

12.7 Commuting Families

Chapter 9 asked when several operators can be diagonalized in one basis and answered: exactly when they commute, provided each of them is diagonalizable on its own. With an inner product in hand the question sharpens in the same way the spectral theorem sharpened diagonalization. Can one **orthonormal** basis diagonalize all of them at once? For normal operators over $\mathbb{C}$ the answer is yes, and the proof is the chapter's signature move applied family-wide: cut the space into the eigenspaces of one member and recurse.

Throughout, V is a finite-dimensional inner product space over $F = \mathbb{R}$ or $\mathbb{C}$, and a family $\mathcal{F} \subseteq \mathcal{L}(V)$ is called **commuting** if $ST = TS$ for all $S, T \in \mathcal{F}$, as in Chapter 9.

12.7.1 What Chapter 9 already gives, and what is missing

Three results from Chapter 9 do the bookkeeping, and it is worth being exact about what each one delivers.

- Theorem 9.9.1: if $ST = TS$ then every eigenspace $E_\lambda(T)$ is S -invariant. No hypothesis at all beyond commuting — not even finite dimension. This is what lets us look for common eigenvectors *inside* an eigenspace.
- Theorem 9.9.3: a commuting family whose characteristic polynomials all split has a common eigenvector. Over $\mathbb{C}$ the splitting is automatic.
- Theorem 9.9.5 and Theorem 9.9.6: a commuting family of **diagonalizable** operators is simultaneously diagonalizable (Definition 9.9.1), and conversely.

Two things are missing from that list. First, diagonalizability is a hypothesis there, and it has to be checked separately for each member. Second, and more to the point here, the common basis produced is just a basis: nothing says its vectors are orthogonal, and in general they are not.

Normality repairs both at once, over $\mathbb{C}$. It repairs the first because a normal operator is automatically diagonalizable, by Theorem 12.4.2. It repairs the second because that same theorem produces an *orthonormal* eigenbasis, so the pieces we cut the space into are mutually orthogonal.

One more ingredient is needed before the recursion can run: after cutting V into the eigenspaces of one member, the other members must still be normal on each piece. Section 3 supplies exactly this. An invariant subspace of a normal operator is automatically invariant under the adjoint as well, and the restriction to it is again normal: that is Proposition 12.3.6, proved there by a trace computation on the block form. Over $\mathbb{R}$ the self-adjoint version is Theorem 12.2.4. (What fails, as the warning of Section 3 records, is the analogous statement for a subspace that is **not** invariant: a principal submatrix of a normal matrix need not be normal.) With that recalled, nothing stands in the way.

12.7.2 One orthonormal basis for the whole family

Theorem 12.7.1. One Orthonormal Eigenbasis for a Commuting Family

Let $V \neq \{0\}$ be a finite-dimensional **complex** inner product space and let $\mathcal{F} \subseteq \mathcal{L}(V)$ be a nonempty commuting family of **normal** operators. Then V has an orthonormal basis $\mathcal{B}$ each of whose vectors is an eigenvector of every $T \in \mathcal{F}$; equivalently, $[T]_{\mathcal{B}}^{\mathcal{B}}$ is diagonal for

every $T \in \mathcal{F}$.

Conversely, if some orthonormal basis $\mathcal{B}$ makes every $[T]_{\mathcal{B}}^{\mathcal{B}}$ diagonal, then $\mathcal{F}$ is commuting and every member of $\mathcal{F}$ is normal.

Idea. Induction on $\dim V$, with the two cases that every such induction has. If every member of the family is a scalar multiple of the identity, any orthonormal basis will do and there is nothing to arrange. Otherwise some $T_0 \in \mathcal{F}$ has at least two distinct eigenvalues, and Theorem 12.4.2 splits V into the eigenspaces of T_0 — **orthogonally**, and into pieces strictly smaller than V . Each piece is invariant under the whole family because the operators commute, and the restriction of a normal operator to an invariant subspace is normal, so each restricted family is again a commuting family of normal operators on a smaller space. Run the induction on each piece and glue.

Proof. Induction on $n = \dim V \geq 1$.

If $n = 1$, any unit vector v of V is a basis, and $Tv \in V = \text{span}(v)$ makes v an eigenvector of every $T \in \mathcal{L}(V)$.

Let $n \geq 2$ and assume the theorem for all complex inner product spaces of dimension less than n .

Case 1: every $T \in \mathcal{F}$ is a scalar multiple of id_V . Then every nonzero vector is an eigenvector of every $T \in \mathcal{F}$, and V has an orthonormal basis by Theorem 11.2.2.

Case 2: some $T_0 \in \mathcal{F}$ is not a scalar multiple of id_V . Since T_0 is normal, Theorem 12.4.2 gives an orthonormal basis $\mathcal{C} = (c_1, \dots, c_n)$ of V with $T_0 c_j = \mu_j c_j$. Let $\lambda_1, \dots, \lambda_k$ be the distinct values among $\mu_1, \dots, \mu_n$, and put

$$V_i = \text{span}\{c_j : \mu_j = \lambda_i\}, \quad i = 1, \dots, k.$$

The V_i are spanned by disjoint sub-lists of an orthonormal basis, so they are pairwise orthogonal, $V = V_1 \oplus \dots \oplus V_k$, and $\dim V_1 + \dots + \dim V_k = n$.

Claim. $V_i = E_{\lambda_i}(T_0)$ for each i .

Proof. The inclusion $V_i \subseteq E_{\lambda_i}(T_0)$ holds because every spanning vector c_j of V_i satisfies $T_0 c_j = \lambda_i c_j$. Conversely let $v \in E_{\lambda_i}(T_0)$ and write $v = v_1 + \dots + v_k$ with $v_j \in V_j$. Applying T_0 , which acts as λ_j on V_j , gives $\sum_j \lambda_j v_j = T_0 v = \lambda_i v = \sum_j \lambda_i v_j$, so $\sum_j (\lambda_j - \lambda_i) v_j = 0$. The sum being direct, each term vanishes, and $\lambda_j \neq \lambda_i$ for $j \neq i$ forces $v_j = 0$. Hence $v = v_i \in V_i$. ■

Because T_0 is not a scalar multiple of id_V , we have $k \geq 2$: if $k = 1$ then $V = V_1 = E_{\lambda_1}(T_0)$ and $T_0 = \lambda_1 \text{id}_V$. Hence $\dim V_i < n$ for every i .

Fix i and let $S \in \mathcal{F}$. By Theorem 9.9.1, $V_i = E_{\lambda_i}(T_0)$ is S -invariant, and therefore $S|_{V_i}$ is a normal operator on V_i by Proposition 12.3.6. Moreover the family

$$\mathcal{F}_i = \{S|_{V_i} : S \in \mathcal{F}\}$$

is commuting, since $(S|_{V_i})(T|_{V_i}) = (ST)|_{V_i} = (TS)|_{V_i} = (T|_{V_i})(S|_{V_i})$ for $S, T \in \mathcal{F}$, all four restrictions being defined because V_i is invariant.

By the inductive hypothesis applied to $V_i \neq \{0\}$ and $\mathcal{F}_i$, there is an orthonormal basis $\mathcal{B}_i$ of V_i whose vectors are eigenvectors of every $S|_{V_i}$; such a vector $v \in V_i$ satisfies $Sv = (S|_{V_i})v = \mu v$, so it is an eigenvector of S itself. Let $\mathcal{B}$ be the concatenation $\mathcal{B}_1, \dots, \mathcal{B}_k$. It is orthonormal, because each $\mathcal{B}_i$ is and because $V_i \perp V_j$ for $i \neq j$, and it has $\sum_i \dim V_i = n$ vectors, so it is an orthonormal basis of V (Theorem 11.1.4 and Theorem 1.6.4). Every vector of $\mathcal{B}$ is an eigenvector of every $T \in \mathcal{F}$. This completes the induction.

For the converse, suppose every $[T]_{\mathcal{B}}^{\mathcal{B}}$ is diagonal with $\mathcal{B}$ orthonormal. Diagonal matrices commute, so $[ST]_{\mathcal{B}}^{\mathcal{B}} = [TS]_{\mathcal{B}}^{\mathcal{B}}$ and hence $ST = TS$ (Theorem 4.1.4). And $[T^*]_{\mathcal{B}}^{\mathcal{B}} = ([T]_{\mathcal{B}}^{\mathcal{B}})^*$ by Theorem 11.6.3 is diagonal as well, so $[T^*T]_{\mathcal{B}}^{\mathcal{B}}$ and $[TT^*]_{\mathcal{B}}^{\mathcal{B}}$ are products of diagonal matrices in the two orders, hence equal; so T is normal. This proves the theorem. ■

Notice which hypothesis did which job. *Commuting* made the eigenspaces of T_0 invariant under the rest of the family; *normal* made those eigenspaces orthogonal and made the restrictions normal again; and *complex* is what makes T_0 have any eigenvalues at all. Drop any one of the three and the argument stops.

In matrix language:

Corollary 12.7.2. Commuting Normal Matrices

Let $\mathbf{A}_1, \dots, \mathbf{A}_r \in M_n(\mathbb{C})$ be normal and pairwise commuting. Then there is a unitary $\mathbf{U} \in U(n)$ such that $\mathbf{U}^* \mathbf{A}_i \mathbf{U}$ is diagonal for every i . The same holds for an arbitrary commuting family of normal matrices in $M_n(\mathbb{C})$.

Proof. Apply Theorem 12.7.1 to the family $\{T_{\mathbf{A}_1}, \dots, T_{\mathbf{A}_r}\}$ on $\mathbb{C}^n$. Each $T_{\mathbf{A}_i}$ is normal because $T_{\mathbf{A}_i}^* = T_{\mathbf{A}_i^*}$ (Theorem 11.6.3, the standard basis being orthonormal) and $\mathbf{A}_i^* \mathbf{A}_i = \mathbf{A}_i \mathbf{A}_i^*$; and the operators commute because the matrices do. Let $\mathbf{U}$ have the resulting orthonormal basis $\mathcal{B}$ as its columns; then $\mathbf{U} \in U(n)$ by Theorem 11.7.3 (f), and $\mathbf{U}^* \mathbf{A}_i \mathbf{U} = \mathbf{U}^{-1} \mathbf{A}_i \mathbf{U} = [T_{\mathbf{A}_i}]_{\mathcal{B}}^{\mathcal{B}}$ is diagonal by Theorem 4.2.3. Nothing in the argument used that the family is finite. ■

Corollary 12.7.3. Commuting Symmetric Matrices

Let V be a finite-dimensional **real** inner product space and let $\mathcal{F} \subseteq \mathcal{L}(V)$ be a nonempty commuting family of self-adjoint operators. Then V has an orthonormal basis of common eigenvectors. Equivalently, if $\mathbf{A}_1, \dots, \mathbf{A}_r \in M_n(\mathbb{R})$ are symmetric and pairwise commuting, there is $\mathbf{Q} \in O(n)$ with every $\mathbf{Q}^T \mathbf{A}_i \mathbf{Q}$ diagonal.

Proof. Repeat the proof of Theorem 12.7.1 word for word, with two substitutions: use Theorem 12.5.1 in place of Theorem 12.4.2 to split V into the eigenspaces of a member T_0 that is not a scalar multiple of id_V , and use Theorem 12.2.4 (b) in place of Proposition 12.3.6 to see that each restriction $S|_{V_i}$ is again self-adjoint. Every other step is unchanged, and the only property of $\mathbb{C}$ used in the original was the existence of an eigenvalue, which Theorem 12.5.1 supplies over $\mathbb{R}$ from self-adjointness. The matrix statement follows as in Corollary 12.7.2,

using Corollary 12.5.2 and $\mathbf{Q}^\top = \mathbf{Q}^{-1}$.

■

Example. Two Commuting Symmetric Matrices, Diagonalized Together

Let

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix} \in M_3(\mathbb{R}).$$

Check that $\mathbf{AB} = \mathbf{BA}$ and find $\mathbf{Q} \in O(3)$ with $\mathbf{Q}^\top \mathbf{A} \mathbf{Q}$ and $\mathbf{Q}^\top \mathbf{B} \mathbf{Q}$ both diagonal.

Solution. They commute. Multiplying out,

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

Both matrices are symmetric, so Corollary 12.7.3 applies.

Diagonalize A. Each row of $\mathbf{A}$ sums to 2, so $\mathbf{A}(1, 1, 1) = (2, 2, 2)$ and 2 is an eigenvalue with eigenvector $(1, 1, 1)$. Also $\mathbf{A} = \mathbf{1}\mathbf{1}^\top - \mathbf{I}_3$, where $\mathbf{1} = (1, 1, 1)$; any $\mathbf{v}$ with $\mathbf{1}^\top \mathbf{v} = 0$, that is with $v_1 + v_2 + v_3 = 0$, satisfies $\mathbf{A}\mathbf{v} = -\mathbf{v}$. So $E_2(\mathbf{A}) = \text{span}((1, 1, 1))$ and $E_{-1}(\mathbf{A})$ is the plane $v_1 + v_2 + v_3 = 0$, and these account for all of $\mathbb{R}^3$.

The one-dimensional eigenspace. By Theorem 9.9.1 the line $E_2(\mathbf{A})$ is $\mathbf{B}$ -invariant, so its spanning vector must already be an eigenvector of $\mathbf{B}$. Indeed $\mathbf{B}(1, 1, 1) = (2, 2, 2)$.

The plane. Inside $E_{-1}(\mathbf{A})$ we must search. The matrix $\mathbf{B}$ has $\mathbf{B}(0, 1, -1) = (0, 0, 0)$ and $\mathbf{B}(-2, 1, 1) = (-4, 2, 2) = 2(-2, 1, 1)$, and both $(0, 1, -1)$ and $(-2, 1, 1)$ have coordinate sum 0, so they lie in $E_{-1}(\mathbf{A})$. They are orthogonal, and independent, so they are a basis of the plane.

The common basis. Normalizing the three vectors,

$$\mathbf{Q} = \begin{pmatrix} 1/\sqrt{3} & -2/\sqrt{6} & 0 \\ 1/\sqrt{3} & 1/\sqrt{6} & 1/\sqrt{2} \\ 1/\sqrt{3} & 1/\sqrt{6} & -1/\sqrt{2} \end{pmatrix},$$

whose columns are orthonormal, so $\mathbf{Q} \in O(3)$ by Theorem 11.7.3 (f). Then

$$\mathbf{Q}^\top \mathbf{A} \mathbf{Q} = \text{diag}(2, -1, -1), \quad \mathbf{Q}^\top \mathbf{B} \mathbf{Q} = \text{diag}(2, 2, 0),$$

the diagonal entries being the eigenvalues found above, column by column. Neither matrix alone forces the choice: $\mathbf{A}$ cannot tell the second column from the third, and $\mathbf{B}$ cannot tell the first from the second. Together they pin down all three lines.

Warning. Neither hypothesis can be dropped.

Commuting is not enough. The matrices $\mathbf{A} = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}$ and $\mathbf{B} = \begin{pmatrix} 1 & 3 \\ 0 & 4 \end{pmatrix}$ satisfy $\mathbf{AB} = \begin{pmatrix} 1 & 7 \\ 0 & 8 \end{pmatrix} = \mathbf{BA}$, and each is diagonalizable, having two distinct eigenvalues. They even share the eigenbasis $\mathbf{e}_1, (1, 1)$: $\mathbf{A}(1, 1) = 2(1, 1)$ and $\mathbf{B}(1, 1) = 4(1, 1)$. But that basis is not orthogonal, and no orthonormal one exists, because an orthonormal eigenbasis would make $\mathbf{A}$ normal (Corollary 12.4.3), whereas

$$\mathbf{A}^* \mathbf{A} = \begin{pmatrix} 1 & 1 \\ 1 & 5 \end{pmatrix} \neq \begin{pmatrix} 2 & 2 \\ 2 & 4 \end{pmatrix} = \mathbf{A} \mathbf{A}^*.$$

Normal is not enough. The matrices $\mathbf{X} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ and $\mathbf{Z} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ are real symmetric, hence normal, and each is orthogonally diagonalizable on its own. They do not commute:

$$\mathbf{XZ} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \neq \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \mathbf{ZX}.$$

In fact they share no eigenvector at all: the eigenvectors of $\mathbf{Z}$ are the multiples of $\mathbf{e}_1$ and of $\mathbf{e}_2$, and $\mathbf{X}\mathbf{e}_1 = \mathbf{e}_2$ is a multiple of neither.

12.7.3 Sums and products stay normal

Normality is not preserved by addition or multiplication. It is preserved by both as soon as the two operators commute, and the common eigenbasis makes this a one-line matter.

Theorem 12.7.4. Sums and Products of Commuting Normal Operators

Let V be a finite-dimensional complex inner product space and let $S, T \in \mathcal{L}(V)$ be normal with $ST = TS$. Then $S + T$ and ST are normal.

Proof. If $V = \{0\}$ there is nothing to prove. Otherwise apply [Theorem 12.7.1](#) to the family $\{S, T\}$: there is an orthonormal basis $\mathcal{B}$ with $[S]_{\mathcal{B}}^{\mathcal{B}} = \mathbf{D}$ and $[T]_{\mathcal{B}}^{\mathcal{B}} = \mathbf{D}'$ both diagonal. Then $[S + T]_{\mathcal{B}}^{\mathcal{B}} = \mathbf{D} + \mathbf{D}'$ and $[ST]_{\mathcal{B}}^{\mathcal{B}} = \mathbf{D}\mathbf{D}'$ are diagonal, so every vector of $\mathcal{B}$ is an eigenvector of both operators, and both are normal by Theorem 12.4.2. ■

Remark. Commutativity is essential, not a convenience. Take $\mathbf{M} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, which is symmetric, and $\mathbf{N} = \text{diag}(1, 2i)$, which is diagonal; both are normal and they do not commute. Then

$$\mathbf{M} + \mathbf{N} = \begin{pmatrix} 1 & 1 \\ 1 & 2i \end{pmatrix}, \quad \mathbf{MN} = \begin{pmatrix} 0 & 2i \\ 1 & 0 \end{pmatrix},$$

and neither is normal: for the sum, $(\mathbf{M} + \mathbf{N})(\mathbf{M} + \mathbf{N})^*$ has $(1, 2)$ entry $1 - 2i$ while

$(\mathbf{M} + \mathbf{N})^*(\mathbf{M} + \mathbf{N})$ has $1 + 2i$ there; for the product, $(\mathbf{MN})(\mathbf{MN})^* = \text{diag}(4, 1)$ while $(\mathbf{MN})^*(\mathbf{MN}) = \text{diag}(1, 4)$.

Quick check. Let $\mathbf{A} \in M_n(\mathbb{C})$ be normal and let $p \in \mathbb{C}[x]$. Explain, without using [Theorem 12.7.1](#), why $\mathbf{A}$ and $p(\mathbf{A})$ are simultaneously unitarily diagonalizable.

Solution. By Corollary 12.4.3, $\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{U}^*$ with $\mathbf{U}$ unitary and $\mathbf{D}$ diagonal. Then $\mathbf{A}^m = \mathbf{U}\mathbf{D}^m\mathbf{U}^*$ for every $m \geq 0$, since the inner factors $\mathbf{U}^*\mathbf{U}$ cancel, so $p(\mathbf{A}) = \mathbf{U}p(\mathbf{D})\mathbf{U}^*$, and $p(\mathbf{D}) = \text{diag}(p(d_1), \dots, p(d_n))$ is diagonal. The *same* $\mathbf{U}$ diagonalizes both. This is consistent with the theorem, since $\mathbf{A}$ and $p(\mathbf{A})$ commute and $p(\mathbf{A})$ is normal, being unitarily diagonalizable.

12.7.4 Without normality: simultaneous triangularization

If we ask only for triangular matrices, normality may be dropped entirely. Chapter 9 proved this with ordinary bases (Theorem 9.9.4); over $\mathbb{C}$ an inner product upgrades the basis to an orthonormal one, exactly as Theorem 12.1.2 upgraded Theorem 9.7.2.

Theorem 12.7.5. Simultaneous Unitary Triangularization

Let $V \neq \{0\}$ be a finite-dimensional complex inner product space and let $\mathcal{F} \subseteq \mathcal{L}(V)$ be a nonempty commuting family. Then there is an orthonormal basis $\mathcal{B}$ of V such that $[T]_{\mathcal{B}}^{\mathcal{B}}$ is upper triangular for **every** $T \in \mathcal{F}$. In particular, if $\mathbf{A}_1, \dots, \mathbf{A}_r \in M_n(\mathbb{C})$ commute pairwise, there is $\mathbf{U} \in \mathbf{U}(n)$ with every $\mathbf{U}^*\mathbf{A}_i\mathbf{U}$ upper triangular.

Idea. The same swap that proves Schur's theorem (Theorem 12.1.2). A common eigenvector v of $\mathcal{F}$ gives an invariant line, but $\text{span}(v)^\perp$ need not be invariant, so the recursion has nowhere to go. Take instead a common eigenvector u of the **adjoints** T^* : they commute too, and $\text{span}(u)^\perp$ is invariant under every $T \in \mathcal{F}$. Recurse there, and put u at the **end** of the basis, where an upper triangular matrix allows an arbitrary column.

Proof. Induction on $n = \dim V \geq 1$. For $n = 1$ every 1×1 matrix is upper triangular, and a unit vector is an orthonormal basis.

Let $n \geq 2$ and assume the theorem in dimension $n - 1$. The family $\mathcal{F}^* = \{T^* : T \in \mathcal{F}\}$ is commuting: for $S, T \in \mathcal{F}$, $S^*T^* = (TS)^* = (ST)^* = T^*S^*$ by Theorem 11.6.2. Since V is a complex space, Theorem 9.9.3 provides $u \neq 0$ with $T^*u = \nu_T u$ for every $T \in \mathcal{F}$; dividing by $\|u\|$, take $\|u\| = 1$.

Put $W = \text{span}(u)$ and $U = W^\perp$, of dimension $n - 1$ by Theorem 11.3.2 (c). For each $T \in \mathcal{F}$ the line W is T^* -invariant, since $T^*u = \nu_T u$, so U is T -invariant by Lemma 12.1.1. The family $\{T|_U : T \in \mathcal{F}\}$ is commuting, by the same computation on restrictions as in

the proof of [Theorem 12.7.1](#). By the inductive hypothesis there is an orthonormal basis $\mathcal{B}_U = (v_1, \dots, v_{n-1})$ of U such that $[T|_U]_{\mathcal{B}_U}^{\mathcal{B}_U}$ is upper triangular for every $T \in \mathcal{F}$.

Let $\mathcal{B} = (v_1, \dots, v_{n-1}, u)$. It is orthonormal, since u is a unit vector orthogonal to U , and it has n vectors, so it is an orthonormal basis of V . Fix $T \in \mathcal{F}$. For $j \leq n-1$, $Tv_j = T|_U v_j \in \text{span}(v_1, \dots, v_j)$, because $[T|_U]_{\mathcal{B}_U}^{\mathcal{B}_U}$ is upper triangular; so the j -th column of $[T]_{\mathcal{B}}^{\mathcal{B}}$ has zeros below row j . The last column is unconstrained, and an upper triangular matrix permits any entries there. Hence $[T]_{\mathcal{B}}^{\mathcal{B}}$ is upper triangular, which completes the induction.

For the matrix statement, apply this to the operators $T_{\mathbf{A}_i}$ on $\mathbb{C}^n$ and let $\mathbf{U}$ have the vectors of $\mathcal{B}$ as columns, as in [Corollary 12.7.2](#). ■

The order matters: the common eigenvector of the adjoints goes last, and the triangular shape is built from the bottom right corner upward. Running the same recursion with a common eigenvector of $\mathcal{F}$ itself and putting it first would work only if $\text{span}(v)^\perp$ were invariant, which is the extra fact that normality supplies and a bare commuting family does not.

12.7.5 Exercises

12.7.5.1 A. Check your understanding

Exercise. A1

- State [Theorem 12.7.1](#), naming every hypothesis.
- Which result of Chapter 9 makes the eigenspaces of one member invariant under the others, and which hypothesis does it use?
- Why is the restriction of a normal operator to an invariant subspace again normal, and which earlier result says so? Why does the same fail for a subspace that is not invariant?
- True or false: two commuting diagonalizable matrices are simultaneously **unitarily** diagonalizable. Justify your answer.
- True or false: if S and T are normal then $S + T$ is normal. Justify your answer.
- Which hypothesis of [Theorem 12.7.1](#) is dropped in [Theorem 12.7.5](#), and what is lost in the conclusion?

Solution.

- Let $V \neq \{0\}$ be a finite-dimensional **complex** inner product space and $\mathcal{F} \subseteq \mathcal{L}(V)$ a **nonempty** family of **normal** operators that **commute pairwise**. Then V has an orthonormal basis of common eigenvectors.
- Theorem 9.9.1: if $ST = TS$, then $E_\lambda(T)$ is S -invariant. The only hypothesis is that the two operators commute.
- Proposition 12.3.6: in an orthonormal basis adapted to U the matrix of a normal T is $\mathbf{A} = \begin{pmatrix} \mathbf{B} & \mathbf{C} \\ \mathbf{O} & \mathbf{D} \end{pmatrix}$, and comparing the traces of the $(1,1)$ blocks of $\mathbf{A}^* \mathbf{A}$ and $\mathbf{A} \mathbf{A}^*$ forces $\mathbf{C} = \mathbf{O}$; the matrix is then block diagonal and its first block is normal. For a subspace that is not invariant the zero block below is absent and the argument collapses: the

Section 3 warning compresses the 3×3 cyclic shift to $\text{span}(\mathbf{e}_1, \mathbf{e}_2)$ and gets a nonnormal 2×2 matrix.

- (d) False. $\begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}$ and $\begin{pmatrix} 1 & 3 \\ 0 & 4 \end{pmatrix}$ commute and are diagonalizable, but the first is not normal, so it has no orthonormal eigenbasis at all (Corollary 12.4.3). Simultaneous diagonalization by an invertible matrix is all that Chapter 9 promises (Theorem 9.9.5).
- (e) False in general, true when they commute (Theorem 12.7.4). For a counterexample take $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ and $\text{diag}(1, 2i)$, as in the Remark after that theorem.
- (f) Normality. What is lost is that the common matrices are only upper triangular, not diagonal.

12.7.5.2 B. Practice

Exercise. B1: Diagonalize two at once

Let

$$\mathbf{A} = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 2 & 0 & -2 \\ 0 & 2 & -2 \\ -2 & -2 & 4 \end{pmatrix} \in M_3(\mathbb{R}).$$

- (a) Verify that $\mathbf{AB} = \mathbf{BA}$ and that both are symmetric.
- (b) Find $\mathbf{Q} \in O(3)$ with $\mathbf{Q}^\top \mathbf{A} \mathbf{Q}$ and $\mathbf{Q}^\top \mathbf{B} \mathbf{Q}$ both diagonal, and write down the two diagonal matrices.
- (c) Explain why $\mathbf{A}$ on its own does not determine $\mathbf{Q}$ up to permuting and rescaling columns, but the pair does.

Solution.

- (a) Both are equal to their transposes by inspection. Each row of $\mathbf{B}$ sums to 0 and each column too, so $\mathbf{1}\mathbf{1}^\top \mathbf{B} = \mathbf{O} = \mathbf{B}\mathbf{1}\mathbf{1}^\top$ where $\mathbf{1} = (1, 1, 1)$. Since $\mathbf{A} = \mathbf{1}\mathbf{1}^\top + \mathbf{I}_3$, we get $\mathbf{AB} = \mathbf{B} = \mathbf{BA}$. (Direct multiplication confirms $\mathbf{AB} = \mathbf{B}$.)
- (b) The eigenvectors of $\mathbf{A} = \mathbf{1}\mathbf{1}^\top + \mathbf{I}_3$ are $\mathbf{1}$, with $\mathbf{A}\mathbf{1} = 3\mathbf{1} + \mathbf{1} = 4\mathbf{1}$, and every $\mathbf{v}$ with $v_1 + v_2 + v_3 = 0$, with $\mathbf{A}\mathbf{v} = \mathbf{v}$. Inside that plane we search with $\mathbf{B}$:

$$\begin{aligned} \mathbf{B}(1, -1, 0) &= (2, -2, 0) = 2(1, -1, 0), \\ \mathbf{B}(1, 1, -2) &= (6, 6, -12) = 6(1, 1, -2), \end{aligned}$$

and both vectors have coordinate sum 0. Also $\mathbf{B}\mathbf{1} = \mathbf{O}$. The three vectors $\mathbf{1}$, $(1, -1, 0)$, $(1, 1, -2)$ are pairwise orthogonal, as one checks from $1 - 1 + 0 = 0$, $1 + 1 - 2 = 0$, and $1 - 1 + 0 = 0$. Normalizing,

$$\mathbf{Q} = \begin{pmatrix} 1/\sqrt{3} & 1/\sqrt{2} & 1/\sqrt{6} \\ 1/\sqrt{3} & -1/\sqrt{2} & 1/\sqrt{6} \\ 1/\sqrt{3} & 0 & -2/\sqrt{6} \end{pmatrix},$$

and $\mathbf{Q}^\top \mathbf{A} \mathbf{Q} = \text{diag}(4, 1, 1)$, $\mathbf{Q}^\top \mathbf{B} \mathbf{Q} = \text{diag}(0, 2, 6)$.

- (c) The eigenvalue 1 of $\mathbf{A}$ has a two-dimensional eigenspace, so *every* orthonormal basis of that plane serves $\mathbf{A}$ equally well, and there are infinitely many, not obtained from

one another by permuting and rescaling columns. The matrix $\mathbf{B}$ has three distinct eigenvalues, so its eigenlines are determined, and the pair therefore determines the three lines of $\mathbf{Q}$ up to order and sign.

Exercise. B2: Check the two warnings

- (a) For $\mathbf{A} = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}$ and $\mathbf{B} = \begin{pmatrix} 1 & 3 \\ 0 & 4 \end{pmatrix}$, find an invertible $\mathbf{P}$ with $\mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ and $\mathbf{P}^{-1}\mathbf{B}\mathbf{P}$ diagonal, and show that no **unitary** $\mathbf{P}$ does this.
- (b) For $\mathbf{X} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ and $\mathbf{Z} = \text{diag}(1, -1)$, write down an orthogonal $\mathbf{Q}_1$ diagonalizing $\mathbf{X}$ and an orthogonal $\mathbf{Q}_2$ diagonalizing $\mathbf{Z}$, and prove that no single invertible matrix diagonalizes both.

Solution.

- (a) The common eigenvectors are $\mathbf{e}_1$ (eigenvalues 1 and 1) and $(1, 1)$ (eigenvalues 2 and 4): $\mathbf{A}(1, 1) = (2, 2)$ and $\mathbf{B}(1, 1) = (4, 4)$. So $\mathbf{P} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ works, with $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \text{diag}(1, 2)$ and $\mathbf{P}^{-1}\mathbf{B}\mathbf{P} = \text{diag}(1, 4)$. If a unitary $\mathbf{P}$ did the same, then $\mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^*$ with $\mathbf{D}$ diagonal would be normal by Corollary 12.4.3; but $\mathbf{A}^*\mathbf{A} = \begin{pmatrix} 1 & 1 \\ 1 & 5 \end{pmatrix}$ and $\mathbf{A}\mathbf{A}^* = \begin{pmatrix} 2 & 2 \\ 2 & 4 \end{pmatrix}$ differ in every entry.
- (b) $\mathbf{X}$ has eigenvectors $(1, 1)$ and $(1, -1)$ for the eigenvalues 1 and -1 , so $\mathbf{Q}_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ gives $\mathbf{Q}_1^\top \mathbf{X} \mathbf{Q}_1 = \text{diag}(1, -1)$. $\mathbf{Z}$ is already diagonal, so $\mathbf{Q}_2 = \mathbf{I}_2$. No single invertible $\mathbf{P}$ diagonalizes both: if it did, then $\mathbf{P}^{-1}\mathbf{X}\mathbf{P}$ and $\mathbf{P}^{-1}\mathbf{Z}\mathbf{P}$ would be diagonal, hence commuting, and multiplying by $\mathbf{P}$ and $\mathbf{P}^{-1}$ would give $\mathbf{XZ} = \mathbf{ZX}$. But $\mathbf{XZ} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ and $\mathbf{ZX} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$.

Exercise. B3: Triangularize two at once

Let $\mathbf{A} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ and $\mathbf{B} = \begin{pmatrix} 2 & 3 \\ 0 & 2 \end{pmatrix}$ in $M_2(\mathbb{C})$.

- (a) Verify that $\mathbf{AB} = \mathbf{BA}$.
- (b) Determine whether $\mathbf{A}$ and $\mathbf{B}$ are simultaneously diagonalizable, and justify your answer.
- (c) Exhibit a unitary $\mathbf{U}$ with $\mathbf{U}^*\mathbf{A}\mathbf{U}$ and $\mathbf{U}^*\mathbf{B}\mathbf{U}$ both upper triangular, as [Theorem 12.7.5](#) promises.

Solution.

- (a) $\mathbf{AB} = \begin{pmatrix} 2 & 5 \\ 0 & 2 \end{pmatrix} = \mathbf{BA}$.
- (b) No. If one basis made both diagonal, then $\mathbf{A}$ alone would be diagonalizable; but $p_{\mathbf{A}} = (x - 1)^2$ and $\mathbf{A} - \mathbf{I}_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \neq \mathbf{O}$, so $\mathbf{A}$ has the single eigenvalue 1 with a one-dimensional eigenspace, and is not diagonalizable. This is the hypothesis that Theorem 9.9.5 needs and that commuting does not supply.
- (c) $\mathbf{U} = \mathbf{I}_2$: both matrices are already upper triangular in the standard basis, which is orthonormal. The content of the theorem is that such a basis always exists, not that it is hard to find in this example.

12.7.5.3 C. Going deeper

Exercise. C1: A special case of a theorem of Fuglede

Let V be a finite-dimensional complex inner product space and let $S, T \in \mathcal{L}(V)$ be **both normal**, with $ST = TS$.

- (a) Prove that $ST^* = T^*S$ and $S^*T = TS^*$.
- (b) Deduce that $S + T^*$ and ST^* are normal.
- (c) Give an example of two commuting operators, **not** both normal, for which the conclusion of (a) fails as stated for the pair — that is, exhibit commuting S, T with $ST^* \neq T^*S$.

Note on the hypotheses: a theorem of Fuglede says that in (a) the operator S need not be assumed normal — it is enough that T be normal and that S commute with T . That stronger statement is true and is not proved in this book; (a) is only the case in which both are normal, and its proof uses the normality of S in an essential way.

Solution.

- (a) By Theorem 12.7.1 applied to the commuting family $\{S, T\}$ of normal operators, there is an orthonormal basis $\mathcal{B}$ with $[S]_{\mathcal{B}}^{\mathcal{B}} = \mathbf{D}$ and $[T]_{\mathcal{B}}^{\mathcal{B}} = \mathbf{D}'$ diagonal. By Theorem 11.6.3, $[T^*]_{\mathcal{B}}^{\mathcal{B}} = (\mathbf{D}')^*$, which is the diagonal matrix with the conjugated entries. Diagonal matrices commute, so

$$[ST^*]_{\mathcal{B}}^{\mathcal{B}} = \mathbf{D}(\mathbf{D}')^* = (\mathbf{D}')^*\mathbf{D} = [T^*S]_{\mathcal{B}}^{\mathcal{B}},$$

and $ST^* = T^*S$ because a map is determined by its matrix (Theorem 4.1.4). Exchanging the roles of S and T gives $TS^* = S^*T$.

- (b) T^* is normal, since $(T^*)^*T^* = TT^* = T^*T = T^*(T^*)^*$. By (a), S and T^* commute, and both are normal, so Theorem 12.7.4 applies to the pair (S, T^*) and gives that $S + T^*$ and ST^* are normal.
- (c) Take $S = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ and $T = \mathbf{I}_2$. They commute, T is normal, and S is not. Here $ST^* = S = T^*S$, so this pair does not refute anything — as the theorem of Fuglede predicts, since T is normal. Take instead $T = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ and $S = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, which

commute, since both are polynomials in T . Then

$$ST^* = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}, \quad T^*S = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix},$$

which differ. Here it is T , not S , that fails to be normal, and that is exactly the hypothesis the theorem of Fuglede keeps.

Exercise. C2: A commuting family of orthogonal projections

Let $V \neq \{0\}$ be a finite-dimensional inner product space over $F = \mathbb{R}$ or $\mathbb{C}$, and let $P_1, \dots, P_r \in \mathcal{L}(V)$ be the orthogonal projections onto subspaces $U_1, \dots, U_r$, pairwise commuting.

- Prove that there is an orthonormal basis $\mathcal{B}$ of V such that every vector of $\mathcal{B}$ lies in U_i or in $U_i^\perp$, for every i .
- Deduce that $P_i P_j$ is the orthogonal projection onto $U_i \cap U_j$.
- Prove that $P_i + P_j$ is an orthogonal projection if and only if $U_i \perp U_j$, and identify the subspace in that case.

Hint for (a): the eigenvalues of an idempotent operator.

Solution.

- Each P_i is self-adjoint, by Example (Four Adjoints) (d). Over $\mathbb{R}$, [Corollary 12.7.3](#) applies directly to the commuting family $\{P_1, \dots, P_r\}$; over $\mathbb{C}$ a self-adjoint operator is normal, since $P^*P = P^2 = PP^*$, so [Theorem 12.7.1](#) applies. Either way there is an orthonormal basis $\mathcal{B}$ of common eigenvectors. Let $v \in \mathcal{B}$ and fix i . If $P_i v = \mu v$, then $\mu^2 v = P_i^2 v = P_i v = \mu v$, so $\mu \in \{0, 1\}$, since $v \neq 0$. If $\mu = 1$ then $v = P_i v \in \text{im } P_i = U_i$; if $\mu = 0$ then $v \in \ker P_i = U_i^\perp$ (Definition 11.3.2).
- Write $\mathcal{B} = (v_1, \dots, v_n)$ and let $W = U_i \cap U_j$. By (a) each v_t satisfies $P_i v_t = a_t v_t$ and $P_j v_t = b_t v_t$ with $a_t, b_t \in \{0, 1\}$, so $P_i P_j v_t = a_t b_t v_t$, which is v_t when $v_t \in U_i \cap U_j$ and 0 otherwise. Let $L = \text{span}\{v_t : a_t b_t = 1\}$; then $L \subseteq W$. Conversely, if $w \in W$, expand $w = \sum_t c_t v_t$; applying P_i gives $w = \sum_t a_t c_t v_t$, so $c_t = a_t c_t$ and hence $c_t = 0$ whenever $a_t = 0$, and likewise for b_t . So $w \in L$ and $W = L$. Now $P_i P_j$ fixes each $v_t \in L$ and kills the rest, and by Theorem 11.3.3 (a) with the orthonormal basis $\{v_t : a_t b_t = 1\}$ of W , so does P_W ; two operators agreeing on a basis are equal, so $P_i P_j = P_W$.
- Write $Q = P_i + P_j$.

($\Rightarrow$) Suppose Q is the orthogonal projection onto some subspace W . Then $Q^2 = Q$ (Definition 11.3.2), and expanding,

$$P_i + P_j = Q^2 = P_i^2 + P_i P_j + P_j P_i + P_j^2 = P_i + P_j + 2P_i P_j,$$

using $P_i^2 = P_i$, $P_j^2 = P_j$, and $P_j P_i = P_i P_j$. Hence $2P_i P_j = 0$, so $P_i P_j = 0$. Now let $v \in U_j$; then $P_i v = P_i(P_j v) = 0$, so $v \in \ker P_i = U_i^\perp$. Therefore $U_j \subseteq U_i^\perp$, which says $U_i \perp U_j$.

($\Leftarrow$) Suppose $U_i \perp U_j$, so $U_j \subseteq U_i^\perp = \ker P_i$ and $U_i \subseteq \ker P_j$. For $v \in U_i$ we get $Qv = v + 0 = v$, and symmetrically $Qv = v$ for $v \in U_j$, so Q is the identity on $U_i \oplus U_j$. For $v \in (U_i \oplus U_j)^\perp = U_i^\perp \cap U_j^\perp$ we get $Qv = 0 + 0 = 0$. Since $V = (U_i \oplus U_j) \oplus (U_i \oplus U_j)^\perp$ by Theorem 11.3.2 (a), Q agrees with $P_{U_i \oplus U_j}$ on both summands, hence everywhere. So

$$P_i + P_j = P_{U_i \oplus U_j}.$$