.2](#) hold — $7 > 0 > -3$ are distinct, $\tau = 1 > 0$, and every coordinate of $\mathbf{v}$ is nonzero — so by part (c) the secular function has exactly three roots and they are the eigenvalues. We have found three distinct roots, so there are no others.
- $21 \geq 7 \geq 2 \geq 0 \geq -2 \geq -3$, all strict, as part (c) of [Theorem 17.5.2](#) predicts.

Exercise. B2: Reading the chain backwards

Let $\mathbf{B} = \begin{pmatrix} 5 & 2 \\ 2 & 5 \end{pmatrix}$ and $\mathbf{v} = (1, 2)$.

- Compute the eigenvalues of $\mathbf{B}$ and of $\mathbf{B} - \mathbf{v}\mathbf{v}^\top$.
- State which chain of inequalities [Theorem 17.5.1](#) gives here, being careful about which matrix plays the role of $\mathbf{A}$.

Solution.

- $\mathbf{B}$ has trace 10 and determinant 21, so its characteristic polynomial is $x^2 - 10x + 21 = (x - 7)(x - 3)$ and its eigenvalues are 7 and 3. Also

$$\mathbf{B} - \mathbf{v}\mathbf{v}^\top = \begin{pmatrix} 5 & 2 \\ 2 & 5 \end{pmatrix} - \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} = \begin{pmatrix} 4 & 0 \\ 0 & 1 \end{pmatrix},$$

with eigenvalues 4 and 1.

- Put $\mathbf{A} = \mathbf{B} - \mathbf{v}\mathbf{v}^\top = \text{diag}(4, 1)$ and $\tau = 1$, so that $\mathbf{B} = \mathbf{A} + \tau\mathbf{v}\mathbf{v}^\top$ is the *bumped* matrix. The theorem then reads $\lambda_1(\mathbf{B}) \geq \lambda_1(\mathbf{A}) \geq \lambda_2(\mathbf{B}) \geq \lambda_2(\mathbf{A})$, that is, $7 \geq 4 \geq 3 \geq 1$. Taking $\mathbf{B}$ for $\mathbf{A}$ and $\tau = -1$ instead would give the reversed chain

$\lambda_i(\mathbf{B}) \geq \lambda_i(\mathbf{B} - \mathbf{v}\mathbf{v}^\top) \geq \lambda_{i+1}(\mathbf{B})$, namely $7 \geq 4 \geq 3 \geq 1$ again: the same four numbers, read from the other side.

Exercise. B3: The missing eigenvalue

Let $\mathbf{A} = \text{diag}(5, 3, 0)$, $\mathbf{v} = (2, 2, 2)$, and $\tau = 1$. It is given that two of the eigenvalues of $\tilde{\mathbf{A}} = \mathbf{A} + \mathbf{v}\mathbf{v}^\top$ are 15 and 4. Find the third without computing a characteristic polynomial, and then confirm it from the secular equation.

Solution. By [Corollary 17.5.3](#) the eigenvalues of $\tilde{\mathbf{A}}$ sum to

$$(5 + 3 + 0) + \|\mathbf{v}\|^2 = 8 + 12 = 20.$$

Hence the third eigenvalue is $20 - 15 - 4 = 1$.

For the confirmation, $|v_i|^2 = 4$ for each i , so

$$f(\lambda) = 1 + \frac{4}{5 - \lambda} + \frac{4}{3 - \lambda} + \frac{4}{-\lambda},$$

and $f(1) = 1 + 1 + 2 - 4 = 0$. The hypotheses of [Theorem 17.5.2](#) hold, and 1 lies in the interval $(0, 3)$ where part (c) puts the third root. As a check on the other two, $15 > 5$ and $4 \in (3, 5)$, exactly as the theorem requires.

17.5.4.3 C. Going deeper

Exercise. C1: What a zero coordinate does

Let $\mathbf{A} = \text{diag}(d_1, \dots, d_n)$ with the d_i real, let $\tau \in \mathbb{R}$, and let $\mathbf{v} \in F^n$.

- Prove that if $v_i = 0$ then $\mathbf{e}_i$ is an eigenvector of $\mathbf{A} + \tau\mathbf{v}\mathbf{v}^*$ with eigenvalue d_i .
- Suppose $\tau > 0$, the d_i are strictly decreasing, and exactly one coordinate v_k vanishes. Show that the remaining $n - 1$ eigenvalues of $\mathbf{A} + \tau\mathbf{v}\mathbf{v}^*$ are the roots of the function obtained from the secular function by deleting the k -th term. *Hint: restrict to the coordinate subspace spanned by the $\mathbf{e}_i$ with $i \neq k$.*
- Illustrate both parts with $\mathbf{A} = \text{diag}(2, 1, 0)$, $\mathbf{v} = (1, 0, 1)$, and $\tau = 1$.

Solution.

- $\mathbf{v}^*\mathbf{e}_i = \overline{v_i} = 0$, so $(\mathbf{A} + \tau\mathbf{v}\mathbf{v}^*)\mathbf{e}_i = \mathbf{A}\mathbf{e}_i + \tau\mathbf{v}(\mathbf{v}^*\mathbf{e}_i) = d_i\mathbf{e}_i$.
- Let $U = \text{span}(\mathbf{e}_i : i \neq k)$, a subspace of dimension $n - 1$. Every column of $\mathbf{A}$ other than the k -th lies in U , and $\mathbf{v} \in U$ because $v_k = 0$; hence $(\mathbf{A} + \tau\mathbf{v}\mathbf{v}^*)U \subseteq U$. In the ordered basis $(\mathbf{e}_i)_{i \neq k}$ of U , the restriction of $\mathbf{A} + \tau\mathbf{v}\mathbf{v}^*$ to U has matrix $\mathbf{A}' + \tau\mathbf{v}'(\mathbf{v}')^*$, where $\mathbf{A}' = \text{diag}(d_i : i \neq k)$ and $\mathbf{v}'$ is $\mathbf{v}$ with its k -th coordinate removed. Since $U \oplus \text{span}(\mathbf{e}_k) = F^n$ and $\text{span}(\mathbf{e}_k)$ is also invariant by (a), the eigenvalues of $\mathbf{A} + \tau\mathbf{v}\mathbf{v}^*$ are those of the restriction to U together with d_k . Now $\mathbf{A}'$ has strictly decreasing diagonal entries, $\tau > 0$, and $\mathbf{v}'$ has no zero coordinate, so [Theorem 17.5.2](#) applies to $\mathbf{A}' + \tau\mathbf{v}'(\mathbf{v}')^*$: its $n - 1$ eigenvalues are exactly the roots of its secular function.

That function is the secular function of the original update with the k -th term deleted, because that term was $\tau|v_k|^2/(d_k - \lambda) = 0$ anyway.

- (c) Here $v_2 = 0$, so $\mathbf{e}_2$ is an eigenvector with eigenvalue $d_2 = 1$ by (a). By (b) the other two eigenvalues are the roots of

$$1 + \frac{1}{2 - \lambda} + \frac{1}{-\lambda} = 0,$$

that is, of $\lambda(2 - \lambda) + \lambda - (2 - \lambda) = 0$, i.e. $\lambda^2 - 4\lambda + 2 = 0$, with roots $2 \pm \sqrt{2}$. So the spectrum is $\{2 + \sqrt{2}, 1, 2 - \sqrt{2}\}$. Directly,

$$\mathbf{A} + \mathbf{v}\mathbf{v}^\top = \begin{pmatrix} 3 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix},$$

whose second row and column confirm the eigenvalue 1, and whose remaining 2×2 block $\begin{pmatrix} 3 & 1 \\ 1 & 1 \end{pmatrix}$ has trace 4 and determinant 2.

Exercise. C2: Only one eigenvalue can escape

Let $\mathbf{A} \in M_n(F)$ be Hermitian, $\mathbf{v} \in F^n$, $\tau \in \mathbb{R}$, and $\tilde{\mathbf{A}} = \mathbf{A} + \tau\mathbf{v}\mathbf{v}^*$.

- Prove that **at most one** eigenvalue of $\tilde{\mathbf{A}}$ lies outside the closed interval $[\lambda_n(\mathbf{A}), \lambda_1(\mathbf{A})]$, and say which one it is and on which side, according to the sign of τ .
- Prove that the escaping eigenvalue moves by at most $|\tau|\|\mathbf{v}\|^2$, and show by examples with $n = 2$ that it may move by exactly that amount, and that it may not escape at all.

Solution.

- Suppose first $\tau \geq 0$. By [Theorem 17.5.1](#), $\lambda_n(\tilde{\mathbf{A}}) \geq \lambda_n(\mathbf{A})$, so no eigenvalue of $\tilde{\mathbf{A}}$ falls below the interval; and for $i \geq 2$, $\lambda_i(\tilde{\mathbf{A}}) \leq \lambda_{i-1}(\mathbf{A}) \leq \lambda_1(\mathbf{A})$, so only $\lambda_1(\tilde{\mathbf{A}})$ can rise above it. So at most one eigenvalue escapes, namely the largest, and it escapes upwards.
For $\tau < 0$, the warning above gives $\lambda_i(\mathbf{A}) \geq \lambda_i(\tilde{\mathbf{A}}) \geq \lambda_{i+1}(\mathbf{A})$ for $i \leq n - 1$ and $\lambda_n(\mathbf{A}) \geq \lambda_n(\tilde{\mathbf{A}})$. Hence $\lambda_1(\tilde{\mathbf{A}}) \leq \lambda_1(\mathbf{A})$, and $\lambda_i(\tilde{\mathbf{A}}) \geq \lambda_{i+1}(\mathbf{A}) \geq \lambda_n(\mathbf{A})$ for $i \leq n - 1$. So only $\lambda_n(\tilde{\mathbf{A}})$ can escape, and it escapes downwards.
- Take $\tau > 0$; the case $\tau < 0$ is the mirror image. By [Theorem 17.5.1](#) every shift $\lambda_i(\tilde{\mathbf{A}}) - \lambda_i(\mathbf{A})$ is ≥ 0 , and by [Corollary 17.5.3](#) they sum to $\tau\|\mathbf{v}\|^2$. So each shift, in particular that of λ_1 , is at most $\tau\|\mathbf{v}\|^2$.

For equality, let $\mathbf{A} = \text{diag}(2, 1)$, $\mathbf{v} = \mathbf{e}_1$, $\tau = 3$: then $\tilde{\mathbf{A}} = \text{diag}(5, 1)$, and λ_1 moves from 2 to 5, by exactly $\tau\|\mathbf{v}\|^2 = 3$. For no escape, let $\mathbf{v} = \mathbf{e}_2$ and $\tau = 1$: then $\tilde{\mathbf{A}} = \text{diag}(2, 2)$, whose eigenvalues both lie in $[1, 2]$. The whole budget went to λ_2 , which [Theorem 17.5.1](#) caps at $\lambda_1(\mathbf{A}) = 2$.

Exercise. C3: Bumping along an eigenvector

Let $\mathbf{A} \in M_n(F)$ be Hermitian with an orthonormal basis of eigenvectors $\mathbf{u}_1, \dots, \mathbf{u}_n$, where $\mathbf{A}\mathbf{u}_i = \lambda_i(\mathbf{A})\mathbf{u}_i$.

- (a) Fix k and $\tau \in \mathbb{R}$. Prove that the eigenvalues of $\mathbf{A} + \tau\mathbf{u}_k\mathbf{u}_k^*$, listed with multiplicity but not necessarily in order, are $\lambda_k(\mathbf{A}) + \tau$ together with $\lambda_i(\mathbf{A})$ for $i \neq k$.
- (b) Deduce that $\mathbf{A} - \lambda_1(\mathbf{A})\mathbf{u}_1\mathbf{u}_1^*$ has eigenvalue list $0, \lambda_2(\mathbf{A}), \dots, \lambda_n(\mathbf{A})$. (This is the **deflation** step: it removes the top eigenvalue and leaves the rest untouched, which is what makes it possible to find eigenvalues one at a time.)
- (c) Explain why (a) does not contradict part (c) of [Theorem 17.5.2](#).

Solution.

- (a) For $i \neq k$, orthonormality gives $\mathbf{u}_k^*\mathbf{u}_i = 0$, so

$$(\mathbf{A} + \tau\mathbf{u}_k\mathbf{u}_k^*)\mathbf{u}_i = \mathbf{A}\mathbf{u}_i + \tau\mathbf{u}_k(\mathbf{u}_k^*\mathbf{u}_i) = \lambda_i(\mathbf{A})\mathbf{u}_i.$$

For $i = k$, $\mathbf{u}_k^*\mathbf{u}_k = 1$ gives $(\mathbf{A} + \tau\mathbf{u}_k\mathbf{u}_k^*)\mathbf{u}_k = (\lambda_k(\mathbf{A}) + \tau)\mathbf{u}_k$. So $(\mathbf{u}_1, \dots, \mathbf{u}_n)$ is a basis of F^n consisting of eigenvectors of $\mathbf{A} + \tau\mathbf{u}_k\mathbf{u}_k^*$, with the stated eigenvalues. A basis of eigenvectors exhibits the whole eigenvalue list with multiplicity.

- (b) Take $k = 1$ and $\tau = -\lambda_1(\mathbf{A})$ in (a).
- (c) There is no contradiction, because the hypothesis of [Theorem 17.5.2](#) that fails here fails as badly as possible. Writing everything in the orthonormal eigenbasis, the update vector is $\mathbf{u}_k$, whose coordinate vector is $\mathbf{e}_k$: **all but one** of its coordinates are zero. Part (a) is the extreme case of [Exercise \(C1: What a zero coordinate does\)](#) (a), with $n - 1$ coordinates deflating at once, and only one eigenvalue is left to move.