

20.9 The Davis–Kahan Theorem

Weyl’s inequality settled the eigenvalues of a Hermitian matrix: under a Hermitian perturbation $\mathbf{E}$, each of them moves by at most $\|\mathbf{E}\|_2$, whatever the matrix and however its eigenvalues are spaced. Eigenvectors are a different matter, and this section is about them. Section 8 supplied the yardstick, the principal angles between two subspaces, and Section 7 supplied the tool, bounds for the equation $\mathbf{A}\mathbf{X} - \mathbf{X}\mathbf{B} = \mathbf{Y}$. The theorem of this section is what happens when the two meet: the perturbation of an invariant subspace satisfies a Sylvester equation, and the bounds of Section 7 then measure it.

Throughout, $\mathbf{A}$ and $\tilde{\mathbf{A}} = \mathbf{A} + \mathbf{E}$ are Hermitian matrices in $M_n(\mathbb{C})$, so $\mathbf{E} = \tilde{\mathbf{A}} - \mathbf{A}$ is Hermitian too, and $1 \leq k \leq n - 1$. As everywhere in this chapter, $\|\cdot\|_2$ is the spectral norm and $\|\cdot\|_F$ the Frobenius norm, and every bound names the norm it holds in. For a subspace we write $\text{col}(\mathbf{Q})$ for the column space of a matrix $\mathbf{Q}$ whose columns span it.

20.9.1 Eigenvectors can move a long way

Before any theorem, here is what goes wrong. The example is as small as it can be, and every number in it is exact.

Example. A tiny perturbation that turns the eigenvectors through 45°

Let $0 < \eta < 1$ and

$$\mathbf{A} = \begin{pmatrix} 1+\eta & 0 \\ 0 & 1-\eta \end{pmatrix}, \quad \mathbf{E} = \eta \begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix}.$$

Compare the eigenvalues and the eigenvectors of $\mathbf{A}$ and of $\tilde{\mathbf{A}} = \mathbf{A} + \mathbf{E}$, and compute $\|\mathbf{E}\|_2$.

Solution. Adding, $\tilde{\mathbf{A}} = \begin{pmatrix} 1 & \eta \\ \eta & 1 \end{pmatrix}$. Since

$$\tilde{\mathbf{A}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = (1+\eta) \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \tilde{\mathbf{A}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = (1-\eta) \begin{pmatrix} 1 \\ -1 \end{pmatrix},$$

the eigenvalues of $\tilde{\mathbf{A}}$ are $1+\eta$ and $1-\eta$: **exactly** those of $\mathbf{A}$. The eigenvectors are another story. For the eigenvalue $1+\eta$, the matrix $\mathbf{A}$ has the unit eigenvector $\mathbf{e}_1$ and $\tilde{\mathbf{A}}$ has $\frac{1}{\sqrt{2}}(1, 1)$; the two lines they span meet at the angle $\pi/4$, because $|\langle \mathbf{e}_1, \frac{1}{\sqrt{2}}(1, 1) \rangle| = \frac{1}{\sqrt{2}} = \cos \frac{\pi}{4}$. The same happens for $1-\eta$, with $\mathbf{e}_2$ and $\frac{1}{\sqrt{2}}(1, -1)$.

The matrix $\mathbf{E}/\eta$ is Hermitian and squares to $2\mathbf{I}_2$, so its eigenvalues are $\pm\sqrt{2}$ (it has trace 0), and by Lemma 17.3.5

$$\|\mathbf{E}\|_2 = \sqrt{2}\eta.$$

So as $\eta \rightarrow 0^+$ the perturbation vanishes, the eigenvalues do not move at all, and yet the eigenvectors of $\tilde{\mathbf{A}}$ stay a full 45° away from those of $\mathbf{A}$.

The rotation does not shrink with η . What does shrink with η is the distance 2η between the two eigenvalues of $\mathbf{A}$, and that is the whole story: the two eigenvalues are so close that a perturbation of their own size can mix their eigenvectors freely. An eigenvector bound must therefore divide by a gap, and the natural guess is “angle $\lesssim$ perturbation / gap”. Here that ratio

is $\sqrt{2}\eta/2\eta = 1/\sqrt{2} = \sin(\pi/4)$, which is exactly the sine of the angle. The theorem below turns the guess into an inequality, and this example will turn out to attain it.

Warning. Well-conditioned eigenvalues do not make well-conditioned eigenvectors. Corollary 17.3.6 guarantees that every eigenvalue of a Hermitian matrix moves by at most $\|\mathbf{E}\|_2$, with no mention of gaps, and it is tempting to expect the same of eigenvectors. **Example (A tiny perturbation that turns the eigenvectors through 45°)** refutes this: a perturbation of size $\sqrt{2}\eta$ leaves every eigenvalue fixed and turns every eigenvector through $\pi/4$, however small η is. No inequality of the form $\sin \theta \leq C\|\mathbf{E}\|_2$, with C depending only on n , can hold.

20.9.2 Two norms, three facts

The proofs of this section and the next shuffle unitary factors and matrices with orthonormal columns in and out of norms, sometimes the spectral one and sometimes the Frobenius one. The facts needed are elementary, but they are needed so often that they are worth collecting once. Part (c) is mostly quoted: its norm identities for a matrix with orthonormal columns are Lemma 20.8.1 (b) of Section 8, read with $U = \text{col}(\mathbf{X})$, and for the Frobenius norm the unitary invariance is Lemma 11.7.5. The rest — (a), (b), and the inequality $\|\mathbf{X}^*\mathbf{N}\| \leq \|\mathbf{N}\|$ — is proved below, and what the lemma is really for is to have all of it, in both norms, in one place.

Lemma 20.9.1. Moving Factors in and out of the Two Norms

Let $\|\cdot\|$ stand for either $\|\cdot\|_2$ or $\|\cdot\|_F$, the same one throughout each statement.

- (a) $\|\mathbf{M}^*\| = \|\mathbf{M}\|$ for every $\mathbf{M} \in M_{m \times n}(\mathbb{C})$.
- (b) $\|\mathbf{XMY}\| \leq \|\mathbf{X}\|_2 \|\mathbf{M}\| \|\mathbf{Y}\|_2$ whenever the product is defined.
- (c) If $\mathbf{X} \in M_{m \times k}(\mathbb{C})$ has orthonormal columns, that is $\mathbf{X}^*\mathbf{X} = \mathbf{I}_k$, then $\|\mathbf{X}\|_2 = 1$, $\|\mathbf{XM}\| = \|\mathbf{M}\|$, and $\|\mathbf{X}^*\mathbf{N}\| \leq \|\mathbf{N}\|$ for all $\mathbf{M} \in M_{k \times n}(\mathbb{C})$ and $\mathbf{N} \in M_{m \times n}(\mathbb{C})$. In particular $\|\mathbf{UMV}\| = \|\mathbf{M}\|$ for unitary $\mathbf{U}$ and $\mathbf{V}$.

Proof.

- (a) For $\|\cdot\|_F$ this holds because $\mathbf{M}^*$ has the same entries as $\mathbf{M}$ up to position and conjugation. For $\|\cdot\|_2$, let $\mathbf{x} \in \mathbb{C}^n$. By Theorem 11.1.5 and Theorem 16.3.2 (a), applied twice,

$$\|\mathbf{Mx}\|^2 = \langle \mathbf{M}^*\mathbf{Mx}, \mathbf{x} \rangle \leq \|\mathbf{M}^*\mathbf{Mx}\| \|\mathbf{x}\| \leq \|\mathbf{M}^*\|_2 \|\mathbf{M}\|_2 \|\mathbf{x}\|^2.$$

Taking the maximum over unit $\mathbf{x}$ gives $\|\mathbf{M}\|_2^2 \leq \|\mathbf{M}^*\|_2 \|\mathbf{M}\|_2$, hence $\|\mathbf{M}\|_2 \leq \|\mathbf{M}^*\|_2$ when $\mathbf{M} \neq \mathbf{O}$ (and trivially when $\mathbf{M} = \mathbf{O}$). Applying this to $\mathbf{M}^*$, whose conjugate transpose is $\mathbf{M}$, gives the reverse inequality.

- (b) For $\|\cdot\|_2$ this is Theorem 16.3.2 (d) used twice. For $\|\cdot\|_F$, let $\mathbf{m}_1, \dots, \mathbf{m}_n$ be the columns of $\mathbf{M}$. The columns of $\mathbf{XM}$ are the $\mathbf{Xm}_j$, and the squared Frobenius norm is the sum of the squared lengths of the columns, so

$$\|\mathbf{XM}\|_F^2 = \sum_j \|\mathbf{Xm}_j\|^2 \leq \|\mathbf{X}\|_2^2 \sum_j \|\mathbf{m}_j\|^2 = \|\mathbf{X}\|_2^2 \|\mathbf{M}\|_F^2.$$

For a factor on the right, (a) and the left-hand case give $\|\mathbf{MY}\|_F = \|\mathbf{Y}^*\mathbf{M}^*\|_F \leq \|\mathbf{Y}^*\|_2 \|\mathbf{M}^*\|_F = \|\mathbf{Y}\|_2 \|\mathbf{M}\|_F$. Combining the two cases gives (b).

(c) For every $\mathbf{v} \in \mathbb{C}^k$, $\|\mathbf{X}\mathbf{v}\|^2 = \mathbf{v}^* \mathbf{X}^* \mathbf{X} \mathbf{v} = \|\mathbf{v}\|^2$. Hence $\|\mathbf{X}\|_2 = 1$, and multiplying by $\mathbf{X}$ changes the length of no column, so $\|\mathbf{X}\mathbf{M}\|_F = \|\mathbf{M}\|_F$ and $\|\mathbf{X}\mathbf{M}\mathbf{v}\| = \|\mathbf{M}\mathbf{v}\|$ for every $\mathbf{v}$, which gives $\|\mathbf{X}\mathbf{M}\|_2 = \|\mathbf{M}\|_2$. Next, $\|\mathbf{X}^*\|_2 = \|\mathbf{X}\|_2 = 1$ by (a), so $\|\mathbf{X}^* \mathbf{N}\| \leq \|\mathbf{N}\|$ by (b). Finally, for unitary $\mathbf{U}$ and $\mathbf{V}$ both $\mathbf{U}$ and $\mathbf{V}^*$ have orthonormal columns, so $\|\mathbf{U}\mathbf{M}\mathbf{V}\| = \|\mathbf{M}\mathbf{V}\| = \|\mathbf{V}^* \mathbf{M}^*\| = \|\mathbf{M}^*\| = \|\mathbf{M}\|$, using (a) twice. ■

Part (c) is the one used most. Its inequality $\|\mathbf{X}^* \mathbf{N}\| \leq \|\mathbf{N}\|$ says that reading $\mathbf{N}$ in only some of the coordinates of an orthonormal basis can only lose mass, never gain it.

20.9.3 The perturbation of a subspace is a Sylvester equation

Fix the data. A Hermitian matrix is normal, so by Corollary 12.4.3 there is a unitary $\mathbf{Q} \in M_n(\mathbb{C})$ with $\mathbf{Q}^* \mathbf{A} \mathbf{Q}$ diagonal, the eigenvalues appearing in any order we like, and they are real by Theorem 12.2.1. Choose k of them, list them first, and split the columns accordingly:

$$\mathbf{Q} = \begin{pmatrix} \mathbf{Q}_1 & \mathbf{Q}_2 \end{pmatrix}, \quad \mathbf{Q}^* \mathbf{A} \mathbf{Q} = \begin{pmatrix} \Lambda_1 & \mathbf{O} \\ \mathbf{O} & \Lambda_2 \end{pmatrix},$$

with $\mathbf{Q}_1 \in M_{n \times k}(\mathbb{C})$ and $\Lambda_1 \in M_k(\mathbb{R})$, $\Lambda_2 \in M_{n-k}(\mathbb{R})$ diagonal. Do the same for $\tilde{\mathbf{A}}$, producing $\tilde{\mathbf{Q}} = \begin{pmatrix} \tilde{\mathbf{Q}}_1 & \tilde{\mathbf{Q}}_2 \end{pmatrix}$, $\tilde{\Lambda}_1$, and $\tilde{\Lambda}_2$. Reading $\mathbf{A} \mathbf{Q} = \mathbf{Q}(\Lambda_1 \oplus \Lambda_2)$ column by column,

$$\mathbf{A} \mathbf{Q}_1 = \mathbf{Q}_1 \Lambda_1, \quad \tilde{\mathbf{A}} \tilde{\mathbf{Q}}_2 = \tilde{\mathbf{Q}}_2 \tilde{\Lambda}_2,$$

so the columns of $\mathbf{Q}_1$ are orthonormal eigenvectors of $\mathbf{A}$, and $U = \text{col}(\mathbf{Q}_1)$ is a k -dimensional subspace invariant under $\mathbf{A}$. Likewise $W = \text{col}(\tilde{\mathbf{Q}}_1)$ is a k -dimensional subspace invariant under $\tilde{\mathbf{A}}$. The question is how far apart U and W are, and Section 8 answers “measure the principal angles $\theta_1 \leq \dots \leq \theta_k$ of Definition 20.8.1”.

Put

$$\mathbf{X} := \tilde{\mathbf{Q}}_2^* \mathbf{Q}_1 \in M_{(n-k) \times k}(\mathbb{C}).$$

It measures the angles, and it satisfies an equation.

The unperturbed subspace, read in the perturbed complement, solves a Sylvester equation whose coefficients are the two spectra and whose right-hand side is the perturbation.

Lemma 20.9.2. The Subspace Perturbation Equation

In the notation above,

$$\tilde{\Lambda}_2 \mathbf{X} - \mathbf{X} \Lambda_1 = \tilde{\mathbf{Q}}_2^* \mathbf{E} \mathbf{Q}_1, \tag{20.9.1}$$

and the nonzero singular values of $\mathbf{X}$ are the nonzero sines $\sin \theta_i$ of the principal angles between $U = \text{col}(\mathbf{Q}_1)$ and $W = \text{col}(\tilde{\mathbf{Q}}_1)$, with multiplicity. In particular

$$\|\mathbf{X}\|_F = \|\sin \Theta(U, W)\|_F, \quad \|\mathbf{X}\|_2 = \sin \theta_k = \|\sin \Theta(U, W)\|_2,$$

where $\sin \Theta(U, W) = \text{diag}(\sin \theta_1, \dots, \sin \theta_k)$.

Idea. Compute $\tilde{\mathbf{Q}}_2^* \tilde{\mathbf{A}} \tilde{\mathbf{Q}}_1$ twice. Letting $\tilde{\mathbf{A}}$ act to the left, on the perturbed eigenvectors, produces $\tilde{\mathbf{\Lambda}}_2$. Writing $\tilde{\mathbf{A}} = \mathbf{A} + \mathbf{E}$ and letting $\mathbf{A}$ act to the right, on the unperturbed ones, produces $\mathbf{\Lambda}_1$, plus an error term that carries $\mathbf{E}$. The difference of the two computations is the identity.

Proof. Taking conjugate transposes in $\tilde{\mathbf{A}} \tilde{\mathbf{Q}}_2 = \tilde{\mathbf{Q}}_2 \tilde{\mathbf{\Lambda}}_2$, and using $\tilde{\mathbf{A}}^* = \tilde{\mathbf{A}}$ and $\tilde{\mathbf{\Lambda}}_2^* = \tilde{\mathbf{\Lambda}}_2$ (a real diagonal matrix), gives $\tilde{\mathbf{Q}}_2^* \tilde{\mathbf{A}} = \tilde{\mathbf{\Lambda}}_2 \tilde{\mathbf{Q}}_2^*$. Hence

$$\tilde{\mathbf{Q}}_2^* \tilde{\mathbf{A}} \tilde{\mathbf{Q}}_1 = \tilde{\mathbf{\Lambda}}_2 \tilde{\mathbf{Q}}_2^* \tilde{\mathbf{Q}}_1 = \tilde{\mathbf{\Lambda}}_2 \mathbf{X}.$$

On the other hand $\tilde{\mathbf{A}} = \mathbf{A} + \mathbf{E}$ and $\mathbf{A} \tilde{\mathbf{Q}}_1 = \tilde{\mathbf{Q}}_1 \mathbf{\Lambda}_1$, so

$$\begin{aligned} \tilde{\mathbf{Q}}_2^* \tilde{\mathbf{A}} \tilde{\mathbf{Q}}_1 &= \tilde{\mathbf{Q}}_2^* \tilde{\mathbf{Q}}_1 \mathbf{\Lambda}_1 + \tilde{\mathbf{Q}}_2^* \mathbf{E} \tilde{\mathbf{Q}}_1 \\ &= \mathbf{X} \mathbf{\Lambda}_1 + \tilde{\mathbf{Q}}_2^* \mathbf{E} \tilde{\mathbf{Q}}_1. \end{aligned}$$

Subtracting the two expressions gives [Eq. 20.9.1](#).

For the angles: the columns of the unitary $\tilde{\mathbf{Q}}$ form an orthonormal basis of $\mathbb{C}^n$ (Theorem 11.7.3), so the $n - k$ columns of $\tilde{\mathbf{Q}}_2$ are orthonormal and orthogonal to every column of $\tilde{\mathbf{Q}}_1$. They therefore span a subspace of $W^\perp$ of dimension $n - k$, which is all of $W^\perp$ because $\dim W^\perp = n - k$ by Theorem 11.3.2 (c). So $\tilde{\mathbf{Q}}_2$ is a matrix whose columns are an orthonormal basis of $W^\perp$, and Theorem 20.8.4 (a) says that the singular values of $\tilde{\mathbf{Q}}_2^* \tilde{\mathbf{Q}}_1 = \mathbf{X}$ are the $\sin \theta_i$, as a multiset completed by zeros; zeros change neither a sum of squares nor a maximum of nonnegative numbers. For the two norms, $\mathbf{I} - P_W = P_{W^\perp} = \tilde{\mathbf{Q}}_2 \tilde{\mathbf{Q}}_2^*$ by Lemma 20.8.1 (d) and (a), so $(\mathbf{I} - P_W) \tilde{\mathbf{Q}}_1 = \tilde{\mathbf{Q}}_2 \mathbf{X}$, and [Lemma 20.9.1](#) (c) removes the factor $\tilde{\mathbf{Q}}_2$:

$$\|\mathbf{X}\|_F = \|(\mathbf{I} - P_W) \tilde{\mathbf{Q}}_1\|_F, \quad \|\mathbf{X}\|_2 = \|(\mathbf{I} - P_W) \tilde{\mathbf{Q}}_1\|_2.$$

By Theorem 20.8.4 (b), applied with U and W , these are $\|\sin \Theta(U, W)\|_F$ and $\sin \theta_k = \|\sin \Theta(U, W)\|_2$. This proves the lemma. ■

Quick check. Take $\mathbf{A} = \tilde{\mathbf{A}} = \mathbf{I}_2$, so $\mathbf{E} = \mathbf{O}$, with $k = 1$, $\tilde{\mathbf{Q}}_1 = \mathbf{e}_1$, and $\tilde{\mathbf{Q}}_2 = \frac{1}{\sqrt{2}}(1, 1)$; both are legitimate choices of unit eigenvector. The subspaces U and W are 45° apart, although the matrices are equal. What does [Eq. 20.9.1](#) say here, and why is there no contradiction?

Solution. Here $\mathbf{\Lambda}_1 = (1)$ and $\tilde{\mathbf{\Lambda}}_2 = (1)$, so the equation reads $1 \cdot \mathbf{X} - \mathbf{X} \cdot 1 = \mathbf{O}$, which every 1×1 matrix satisfies. The equation holds, as it must, but it does not determine $\mathbf{X}$, and $\mathbf{X} = \tilde{\mathbf{Q}}_2^* \tilde{\mathbf{Q}}_1 = \frac{1}{\sqrt{2}}(1, -1) \cdot \mathbf{e}_1 = \frac{1}{\sqrt{2}} = \sin \frac{\pi}{4}$ is one of its many solutions. The trouble is that the spectra of $\tilde{\mathbf{\Lambda}}_2$ and $\mathbf{\Lambda}_1$ overlap. The theorems below assume that they are separated, and then the equation has exactly one solution (Theorem 12.10.2), which is small when the right-hand side is.

20.9.4 The $\sin \Theta$ theorem in the Frobenius norm

With the identity in hand, the theorem is a matter of dividing by the gap. The gap that matters compares **perturbed** eigenvalues outside the subspace with **unperturbed** eigenvalues inside it, because those are the two coefficient matrices of the equation.

Theorem 20.9.3. The Davis–Kahan $\sin \Theta$ Theorem

Let $\mathbf{A}, \tilde{\mathbf{A}} = \mathbf{A} + \mathbf{E} \in M_n(\mathbb{C})$ be Hermitian, let $1 \leq k \leq n - 1$, and let $\mathbf{Q} = \begin{pmatrix} \mathbf{Q}_1 & \mathbf{Q}_2 \end{pmatrix}$ and $\tilde{\mathbf{Q}} = \begin{pmatrix} \tilde{\mathbf{Q}}_1 & \tilde{\mathbf{Q}}_2 \end{pmatrix}$ be unitary, with $\mathbf{Q}_1, \tilde{\mathbf{Q}}_1 \in M_{n \times k}(\mathbb{C})$, such that

$$\mathbf{Q}^* \mathbf{A} \mathbf{Q} = \mathbf{\Lambda}_1 \oplus \mathbf{\Lambda}_2, \quad \tilde{\mathbf{Q}}^* \tilde{\mathbf{A}} \tilde{\mathbf{Q}} = \tilde{\mathbf{\Lambda}}_1 \oplus \tilde{\mathbf{\Lambda}}_2$$

are real diagonal. Suppose there is a $\delta > 0$ such that $|\mu - \lambda| \geq \delta$ for **every** diagonal entry μ of $\tilde{\mathbf{\Lambda}}_2$ and **every** diagonal entry λ of $\mathbf{\Lambda}_1$. Then the principal angles $\Theta = \Theta(\text{col}(\mathbf{Q}_1), \text{col}(\tilde{\mathbf{Q}}_1))$ satisfy

$$\|\sin \Theta\|_F \leq \frac{\|\mathbf{E} \mathbf{Q}_1\|_F}{\delta} \leq \frac{\sqrt{k} \|\mathbf{E}\|_2}{\delta}.$$

Idea. Eq. 20.9.1 is a Sylvester equation whose two coefficient matrices are diagonal, hence normal, and for a normal pair the separation of Section 7 is exactly the distance between the two spectra. The hypothesis says that this distance is at least δ , so Section 7's bound divides the right-hand side by δ . What is left is Lemma 20.9.2, which reads the size of $\mathbf{X}$ as the size of $\sin \Theta$.

Proof. Write $\mu_1, \dots, \mu_{n-k}$ for the diagonal entries of $\tilde{\mathbf{\Lambda}}_2$, $\lambda_1, \dots, \lambda_k$ for those of $\mathbf{\Lambda}_1$, and $\mathbf{Y} = \tilde{\mathbf{Q}}_2^* \mathbf{E} \mathbf{Q}_1$, so that $\tilde{\mathbf{\Lambda}}_2 \mathbf{X} - \mathbf{X} \mathbf{\Lambda}_1 = \mathbf{Y}$ by Lemma 20.9.2. A diagonal matrix is normal, and its eigenvalues are its diagonal entries, so Theorem 20.7.2 gives

$$\text{sep}_F(\tilde{\mathbf{\Lambda}}_2, \mathbf{\Lambda}_1) = \min_{i,j} |\mu_i - \lambda_j| \geq \delta > 0.$$

In particular the two spectra are disjoint, so $\mathbf{X}$ is the **unique** solution of that equation (Theorem 12.10.2), and Proposition 20.7.1 (d) applies to it:

$$\|\mathbf{X}\|_F \leq \frac{\|\mathbf{Y}\|_F}{\text{sep}_F(\tilde{\mathbf{\Lambda}}_2, \mathbf{\Lambda}_1)} \leq \frac{\|\mathbf{Y}\|_F}{\delta}.$$

The columns of $\tilde{\mathbf{Q}}_2$ are orthonormal, so Lemma 20.9.1 (c) gives $\|\mathbf{Y}\|_F = \|\tilde{\mathbf{Q}}_2^* (\mathbf{E} \mathbf{Q}_1)\|_F \leq \|\mathbf{E} \mathbf{Q}_1\|_F$. By Lemma 20.9.1 (b), $\|\mathbf{E} \mathbf{Q}_1\|_F \leq \|\mathbf{E}\|_2 \|\mathbf{Q}_1\|_F$, and $\|\mathbf{Q}_1\|_F = \sqrt{k}$ because $\mathbf{Q}_1$ has k unit columns. Finally $\|\mathbf{X}\|_F = \|\sin \Theta\|_F$ by Lemma 20.9.2. Chaining the three inequalities proves the theorem. ■

The right-hand side is a Lipschitz bound with constant $1/\delta$, and it is sharp. In Example (A tiny perturbation that turns the eigenvectors through 45°), take $k = 1$, $\mathbf{Q}_1 = \mathbf{e}_1$ with $\mathbf{\Lambda}_1 = (1 + \eta)$, and $\tilde{\mathbf{Q}}_2 = \frac{1}{\sqrt{2}}(1, -1)$ with $\tilde{\mathbf{\Lambda}}_2 = (1 - \eta)$. Then $\delta = 2\eta$ and $\mathbf{E} \mathbf{e}_1 = \eta(-1, 1)$ has length $\sqrt{2}\eta$, so

the theorem gives

$$\sin \theta_1 \leq \frac{\sqrt{2}\eta}{2\eta} = \frac{1}{\sqrt{2}},$$

and the true angle is $\pi/4$, whose sine is exactly $1/\sqrt{2}$. The large rotation there was not a failure of the theorem. It is what the theorem predicts when the gap is no bigger than the perturbation.

Remark. The roles of $\mathbf{A}$ and $\tilde{\mathbf{A}}$ may be exchanged, since $\mathbf{A} = \tilde{\mathbf{A}} + (-\mathbf{E})$ and the principal angles between U and W are those between W and U (a matrix and its conjugate transpose have the same singular values). This gives the companion bound $\|\sin \Theta\|_F \leq \|\mathbf{E}\tilde{\mathbf{Q}}_1\|_F/\delta'$, where δ' separates the diagonal of $\mathbf{\Lambda}_2$ from that of $\tilde{\mathbf{\Lambda}}_1$. Use whichever one's data are at hand.

20.9.5 A one-sided version in the spectral norm

The Frobenius theorem controls all k angles at once. Often only the largest angle matters, $\|\sin \Theta\|_2 = \sin \theta_k$, and the Frobenius theorem bounds it too, since $\|\sin \Theta\|_2 \leq \|\sin \Theta\|_F$. But its right-hand side $\|\mathbf{E}\mathbf{Q}_1\|_F$ can be $\sqrt{k}$ times $\|\mathbf{E}\mathbf{Q}_1\|_2$, and for large k that factor is a real loss. When the two spectra are separated by a single point, one entirely above it and the other entirely below, the one-sided bound of Section 7 removes it.

Theorem 20.9.4. The $\sin \Theta$ Theorem, One-Sided Case

In the setting of Theorem 20.9.3, suppose instead that there are $a \in \mathbb{R}$ and $\delta > 0$ such that either

- (i) every eigenvalue in $\mathbf{\Lambda}_1$ is $\geq a$ and every eigenvalue in $\tilde{\mathbf{\Lambda}}_2$ is $\leq a - \delta$, or
- (ii) every eigenvalue in $\mathbf{\Lambda}_1$ is $\leq a$ and every eigenvalue in $\tilde{\mathbf{\Lambda}}_2$ is $\geq a + \delta$.

Then

$$\|\sin \Theta\|_2 = \sin \theta_k \leq \frac{\|\mathbf{E}\mathbf{Q}_1\|_2}{\delta} \leq \frac{\|\mathbf{E}\|_2}{\delta}.$$

Proof. Let $\mathbf{X} = \tilde{\mathbf{Q}}_2^* \mathbf{Q}_1$ and $\mathbf{Y} = \tilde{\mathbf{Q}}_2^* \mathbf{E}\mathbf{Q}_1$, so that $\tilde{\mathbf{\Lambda}}_2 \mathbf{X} - \mathbf{X} \mathbf{\Lambda}_1 = \mathbf{Y}$ by Lemma 20.9.2. For a real diagonal matrix, $\mathbf{\Lambda} \succeq \alpha \mathbf{I}$ means that every diagonal entry is at least α (Theorem 13.1.4 (b), applied to $\mathbf{\Lambda} - \alpha \mathbf{I}$).

Case (ii). Here $\tilde{\mathbf{\Lambda}}_2 \succeq (a + \delta) \mathbf{I}$ and $\mathbf{\Lambda}_1 \preceq a \mathbf{I}$, with $a + \delta > a$. By Theorem 20.7.3, applied with $\tilde{\mathbf{\Lambda}}_2$ and $\mathbf{\Lambda}_1$ in the roles of its two Hermitian coefficients, the solution of $\tilde{\mathbf{\Lambda}}_2 \mathbf{X} - \mathbf{X} \mathbf{\Lambda}_1 = \mathbf{Y}$, which is unique because the two spectra are disjoint, satisfies $\|\mathbf{X}\|_2 \leq \|\mathbf{Y}\|_2/\delta$.

Case (i). Multiply the equation by -1 : $(-\tilde{\mathbf{\Lambda}}_2) \mathbf{X} - \mathbf{X} (-\mathbf{\Lambda}_1) = -\mathbf{Y}$, where now $-\tilde{\mathbf{\Lambda}}_2 \preceq (\delta - a) \mathbf{I}$ and $-\mathbf{\Lambda}_1 \preceq -a \mathbf{I}$, and $(\delta - a) - (-a) = \delta > 0$. So case (ii)'s argument applies and gives $\|\mathbf{X}\|_2 \leq \|-\mathbf{Y}\|_2/\delta = \|\mathbf{Y}\|_2/\delta$.

In both cases, Lemma 20.9.1 (c) and (b) give $\|\mathbf{Y}\|_2 \leq \|\mathbf{E}\mathbf{Q}_1\|_2 \leq \|\mathbf{E}\|_2 \|\mathbf{Q}_1\|_2 = \|\mathbf{E}\|_2$, and $\|\mathbf{X}\|_2 = \sin \theta_k$ by Lemma 20.9.2. This proves the theorem. ■

The typical use is the top of the spectrum: $\mathbf{\Lambda}_1$ holds the k largest eigenvalues of $\mathbf{A}$, and $\tilde{\mathbf{\Lambda}}_2$ the

$n - k$ smallest of $\tilde{\mathbf{A}}$. Then (i) holds with $a = \lambda_k(\mathbf{A})$ and $\delta = \lambda_k(\mathbf{A}) - \lambda_{k+1}(\tilde{\mathbf{A}})$, provided that number is positive.

Remark. Davis and Kahan proved more. If the eigenvalues in $\mathbf{A}_1$ lie in an interval $[a, b]$ and every eigenvalue in $\tilde{\mathbf{A}}_2$ lies outside $(a - \delta, b + \delta)$, on either side, then the spectral bound $\sin \theta_k \leq \|\mathbf{E}\mathbf{Q}_1\|_2/\delta$ still holds. The proof needs more than the one-sided Sylvester bound, and we do not give it. Nothing in this book depends on it. When the two spectra merely avoid each other by δ , interleaved in any pattern, what we have proved is the Frobenius theorem, and through $\|\sin \Theta\|_2 \leq \|\sin \Theta\|_F$ the spectral consequence $\sin \theta_k \leq \|\mathbf{E}\mathbf{Q}_1\|_F/\delta$.

20.9.6 Single eigenvectors

The most common question is about one eigenvector. A simple eigenvalue has a one-dimensional eigenspace, so its unit eigenvector is determined up to a unimodular factor, and the right measure of how far it moves is the angle between the two **lines**. For unit vectors $\mathbf{x}$ and $\tilde{\mathbf{x}}$ we write $\theta(\mathbf{x}, \tilde{\mathbf{x}}) \in [0, \pi/2]$ for that angle, so that $\cos \theta(\mathbf{x}, \tilde{\mathbf{x}}) = |\langle \tilde{\mathbf{x}}, \mathbf{x} \rangle|$. This is the single principal angle between $\text{span}(\mathbf{x})$ and $\text{span}(\tilde{\mathbf{x}})$, since the 1×1 matrix $\mathbf{x}^* \tilde{\mathbf{x}}$ has the singular value $|\mathbf{x}^* \tilde{\mathbf{x}}|$.

The theorem needs a separation between the perturbed eigenvalues of $\tilde{\mathbf{A}}$ and the unperturbed one of $\mathbf{A}$. In practice only $\mathbf{A}$ is known, and Weyl's inequality converts a gap in the spectrum of $\mathbf{A}$ into the separation the theorem wants.

Corollary 20.9.5. Perturbation of a Single Eigenvector

Let $\mathbf{A}, \tilde{\mathbf{A}} = \mathbf{A} + \mathbf{E} \in M_n(\mathbb{C})$ be Hermitian, let $\lambda_i(\mathbf{A})$ be a **simple** eigenvalue of $\mathbf{A}$ with unit eigenvector $\mathbf{x}$, and put

$$\gamma := \min_{j \neq i} |\lambda_i(\mathbf{A}) - \lambda_j(\mathbf{A})| > 0.$$

If $\|\mathbf{E}\|_2 < \gamma/2$, then $\lambda_i(\tilde{\mathbf{A}})$ is a simple eigenvalue of $\tilde{\mathbf{A}}$, and every unit eigenvector $\tilde{\mathbf{x}}$ of $\tilde{\mathbf{A}}$ for it satisfies

$$\sin \theta(\mathbf{x}, \tilde{\mathbf{x}}) \leq \frac{\|\mathbf{E}\|_2}{\gamma - \|\mathbf{E}\|_2} \leq \frac{2\|\mathbf{E}\|_2}{\gamma}.$$

Idea. Apply the theorem with $k = 1$. The only thing to supply is δ : every eigenvalue $\lambda_j(\tilde{\mathbf{A}})$ with $j \neq i$ sits within $\|\mathbf{E}\|_2$ of $\lambda_j(\mathbf{A})$ by Weyl, which is at least γ away from $\lambda_i(\mathbf{A})$. So the separation is at least $\gamma - \|\mathbf{E}\|_2$, by the triangle inequality. The hypothesis $\|\mathbf{E}\|_2 < \gamma/2$ keeps $\lambda_i(\tilde{\mathbf{A}})$ away from its neighbors too, so that “the” eigenvector makes sense.

Proof. Put $\varepsilon = \|\mathbf{E}\|_2$. By Corollary 17.3.6, $|\lambda_j(\tilde{\mathbf{A}}) - \lambda_j(\mathbf{A})| \leq \varepsilon$ for every j . Hence for $j \neq i$, by the triangle inequality in $\mathbb{R}$ (Theorem 0.6.4),

$$\begin{aligned} |\lambda_j(\tilde{\mathbf{A}}) - \lambda_i(\mathbf{A})| &\geq |\lambda_j(\mathbf{A}) - \lambda_i(\mathbf{A})| - |\lambda_j(\tilde{\mathbf{A}}) - \lambda_j(\mathbf{A})| \\ &\geq \gamma - \varepsilon, \end{aligned}$$

and similarly $|\lambda_j(\tilde{\mathbf{A}}) - \lambda_i(\tilde{\mathbf{A}})| \geq \gamma - 2\varepsilon > 0$. So $\lambda_i(\tilde{\mathbf{A}})$ differs from every other $\lambda_j(\tilde{\mathbf{A}})$, and it is a simple eigenvalue of $\tilde{\mathbf{A}}$.

By Corollary 12.4.3 there is a unitary matrix whose columns are eigenvectors of $\tilde{\mathbf{A}}$, with the eigenvalue $\lambda_i(\tilde{\mathbf{A}})$ first. That eigenspace is one-dimensional, so the first column is $c\tilde{\mathbf{x}}$ with $|c| = 1$, and replacing it by $\tilde{\mathbf{x}}$ keeps the columns orthonormal eigenvectors. Call the result $\tilde{\mathbf{Q}} = (\tilde{\mathbf{x}} \quad \tilde{\mathbf{Q}}_2)$. Then $\tilde{\mathbf{\Lambda}}_2$ is diagonal with entries $\lambda_j(\tilde{\mathbf{A}})$, $j \neq i$. In the same way build $\mathbf{Q} = (\mathbf{x} \quad \mathbf{Q}_2)$ for $\mathbf{A}$, with $\mathbf{\Lambda}_1 = (\lambda_i(\mathbf{A}))$. By the display, the hypothesis of [Theorem 20.9.3](#) holds with $k = 1$ and $\delta = \gamma - \varepsilon > 0$, so

$$\sin \theta(\mathbf{x}, \tilde{\mathbf{x}}) = \|\sin \Theta\|_F \leq \frac{\|\mathbf{E}\mathbf{x}\|}{\gamma - \varepsilon} \leq \frac{\varepsilon}{\gamma - \varepsilon},$$

using $\|\mathbf{E}\mathbf{x}\| \leq \|\mathbf{E}\|_2 \|\mathbf{x}\| = \varepsilon$. Finally $\varepsilon < \gamma/2$ gives $\gamma - \varepsilon > \gamma/2$, so $\varepsilon/(\gamma - \varepsilon) \leq 2\varepsilon/\gamma$. This proves the corollary. ■

To first order in $\|\mathbf{E}\|_2$ the bound reads $\sin \theta \lesssim \|\mathbf{E}\|_2/\gamma$: *an eigenvector is as well conditioned as its eigenvalue is isolated*. The constant cannot be improved, and [Exercise \(C2: The constant in the eigenvector bound is sharp\)](#) shows it. The factor 2 in the cruder form is the price of the hypothesis $\|\mathbf{E}\|_2 < \gamma/2$, nothing more. In [Example \(A tiny perturbation that turns the eigenvectors through 45°\)](#), $\gamma = 2\eta$ and $\|\mathbf{E}\|_2 = \sqrt{2}\eta > \gamma/2$, so the corollary does not apply there.

Example. The theorem and the corollary, side by side

Let $\mathbf{A} = \text{diag}(2, 0)$ and $\mathbf{E} = \frac{3}{4} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. Find the exact angle between the eigenvectors of $\mathbf{A}$ and of $\tilde{\mathbf{A}} = \mathbf{A} + \mathbf{E}$ for the larger eigenvalue, and compare it with the bounds of [Theorem 20.9.3](#) and [Corollary 20.9.5](#).

Solution. Exact. $\tilde{\mathbf{A}} = \begin{pmatrix} 2 & 3/4 \\ 3/4 & 0 \end{pmatrix}$ has trace 2 and determinant $-\frac{9}{16}$, so its eigenvalues are $1 \pm \sqrt{1 + 9/16} = 1 \pm \frac{5}{4}$, that is $\frac{9}{4}$ and $-\frac{1}{4}$. For $\frac{9}{4}$, the first row of $\tilde{\mathbf{A}} - \frac{9}{4}\mathbf{I}$ is $(-\frac{1}{4}, \frac{3}{4})$, which is orthogonal to $(3, 1)$, and the second row $(\frac{3}{4}, -\frac{9}{4})$ is a multiple of the first; so $\tilde{\mathbf{x}} = \frac{1}{\sqrt{10}}(3, 1)$. With $\mathbf{x} = \mathbf{e}_1$,

$$\cos \theta = \frac{3}{\sqrt{10}}, \quad \sin \theta = \frac{1}{\sqrt{10}} \approx 0.316.$$

The theorem. Take $\mathbf{\Lambda}_1 = (2)$ and $\tilde{\mathbf{\Lambda}}_2 = (-\frac{1}{4})$, so $\delta = \frac{9}{4}$. Since $\mathbf{E}\mathbf{e}_1 = (0, \frac{3}{4})$, the bound is $\frac{3/4}{9/4} = \frac{1}{3} \approx 0.333$, within six percent of the truth.

The corollary. The gap is $\gamma = 2$, and $\mathbf{E}$ is $\frac{3}{4}$ times a Hermitian matrix with eigenvalues ± 1 , so $\|\mathbf{E}\|_2 = \frac{3}{4} < 1 = \gamma/2$. The bound is $\frac{3/4}{2 - 3/4} = \frac{3}{5} = 0.6$, and its cruder form is $2 \cdot \frac{3/4}{2} = \frac{3}{4}$.

The theorem is nearly exact because it uses the perturbed eigenvalue $-\frac{1}{4}$ itself. The corollary knows only that this eigenvalue lies within $\frac{3}{4}$ of 0, and it must allow for the worst case $\frac{3}{4}$, which would leave a separation of only $\frac{5}{4}$.

20.9.7 Clusters: the subspace is stable when the vectors are not

Example (A tiny perturbation that turns the eigenvectors through 45°) looks like bad news for eigenvectors, and it is, for **individual** eigenvectors of close eigenvalues. But the theorem was stated for subspaces, and that is its real strength. The subspace spanned by the eigenvectors of a cluster of eigenvalues is governed by the gap between the cluster and the rest of the spectrum. The gaps inside the cluster do not enter.

Example. A stable plane with unstable lines inside it

Let $0 < \eta < 1$ and

$$\mathbf{A} = \begin{pmatrix} 1+\eta & 0 & 0 \\ 0 & 1-\eta & 0 \\ 0 & 0 & -2 \end{pmatrix}, \quad \mathbf{E} = \eta \begin{pmatrix} -1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Compare what happens to the eigenvectors for $1 \pm \eta$ with what happens to the plane they span. Then bound the motion of that plane under **any** Hermitian perturbation $\mathbf{E}'$ with $\|\mathbf{E}'\|_2 = \varepsilon < 3 - \eta$.

Solution. The upper-left 2×2 block is [Example \(A tiny perturbation that turns the eigenvectors through \$45^\circ\$ \)](#). So the eigenvalues of $\tilde{\mathbf{A}} = \mathbf{A} + \mathbf{E}$ are again $1 + \eta, 1 - \eta, -2$, and the eigenvectors for $1 \pm \eta$ turn through $\pi/4$, to $\frac{1}{\sqrt{2}}(1, \pm 1, 0)$. But they turn **inside** the plane $U = \text{span}(\mathbf{e}_1, \mathbf{e}_2)$, which is invariant under both matrices. So the plane has not moved at all: both principal angles between U and the new plane $W = U$ are 0.

[Theorem 20.9.3](#) with $k = 2$ sees this. Take $\mathbf{Q}_1 = (\mathbf{e}_1 \ \mathbf{e}_2)$, $\mathbf{\Lambda}_1 = \text{diag}(1 + \eta, 1 - \eta)$, and $\tilde{\mathbf{\Lambda}}_2 = (-2)$. The separation is $\delta = (1 - \eta) - (-2) = 3 - \eta$, and $\|\mathbf{E}\mathbf{Q}_1\|_F = \|\mathbf{E}\|_F = 2\eta$, so

$$\|\sin \Theta\|_F \leq \frac{2\eta}{3 - \eta}.$$

This is small with η , and it has no 2η in its denominator. (The true value is 0; the theorem does not know that $\mathbf{E}$ happens to preserve U .)

For a general Hermitian $\mathbf{E}'$ with $\|\mathbf{E}'\|_2 = \varepsilon$, [Corollary 17.3.6](#) puts the smallest eigenvalue of $\mathbf{A} + \mathbf{E}'$ at most $-2 + \varepsilon$, while the eigenvalues of $\mathbf{\Lambda}_1$ are at least $1 - \eta$. So case (i) of [Theorem 20.9.4](#) holds with $a = 1 - \eta$ and $\delta = 3 - \eta - \varepsilon > 0$, and the plane moves by

$$\sin \theta_2 \leq \frac{\varepsilon}{3 - \eta - \varepsilon}.$$

The individual eigenvectors are covered by [Corollary 20.9.5](#) only when $\varepsilon < \eta$, and then only with a bound of order $\varepsilon/(2\eta)$.

This is the practical lesson of the section. When a matrix has a cluster of close eigenvalues, do not ask for the individual eigenvectors of the cluster. They are ill-conditioned and may be meaningless at the precision of the data. Ask for an orthonormal basis of the invariant subspace of the whole cluster, which is as stable as the cluster is isolated. Numerical eigensolvers act on

this advice.

20.9.8 Exercises

20.9.8.1 A. Check your understanding

Exercise. A1

- State the subspace perturbation equation [Eq. 20.9.1](#), and say which matrix in it measures the principal angles.
- In [Theorem 20.9.3](#), which two lists of eigenvalues must be separated by δ ? Why not the eigenvalues of $\mathbf{A}$ inside and outside the subspace?
- True or false: “if $\|\mathbf{E}\|_2$ is tiny compared with $\|\mathbf{A}\|_2$, then every unit eigenvector of $\mathbf{A} + \mathbf{E}$ is close to a unit eigenvector of $\mathbf{A}$ ”. Justify.
- Why is $\|\sin \Theta\|_2 \leq \|\sin \Theta\|_F$, and when does [Theorem 20.9.4](#) improve on what the Frobenius theorem gives for $\|\sin \Theta\|_2$?
- Which earlier result turns the bound on the matrix $\mathbf{X} = \tilde{\mathbf{Q}}_2^* \mathbf{Q}_1$ into a bound on angles?

Solution.

- $\tilde{\mathbf{\Lambda}}_2 \mathbf{X} - \mathbf{X} \mathbf{\Lambda}_1 = \tilde{\mathbf{Q}}_2^* \mathbf{E} \mathbf{Q}_1$, with $\mathbf{X} = \tilde{\mathbf{Q}}_2^* \mathbf{Q}_1$; by [Lemma 20.9.2](#) the singular values of $\mathbf{X}$ are the sines of the principal angles between $\text{col}(\mathbf{Q}_1)$ and $\text{col}(\tilde{\mathbf{Q}}_1)$, together with zeros.
- The diagonal of $\tilde{\mathbf{\Lambda}}_2$, the eigenvalues of the **perturbed** matrix belonging to the complement, against the diagonal of $\mathbf{\Lambda}_1$, the eigenvalues of the **unperturbed** matrix belonging to the subspace. These are the two coefficient matrices of the Sylvester equation, and the proof divides by their separation, which [Theorem 20.7.2](#) computes as the smallest of those differences. A gap inside the spectrum of $\mathbf{A}$ alone is useful only after [Corollary 17.3.6](#) converts it, as in [Corollary 20.9.5](#).
- False. In [Example \(A tiny perturbation that turns the eigenvectors through \$45^\circ\$ \)](#), $\|\mathbf{E}\|_2 = \sqrt{2}\eta$ can be as small as we like compared with $\|\mathbf{A}\|_2 = 1 + \eta$, yet each unit eigenvector of $\mathbf{A} + \mathbf{E}$ makes the angle $\pi/4$ with every unit eigenvector of $\mathbf{A}$, which are the multiples of $\mathbf{e}_1$ and $\mathbf{e}_2$. What controls eigenvectors is the perturbation relative to the gap, not relative to the matrix.
- $\|\sin \Theta\|_2 = \sin \theta_k$ is one of the numbers whose squares sum to $\|\sin \Theta\|_F^2$. The Frobenius theorem gives $\sin \theta_k \leq \|\mathbf{E} \mathbf{Q}_1\|_F / \delta$, and $\|\mathbf{E} \mathbf{Q}_1\|_F$ can be as large as $\sqrt{k} \|\mathbf{E} \mathbf{Q}_1\|_2$. When the two spectra lie on opposite sides of a gap of width δ , [Theorem 20.9.4](#) replaces it by $\|\mathbf{E} \mathbf{Q}_1\|_2$, which is never worse and can be $\sqrt{k}$ times smaller.
- [Theorem 20.8.4](#), applied with $W = \text{col}(\tilde{\mathbf{Q}}_1)$ and the orthonormal basis $\tilde{\mathbf{Q}}_2$ of $W^\perp$.

20.9.8.2 B. Practice

Exercise. B1: How far does the top eigenvector move?

Let

$$\mathbf{A} = \begin{pmatrix} 5 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \quad \mathbf{E} = \frac{1}{2} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix},$$

and let $\tilde{\mathbf{x}}$ be a unit eigenvector of $\tilde{\mathbf{A}} = \mathbf{A} + \mathbf{E}$ for its largest eigenvalue.

- Compute $\|\mathbf{E}\|_2$ and the gap γ of $\lambda_1(\mathbf{A}) = 5$.
- Use [Corollary 20.9.5](#) to bound $\sin \theta(\mathbf{e}_1, \tilde{\mathbf{x}})$.
- Use [Theorem 20.9.3](#), with the separation estimated by Corollary 17.3.6, to obtain a better bound. Hence show that the angle is less than 13° .
- Check that 2 is an eigenvalue of $\tilde{\mathbf{A}}$ and find the other two. Recompute the bound of (c) with the exact separation.

Solution.

- The matrix $\mathbf{M} = 2\mathbf{E}$ is symmetric with characteristic polynomial $t^3 - 2t = t(t^2 - 2)$, so its eigenvalues are $0, \pm\sqrt{2}$ and $\|\mathbf{E}\|_2 = \frac{1}{2}\sqrt{2} = \frac{\sqrt{2}}{2} \approx 0.707$ by Lemma 17.3.5. The eigenvalues of $\mathbf{A}$ are 5, 2, -1, so $\gamma = 5 - 2 = 3$.
- $\|\mathbf{E}\|_2 \approx 0.707 < 1.5 = \gamma/2$, so the corollary applies:

$$\sin \theta \leq \frac{\sqrt{2}/2}{3 - \sqrt{2}/2} \approx 0.308.$$

- By Corollary 17.3.6, $\lambda_2(\tilde{\mathbf{A}}) \leq 2 + \frac{\sqrt{2}}{2}$ and $\lambda_3(\tilde{\mathbf{A}}) \leq -1 + \frac{\sqrt{2}}{2}$, so both lie at distance at least $\delta = 3 - \frac{\sqrt{2}}{2} \approx 2.293$ from 5. Apply [Theorem 20.9.3](#) with $k = 1$, $\mathbf{Q}_1 = \mathbf{e}_1$, $\Lambda_1 = (5)$, and $\tilde{\Lambda}_2 = \text{diag}(\lambda_2(\tilde{\mathbf{A}}), \lambda_3(\tilde{\mathbf{A}}))$. Since $\mathbf{E}\mathbf{e}_1 = (0, \frac{1}{2}, 0)$,

$$\sin \theta \leq \frac{\|\mathbf{E}\mathbf{e}_1\|}{\delta} = \frac{1/2}{3 - \sqrt{2}/2} \approx 0.218.$$

The gain over (b) comes from $\|\mathbf{E}\mathbf{e}_1\| = \frac{1}{2}$ replacing $\|\mathbf{E}\|_2 \approx 0.707$. Since $\sin 13^\circ \approx 0.225 > 0.218$ and $\sin$ increases on $[0, \pi/2]$, the angle is less than 13° .

- $\tilde{\mathbf{A}} - 2\mathbf{I} = \begin{pmatrix} 3 & 1/2 & 0 \\ 1/2 & 0 & 1/2 \\ 0 & 1/2 & -3 \end{pmatrix}$ sends $(1, -6, -1)$ to $(3 - 3, \frac{1}{2} - \frac{1}{2}, -3 + 3) = \mathbf{0}$, so 2 is an eigenvalue. The other two have sum $\text{tr } \tilde{\mathbf{A}} - 2 = 4$ and product $\det \tilde{\mathbf{A}}/2$, where expanding along the first row gives $\det \tilde{\mathbf{A}} = 5(-2 - \frac{1}{4}) - \frac{1}{2}(-\frac{1}{2}) = -11$. So they are the roots of $t^2 - 4t - \frac{11}{2}$, namely $2 \pm \sqrt{19/2} \approx 5.082, -1.082$. The exact separation of $\{2, 2 - \sqrt{19/2}\}$ from 5 is 3, and the bound becomes $\frac{1/2}{3} = \frac{1}{6} \approx 0.167$. (A direct computation of $\tilde{\mathbf{x}}$ gives $\sin \theta \approx 0.163$.)

Exercise. B2: Which theorem applies?

In each case $\mathbf{A}$ and $\tilde{\mathbf{A}}$ are Hermitian 4×4 matrices, $k = 2$, Λ_1 lists the eigenvalues of $\mathbf{A}$ belonging to $\mathbf{Q}_1$, $\tilde{\Lambda}_2$ lists the eigenvalues of $\tilde{\mathbf{A}}$ belonging to $\tilde{\mathbf{Q}}_2$, and $\|\mathbf{E}\mathbf{Q}_1\|_2 = 0.2$, $\|\mathbf{E}\mathbf{Q}_1\|_F = 0.25$. Determine which of [Theorem 20.9.3](#) and [Theorem 20.9.4](#) apply, and give the best bound on $\sin \theta_k$ that they yield. Justify your answer.

- $\Lambda_1 = \text{diag}(3, 2)$, $\tilde{\Lambda}_2 = \text{diag}(0.4, -1)$.
- $\Lambda_1 = \text{diag}(3, 0)$, $\tilde{\Lambda}_2 = \text{diag}(1.9, 5)$.
- $\Lambda_1 = \text{diag}(2, 1)$, $\tilde{\Lambda}_2 = \text{diag}(2, -1)$.
- $\Lambda_1 = \text{diag}(-1, -2)$, $\tilde{\Lambda}_2 = \text{diag}(0.5, 4)$.

Solution.

- (a) Every eigenvalue in $\mathbf{A}_1$ is ≥ 2 and both in $\tilde{\mathbf{A}}_2$ are $\leq 0.4 = 2 - 1.6$. So case (i) of the one-sided theorem holds with $a = 2$, $\delta = 1.6$, and $\sin \theta_2 \leq 0.2/1.6 = 0.125$. The Frobenius theorem also applies, with the smallest of the four differences, $\delta = 2 - 0.4 = 1.6$, but gives only $\sin \theta_2 \leq \|\sin \Theta\|_F \leq 0.25/1.6 \approx 0.156$.
- (b) The value 1.9 lies **between** 0 and 3, so no point has one list entirely above it and the other entirely below, and the one-sided theorem does not apply. The Frobenius theorem applies with $\delta = \min(1.1, 1.9, 2, 5) = 1.1$, the four differences being $|1.9 - 3|$, $|1.9 - 0|$, $|5 - 3|$, and $|5 - 0|$. It gives $\sin \theta_2 \leq \|\sin \Theta\|_F \leq 0.25/1.1 \approx 0.227$.
- (c) $|2 - 2| = 0$, so no $\delta > 0$ exists and **neither** theorem applies. The check after [Lemma 20.9.2](#) shows why nothing can be said in general: an eigenvalue of $\tilde{\mathbf{A}}$ outside the subspace coincides with one of $\mathbf{A}$ inside it, and then the equation need not determine $\mathbf{X}$.
- (d) Every eigenvalue in $\mathbf{A}_1$ is ≤ -1 and both in $\tilde{\mathbf{A}}_2$ are $\geq 0.5 = -1 + 1.5$, so case (ii) holds with $a = -1$ and $\delta = 1.5$: $\sin \theta_2 \leq 0.2/1.5 \approx 0.133$. The Frobenius theorem, with $\delta = 0.5 - (-1) = 1.5$ again, gives $0.25/1.5 \approx 0.167$.

Exercise. B3: The same bounds for the projections

In the setting of [Theorem 20.9.3](#), let $U = \text{col}(\mathbf{Q}_1)$ and $W = \text{col}(\tilde{\mathbf{Q}}_1)$, with orthogonal projections P_U and P_W . Prove that

$$\|P_U - P_W\|_F \leq \frac{\sqrt{2} \|\mathbf{E}\mathbf{Q}_1\|_F}{\delta},$$

and that under the hypotheses of [Theorem 20.9.4](#), $\|P_U - P_W\|_2 \leq \|\mathbf{E}\mathbf{Q}_1\|_2/\delta$.

Solution. Both subspaces have dimension k , so Theorem 20.8.5 applies: $\|P_U - P_W\|_F = \sqrt{2} \|\sin \Theta\|_F$ and $\|P_U - P_W\|_2 = \sin \theta_k$. By [Theorem 20.9.3](#), $\sqrt{2} \|\sin \Theta\|_F \leq \sqrt{2} \|\mathbf{E}\mathbf{Q}_1\|_F/\delta$, which is the first claim. By [Theorem 20.9.4](#), $\sin \theta_k \leq \|\mathbf{E}\mathbf{Q}_1\|_2/\delta$, which is the second.

20.9.8.3 C. Going deeper

Exercise. C1: A bound that needs only the perturbed matrix

Let $\tilde{\mathbf{A}} \in M_n(\mathbb{C})$ be Hermitian with $\tilde{\mathbf{Q}} = \begin{pmatrix} \tilde{\mathbf{Q}}_1 & \tilde{\mathbf{Q}}_2 \end{pmatrix}$ and $\tilde{\mathbf{A}}_2$ as in the text. Let $\mathbf{Z} \in M_{n \times k}(\mathbb{C})$ be **any** matrix with orthonormal columns, not necessarily eigenvectors of anything, and put

$$\mathbf{M} = \mathbf{Z}^* \tilde{\mathbf{A}} \mathbf{Z} \in M_k(\mathbb{C}), \quad \mathbf{R} = \tilde{\mathbf{A}} \mathbf{Z} - \mathbf{Z} \mathbf{M}.$$

- (a) Prove that if every eigenvalue of the Hermitian matrix $\mathbf{M}$ differs from every diagonal entry of $\tilde{\mathbf{A}}_2$ by at least $\delta > 0$, then the principal angles between $\text{col}(\mathbf{Z})$ and $\text{col}(\tilde{\mathbf{Q}}_1)$ satisfy $\|\sin \Theta\|_F \leq \|\mathbf{R}\|_F/\delta$.
- (b) Deduce: let $\lambda_i(\tilde{\mathbf{A}})$ be a **simple** eigenvalue of $\tilde{\mathbf{A}}$, let $\mathbf{z}$ be a unit vector, $\mu = \mathbf{z}^* \tilde{\mathbf{A}} \mathbf{z}$, and $\mathbf{r} = \tilde{\mathbf{A}} \mathbf{z} - \mu \mathbf{z}$. If every eigenvalue of $\tilde{\mathbf{A}}$ other than $\lambda_i(\tilde{\mathbf{A}})$ lies at distance at least δ from μ , then a unit eigenvector $\tilde{\mathbf{x}}$ for $\lambda_i(\tilde{\mathbf{A}})$ satisfies $\sin \theta(\mathbf{z}, \tilde{\mathbf{x}}) \leq \|\mathbf{r}\|/\delta$.

Hint: diagonalize $\mathbf{M}$ and compute $\tilde{\mathbf{Q}}_2^* \mathbf{R}$ in two ways.

Solution.

- (a) $\mathbf{M}^* = \mathbf{Z}^* \tilde{\mathbf{A}}^* \mathbf{Z} = \mathbf{M}$, so by Corollary 12.4.3 and Theorem 12.2.1, $\mathbf{M} = \mathbf{V} \mathbf{D} \mathbf{V}^*$ with $\mathbf{V} \in M_k(\mathbb{C})$ unitary and $\mathbf{D}$ real diagonal, holding the eigenvalues of $\mathbf{M}$. Put $\mathbf{Z}' = \mathbf{Z} \mathbf{V}$, which has orthonormal columns, since $\mathbf{Z}'^* \mathbf{Z}' = \mathbf{V}^* \mathbf{V} = \mathbf{I}_k$, and the same column space as $\mathbf{Z}$, since $\mathbf{V}$ is invertible. Then $\mathbf{R} \mathbf{V} = \tilde{\mathbf{A}} \mathbf{Z}' - \mathbf{Z} \mathbf{V} \mathbf{D} = \tilde{\mathbf{A}} \mathbf{Z}' - \mathbf{Z}' \mathbf{D}$. As in the proof of Lemma 20.9.2, $\tilde{\mathbf{Q}}_2^* \tilde{\mathbf{A}} = \tilde{\Lambda}_2 \tilde{\mathbf{Q}}_2^*$, so with $\mathbf{X}' = \tilde{\mathbf{Q}}_2^* \mathbf{Z}'$,

$$\tilde{\mathbf{Q}}_2^* \mathbf{R} \mathbf{V} = \tilde{\Lambda}_2 \mathbf{X}' - \mathbf{X}' \mathbf{D}.$$

Both coefficient matrices are real diagonal, hence normal, and their diagonal entries differ by at least δ , so, as in the proof of Theorem 20.9.3, Theorem 20.7.2 and Proposition 20.7.1 (d) give $\|\mathbf{X}'\|_F \leq \|\tilde{\mathbf{Q}}_2^* \mathbf{R} \mathbf{V}\|_F / \delta$. By Lemma 20.9.1 (c), $\|\tilde{\mathbf{Q}}_2^* \mathbf{R} \mathbf{V}\|_F \leq \|\mathbf{R} \mathbf{V}\|_F = \|\mathbf{R}\|_F$. Finally, by Theorem 20.8.4 applied as in Lemma 20.9.2, with $\text{col}(\mathbf{Z}') = \text{col}(\mathbf{Z})$ in place of $\text{col}(\mathbf{Q}_1)$, $\|\mathbf{X}'\|_F = \|\sin \Theta\|_F$. This proves (a).

- (b) Take $k = 1$ and $\mathbf{Z} = \mathbf{z}$. Then $\mathbf{M} = (\mu)$, $\mathbf{R} = \mathbf{r}$, and the only eigenvalue of $\mathbf{M}$ is μ . Since $\lambda_i(\tilde{\mathbf{A}})$ is simple, its eigenspace is the line $\text{span}(\tilde{\mathbf{x}})$, so we may choose $\tilde{\mathbf{Q}}$ with first column $\tilde{\mathbf{x}}$ and $\tilde{\Lambda}_2$ holding the other eigenvalues of $\tilde{\mathbf{A}}$, as in the proof of Corollary 20.9.5. Part (a) gives $\sin \theta(\mathbf{z}, \tilde{\mathbf{x}}) \leq \|\mathbf{r}\| / \delta$. This is the eigenvector companion of a residual bound: everything on the right is computable from $\mathbf{z}$ and $\tilde{\mathbf{A}}$ and a gap estimate, with no unperturbed matrix in sight.

Exercise. C2: The constant in the eigenvector bound is sharp

Let $g > 0$ and $e > 0$, and let $\mathbf{A} = \text{diag}(g, 0)$, $\mathbf{E} = e \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

- (a) Show that a unit eigenvector $\tilde{\mathbf{x}}$ of $\mathbf{A} + \mathbf{E}$ for its larger eigenvalue makes an angle $\theta \in [0, \pi/4)$ with $\mathbf{e}_1$ satisfying $\tan 2\theta = 2e/g$.
 (b) Deduce that $\sin \theta \cdot g/e \rightarrow 1$ as $e \rightarrow 0^+$.
 (c) Hence show that for no constant $c < 1$ is it true that $\sin \theta(\mathbf{x}, \tilde{\mathbf{x}}) \leq c \|\mathbf{E}\|_2 / \gamma$ for all sufficiently small Hermitian $\mathbf{E}$, in the setting of Corollary 20.9.5.

Solution.

- (a) The larger eigenvalue of $\tilde{\mathbf{A}} = \begin{pmatrix} g & e \\ e & 0 \end{pmatrix}$ is $\lambda_+ = h + s$, where $h = g/2$ and $s = \sqrt{h^2 + e^2}$. The first row of $\tilde{\mathbf{A}} - \lambda_+ \mathbf{I}$ is $(g - \lambda_+, e) = (h - s, e)$, so $(e, s - h)$ spans the eigenline, and $s - h > 0$. The angle with $\mathbf{e}_1$ therefore satisfies $\tan \theta = t := (s - h)/e > 0$. Using $s^2 = h^2 + e^2$,

$$1 - t^2 = \frac{e^2 - (s - h)^2}{e^2} = \frac{2h(s - h)}{e^2} > 0,$$

so $t < 1$, that is $\theta < \pi/4$, and

$$\begin{aligned} \tan 2\theta &= \frac{2t}{1 - t^2} = \frac{2(s - h)}{e} \cdot \frac{e^2}{2h(s - h)} \\ &= \frac{e}{h} = \frac{2e}{g}. \end{aligned}$$

(b) From (a), $\cos 2\theta = 1/\sqrt{1 + \tan^2 2\theta} = g/\sqrt{g^2 + 4e^2}$, since $2\theta \in [0, \pi/2)$. Hence

$$\begin{aligned} \sin^2 \theta &= \frac{1 - \cos 2\theta}{2} = \frac{\sqrt{g^2 + 4e^2} - g}{2\sqrt{g^2 + 4e^2}} \\ &= \frac{2e^2}{\sqrt{g^2 + 4e^2}(\sqrt{g^2 + 4e^2} + g)}, \end{aligned}$$

multiplying above and below by $\sqrt{g^2 + 4e^2} + g$. So

$$\left(\frac{g \sin \theta}{e}\right)^2 = \frac{2g^2}{\sqrt{g^2 + 4e^2}(\sqrt{g^2 + 4e^2} + g)},$$

which tends to $2g^2/(g \cdot 2g) = 1$ as $e \rightarrow 0^+$, by continuity of the square root.

(c) Here $\mathbf{x} = \mathbf{e}_1$, $\gamma = g$, and $\|\mathbf{E}\|_2 = e$, since $\mathbf{E}/e$ is Hermitian with eigenvalues ± 1 . If $\sin \theta \leq ce/g$ held for all small $e > 0$ with some $c < 1$, then $\sin \theta \cdot g/e \leq c$ for those e , contradicting (b). So to first order in $\|\mathbf{E}\|_2$, the bound $\|\mathbf{E}\|_2/(\gamma - \|\mathbf{E}\|_2)$ of [Corollary 20.9.5](#) cannot be improved by any constant factor.