7B Spectral Theorem

Recall that a diagonal matrix is a square matrix that is 0 everywhere except possibly on the diagonal. Recall that an operator on V is called diagonalizable if the operator has a diagonal matrix with respect to some basis of V. Recall also that this happens if and only if there is a basis of V consisting of eigenvectors of the operator (see 5.55).

The nicest operators on V are those for which there is an *orthonormal* basis of V with respect to which the operator has a diagonal matrix. These are precisely the operators $T \in \mathcal{L}(V)$ such that there is an orthonormal basis of V consisting of eigenvectors of T. Our goal in this section is to prove the spectral theorem, which characterizes these operators as the self-adjoint operators when $\mathbf{F} = \mathbf{R}$ and as the normal operators when $\mathbf{F} = \mathbf{C}$.

The spectral theorem is probably the most useful tool in the study of operators on inner product spaces. Its extension to certain infinite-dimensional inner product spaces (see, for example, Section 10D of the author's book *Measure, Integration & Real Analysis*) plays a key role in functional analysis.

Because the conclusion of the spectral theorem depends on $\mathbf{F}$, we will break the spectral theorem into two pieces, called the real spectral theorem and the complex spectral theorem.

Real Spectral Theorem

To prove the real spectral theorem, we will need two preliminary results. These preliminary results hold on both real and complex inner product spaces, but they are not needed for the proof of the complex spectral theorem.

You could guess that the next result is true and even discover its proof by thinking about quadratic polynomials with real coefficients. Specifically, suppose $b, c \in \mathbf{R}$ and $b^2 < 4c$. Let x be a real number. Then

This completing-the-square technique can be used to derive the quadratic formula.

$$x^2 + bx + c = \left(x + \frac{b}{2}\right)^2 + \left(c - \frac{b^2}{4}\right) > 0.$$

In particular, $x^2 + bx + c$ is an invertible real number (a convoluted way of saying that it is not 0). Replacing the real number x with a self-adjoint operator (recall the analogy between real numbers and self-adjoint operators) leads to the next result.

7.26 *invertible quadratic expressions*

Suppose $T \in \mathcal{L}(V)$ is self-adjoint and $b, c \in \mathbf{R}$ are such that $b^2 < 4c$. Then

$$T^2 + bT + cI$$

is an invertible operator.

Proof Let v be a nonzero vector in V. Then

$$\begin{aligned}\langle (T^2 + bT + cI)v, v\rangle &= \langle T^2 v, v\rangle + b\langle Tv, v\rangle + c\langle v, v\rangle \\ &= \langle Tv, Tv\rangle + b\langle Tv, v\rangle + c\|v\|^2 \\ &\geq \|Tv\|^2 - |b|\,\|Tv\|\,\|v\| + c\|v\|^2 \\ &= \left(\|Tv\| - \frac{|b|\,\|v\|}{2}\right)^2 + \left(c - \frac{b^2}{4}\right)\|v\|^2 \\ &> 0,\end{aligned}$$

where the third line above holds by the Cauchy–Schwarz inequality (6.14). The last inequality implies that $(T^2 + bT + cI)v \neq 0$. Thus $T^2 + bT + cI$ is injective, which implies that it is invertible (see 3.65). ■

The next result will be a key tool in our proof of the real spectral theorem.

7.27 *minimal polynomial of self-adjoint operator*

Suppose $T \in \mathcal{L}(V)$ is self-adjoint. Then the minimal polynomial of T equals $(z - \lambda_1) \cdots (z - \lambda_m)$ for some $\lambda_1, \ldots, \lambda_m \in \mathbf{R}$.

Proof First suppose $\mathbf{F} = \mathbf{C}$. The zeros of the minimal polynomial of T are the eigenvalues of T [by 5.27(a)]. All eigenvalues of T are real (by 7.12). Thus the second version of the fundamental theorem of algebra (see 4.13) tells us that the minimal polynomial of T has the desired form.

Now suppose $\mathbf{F} = \mathbf{R}$. By the factorization of a polynomial over $\mathbf{R}$ (see 4.16) there exist $\lambda_1, \ldots, \lambda_m \in \mathbf{R}$ and $b_1, \ldots, b_N, c_1, \ldots, c_N \in \mathbf{R}$ with $b_k^2 < 4c_k$ for each k such that the minimal polynomial of T equals

7.28 $$(z - \lambda_1)\cdots(z - \lambda_m)(z^2 + b_1 z + c_1)\cdots(z^2 + b_N z + c_N);$$

here either m or N might equal 0, meaning that there are no terms of the corresponding form. Now

$$(T - \lambda_1 I)\cdots(T - \lambda_m I)(T^2 + b_1 T + c_1 I)\cdots(T^2 + b_N T + c_N I) = 0.$$

If $N > 0$, then we could multiply both sides of the equation above on the right by the inverse of $T^2 + b_N T + c_N I$ (which is an invertible operator by 7.26) to obtain a polynomial expression of T that equals 0. The corresponding polynomial would have degree two less than the degree of 7.28, violating the minimality of the degree of the polynomial with this property. Thus we must have $N = 0$, which means that the minimal polynomial in 7.28 has the form $(z - \lambda_1) \cdots (z - \lambda_m)$, as desired. ■

The result above along with 5.27(a) implies that every self-adjoint operator has an eigenvalue. In fact, as we will see in the next result, self-adjoint operators have enough eigenvectors to form a basis.

The next result, which gives a complete description of the self-adjoint operators on a real inner product space, is one of the major theorems in linear algebra.

7.29 *real spectral theorem*

Suppose $\mathbf{F} = \mathbf{R}$ and $T \in \mathcal{L}(V)$. Then the following are equivalent.

(a) T is self-adjoint.

(b) T has a diagonal matrix with respect to some orthonormal basis of V.

(c) V has an orthonormal basis consisting of eigenvectors of T.

Proof First suppose (a) holds, so T is self-adjoint. Our results on minimal polynomials, specifically 6.37 and 7.27, imply that T has an upper-triangular matrix with respect to some orthonormal basis of V. With respect to this orthonormal basis, the matrix of T^* is the transpose of the matrix of T. However, $T^* = T$. Thus the transpose of the matrix of T equals the matrix of T. Because the matrix of T is upper-triangular, this means that all entries of the matrix above and below the diagonal are 0. Hence the matrix of T is a diagonal matrix with respect to the orthonormal basis. Thus (a) implies (b).

Conversely, now suppose (b) holds, so T has a diagonal matrix with respect to some orthonormal basis of V. That diagonal matrix equals its transpose. Thus with respect to that basis, the matrix of T^* equals the matrix of T. Hence $T^* = T$, proving that (b) implies (a).

The equivalence of (b) and (c) follows from the definitions [or see the proof that (a) and (b) are equivalent in 5.55]. ∎

7.30 example: *an orthonormal basis of eigenvectors for an operator*

Consider the operator T on $\mathbf{R}^3$ whose matrix (with respect to the standard basis) is

$$\begin{pmatrix} 14 & -13 & 8 \\ -13 & 14 & 8 \\ 8 & 8 & -7 \end{pmatrix}.$$

This matrix with real entries equals its transpose; thus T is self-adjoint. As you can verify,

$$\frac{(1,-1,0)}{\sqrt{2}}, \frac{(1,1,1)}{\sqrt{3}}, \frac{(1,1,-2)}{\sqrt{6}}$$

is an orthonormal basis of $\mathbf{R}^3$ consisting of eigenvectors of T. With respect to this basis, the matrix of T is the diagonal matrix

$$\begin{pmatrix} 27 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & -15 \end{pmatrix}.$$

See Exercise 17 for a version of the real spectral theorem that applies simultaneously to more than one operator.

Complex Spectral Theorem

The next result gives a complete description of the normal operators on a complex inner product space.

7.31 *complex spectral theorem*

Suppose $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$. Then the following are equivalent.

(a) T is normal.

(b) T has a diagonal matrix with respect to some orthonormal basis of V.

(c) V has an orthonormal basis consisting of eigenvectors of T.

Proof First suppose (a) holds, so T is normal. By Schur's theorem (6.38), there is an orthonormal basis $e_1, \ldots, e_n$ of V with respect to which T has an upper-triangular matrix. Thus we can write

7.32
$$\mathcal{M}\bigl(T, (e_1, \ldots, e_n)\bigr) = \begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ & \ddots & \vdots \\ 0 & & a_{n,n} \end{pmatrix}.$$

We will show that this matrix is actually a diagonal matrix.

We see from the matrix above that

$$\|Te_1\|^2 = |a_{1,1}|^2,$$
$$\|T^*e_1\|^2 = |a_{1,1}|^2 + |a_{1,2}|^2 + \cdots + |a_{1,n}|^2.$$

Because T is normal, $\|Te_1\| = \|T^*e_1\|$ (see 7.20). Thus the two equations above imply that all entries in the first row of the matrix in 7.32, except possibly the first entry $a_{1,1}$, equal 0.

Now 7.32 implies

$$\|Te_2\|^2 = |a_{2,2}|^2$$

(because $a_{1,2} = 0$, as we showed in the paragraph above) and

$$\|T^*e_2\|^2 = |a_{2,2}|^2 + |a_{2,3}|^2 + \cdots + |a_{2,n}|^2.$$

Because T is normal, $\|Te_2\| = \|T^*e_2\|$. Thus the two equations above imply that all entries in the second row of the matrix in 7.32, except possibly the diagonal entry $a_{2,2}$, equal 0.

Continuing in this fashion, we see that all nondiagonal entries in the matrix 7.32 equal 0. Thus (b) holds, completing the proof that (a) implies (b).

Now suppose (b) holds, so T has a diagonal matrix with respect to some orthonormal basis of V. The matrix of T^* (with respect to the same basis) is obtained by taking the conjugate transpose of the matrix of T; hence T^* also has a diagonal matrix. Any two diagonal matrices commute; thus T commutes with T^*, which means that T is normal. In other words, (a) holds, completing the proof that (b) implies (a).

The equivalence of (b) and (c) follows from the definitions (also see 5.55). ■

See Exercises 13 and 20 for alternative proofs that (a) implies (b) in the previous result.

Exercises 14 and 15 interpret the real spectral theorem and the complex spectral theorem by expressing the domain space as an orthogonal direct sum of eigenspaces.

See Exercise 16 for a version of the complex spectral theorem that applies simultaneously to more than one operator.

The main conclusion of the complex spectral theorem is that every normal operator on a complex finite-dimensional inner product space is diagonalizable by an orthonormal basis, as illustrated by the next example.

7.33 example: *an orthonormal basis of eigenvectors for an operator*

Consider the operator $T \in \mathcal{L}(\mathbf{C}^2)$ defined by $T(w, z) = (2w - 3z, 3w + 2z)$. The matrix of T (with respect to the standard basis) is

$$\begin{pmatrix} 2 & -3 \\ 3 & 2 \end{pmatrix}.$$

As we saw in Example 7.19, T is a normal operator.

As you can verify,

$$\tfrac{1}{\sqrt{2}}(i, 1), \tfrac{1}{\sqrt{2}}(-i, 1)$$

is an orthonormal basis of $\mathbf{C}^2$ consisting of eigenvectors of T, and with respect to this basis the matrix of T is the diagonal matrix

$$\begin{pmatrix} 2+3i & 0 \\ 0 & 2-3i \end{pmatrix}.$$

Exercises 7B

1 Prove that a normal operator on a complex inner product space is self-adjoint if and only if all its eigenvalues are real.

This exercise strengthens the analogy (for normal operators) between self-adjoint operators and real numbers.

2 Suppose $\mathbf{F} = \mathbf{C}$. Suppose $T \in \mathcal{L}(V)$ is normal and has only one eigenvalue. Prove that T is a scalar multiple of the identity operator.

3 Suppose $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$ is normal. Prove that the set of eigenvalues of T is contained in $\{0, 1\}$ if and only if there is a subspace U of V such that $T = P_U$.

4 Prove that a normal operator on a complex inner product space is skew (meaning it equals the negative of its adjoint) if and only if all its eigenvalues are purely imaginary (meaning that they have real part equal to 0).

5 Prove or give a counterexample: If $T \in \mathcal{L}(\mathbf{C}^3)$ is a diagonalizable operator, then T is normal (with respect to the usual inner product).

6 Suppose V is a complex inner product space and $T \in \mathcal{L}(V)$ is a normal operator such that $T^9 = T^8$. Prove that T is self-adjoint and $T^2 = T$.

7 Give an example of an operator T on a complex vector space such that $T^9 = T^8$ but $T^2 \neq T$.

8 Suppose $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$. Prove that T is normal if and only if every eigenvector of T is also an eigenvector of T^*.

9 Suppose $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$. Prove that T is normal if and only if there exists a polynomial $p \in \mathcal{P}(\mathbf{C})$ such that $T^* = p(T)$.

10 Suppose V is a complex inner product space. Prove that every normal operator on V has a square root.

*An operator $S \in \mathcal{L}(V)$ is called a **square root** of $T \in \mathcal{L}(V)$ if $S^2 = T$. We will discuss more about square roots of operators in Sections 7C and 8C.*

11 Prove that every self-adjoint operator on V has a cube root.

*An operator $S \in \mathcal{L}(V)$ is called a **cube root** of $T \in \mathcal{L}(V)$ if $S^3 = T$.*

12 Suppose V is a complex vector space and $T \in \mathcal{L}(V)$ is normal. Prove that if S is an operator on V that commutes with T, then S commutes with T^*.

The result in this exercise is called Fuglede's theorem.

13 Without using the complex spectral theorem, use the version of Schur's theorem that applies to two commuting operators (take $\mathcal{E} = \{T, T^*\}$ in Exercise 20 in Section 6B) to give a different proof that if $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$ is normal, then T has a diagonal matrix with respect to some orthonormal basis of V.

14 Suppose $\mathbf{F} = \mathbf{R}$ and $T \in \mathcal{L}(V)$. Prove that T is self-adjoint if and only if all pairs of eigenvectors corresponding to distinct eigenvalues of T are orthogonal and $V = E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T)$, where $\lambda_1, \ldots, \lambda_m$ denote the distinct eigenvalues of T.

15 Suppose $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$. Prove that T is normal if and only if all pairs of eigenvectors corresponding to distinct eigenvalues of T are orthogonal and $V = E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T)$, where $\lambda_1, \ldots, \lambda_m$ denote the distinct eigenvalues of T.

16 Suppose $\mathbf{F} = \mathbf{C}$ and $\mathcal{E} \subseteq \mathcal{L}(V)$. Prove that there is an orthonormal basis of V with respect to which every element of $\mathcal{E}$ has a diagonal matrix if and only if S and T are commuting normal operators for all $S, T \in \mathcal{E}$.

This exercise extends the complex spectral theorem to the context of a collection of commuting normal operators.

17 Suppose $\mathbf{F} = \mathbf{R}$ and $\mathcal{E} \subseteq \mathcal{L}(V)$. Prove that there is an orthonormal basis of V with respect to which every element of $\mathcal{E}$ has a diagonal matrix if and only if S and T are commuting self-adjoint operators for all $S, T \in \mathcal{E}$.

This exercise extends the real spectral theorem to the context of a collection of commuting self-adjoint operators.

18 Give an example of a real inner product space V, an operator $T \in \mathcal{L}(V)$, and real numbers b, c with $b^2 < 4c$ such that

$$T^2 + bT + cI$$

is not invertible.

This exercise shows that the hypothesis that T is self-adjoint cannot be deleted in 7.26, even for real vector spaces.

19 Suppose $T \in \mathcal{L}(V)$ is self-adjoint and U is a subspace of V that is invariant under T.

(a) Prove that $U^\perp$ is invariant under T.
(b) Prove that $T|_U \in \mathcal{L}(U)$ is self-adjoint.
(c) Prove that $T|_{U^\perp} \in \mathcal{L}(U^\perp)$ is self-adjoint.

20 Suppose $T \in \mathcal{L}(V)$ is normal and U is a subspace of V that is invariant under T.

(a) Prove that $U^\perp$ is invariant under T.
(b) Prove that U is invariant under T^*.
(c) Prove that $(T|_U)^* = (T^*)|_U$.
(d) Prove that $T|_U \in \mathcal{L}(U)$ and $T|_{U^\perp} \in \mathcal{L}(U^\perp)$ are normal operators.

This exercise can be used to give yet another proof of the complex spectral theorem (use induction on $\dim V$ *and the result that T has an eigenvector).*

21 Suppose that T is a self-adjoint operator on a finite-dimensional inner product space and that 2 and 3 are the only eigenvalues of T. Prove that

$$T^2 - 5T + 6I = 0.$$

22 Give an example of an operator $T \in \mathcal{L}(\mathbf{C}^3)$ such that 2 and 3 are the only eigenvalues of T and $T^2 - 5T + 6I \neq 0$.

23 Suppose $T \in \mathcal{L}(V)$ is self-adjoint, $\lambda \in \mathbf{F}$, and $\epsilon > 0$. Suppose there exists $v \in V$ such that $\|v\| = 1$ and

$$\|Tv - \lambda v\| < \epsilon.$$

Prove that T has an eigenvalue λ' such that $|\lambda - \lambda'| < \epsilon$.

This exercise shows that for a self-adjoint operator, a number that is close to satisfying an equation that would make it an eigenvalue is close to an eigenvalue.

24 Suppose U is a finite-dimensional vector space and $T \in \mathcal{L}(U)$.

(a) Suppose $\mathbf{F} = \mathbf{R}$. Prove that T is diagonalizable if and only if there is a basis of U such that the matrix of T with respect to this basis equals its transpose.

(b) Suppose $\mathbf{F} = \mathbf{C}$. Prove that T is diagonalizable if and only if there is a basis of U such that the matrix of T with respect to this basis commutes with its conjugate transpose.

This exercise adds another equivalence to the list of conditions equivalent to diagonalizability in 5.55.

25 Suppose that $T \in \mathcal{L}(V)$ and there is an orthonormal basis $e_1, \dots, e_n$ of V consisting of eigenvectors of T, with corresponding eigenvalues $\lambda_1, \dots, \lambda_n$. Show that if $k \in \{1, \dots, n\}$, then the pseudoinverse $T^\dagger$ satisfies the equation

$$T^\dagger e_k = \begin{cases} \frac{1}{\lambda_k} e_k & \text{if } \lambda_k \neq 0, \\ 0 & \text{if } \lambda_k = 0. \end{cases}$$

7C Positive Operators

7.34 definition: *positive operator*

An operator $T \in \mathcal{L}(V)$ is called *positive* if T is self-adjoint and

$$\langle Tv, v \rangle \geq 0$$

for all $v \in V$.

If V is a complex vector space, then the requirement that T be self-adjoint can be dropped from the definition above (by 7.14).

7.35 example: *positive operators*

(a) Let $T \in \mathcal{L}(\mathbf{F}^2)$ be the operator whose matrix (using the standard basis) is $\begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}$. Then T is self-adjoint and $\langle T(w,z), (w,z) \rangle = 2|w|^2 - 2\operatorname{Re}(w\bar{z}) + |z|^2 = |w-z|^2 + |w|^2 \geq 0$ for all $(w,z) \in \mathbf{F}^2$. Thus T is a positive operator.

(b) If U is a subspace of V, then the orthogonal projection P_U is a positive operator, as you should verify.

(c) If $T \in \mathcal{L}(V)$ is self-adjoint and $b, c \in \mathbf{R}$ are such that $b^2 < 4c$, then $T^2 + bT + cI$ is a positive operator, as shown by the proof of 7.26.

7.36 definition: *square root*

An operator R is called a *square root* of an operator T if $R^2 = T$.

7.37 example: *square root of an operator*

If $T \in \mathcal{L}(\mathbf{F}^3)$ is defined by $T(z_1, z_2, z_3) = (z_3, 0, 0)$, then the operator $R \in \mathcal{L}(\mathbf{F}^3)$ defined by $R(z_1, z_2, z_3) = (z_2, z_3, 0)$ is a square root of T because $R^2 = T$, as you can verify.

*Because positive operators correspond to nonnegative numbers, better terminology would use the term nonnegative operators. However, operator theorists consistently call these positive operators, so we follow that custom. Some mathematicians use the term **positive semidefinite operator**, which means the same as positive operator.*

The characterizations of the positive operators in the next result correspond to characterizations of the nonnegative numbers among $\mathbf{C}$. Specifically, a number $z \in \mathbf{C}$ is nonnegative if and only if it has a nonnegative square root, corresponding to condition (d). Also, z is nonnegative if and only if it has a real square root, corresponding to condition (e). Finally, z is nonnegative if and only if there exists $w \in \mathbf{C}$ such that $z = \bar{w}w$, corresponding to condition (f). See Exercise 20 for another condition that is equivalent to being a positive operator.