Chapter 7
Operators on Inner Product Spaces

The deepest results related to inner product spaces deal with the subject to which we now turn—linear maps and operators on inner product spaces. As we will see, good theorems can be proved by exploiting properties of the adjoint.

The hugely important spectral theorem will provide a complete description of self-adjoint operators on real inner product spaces and of normal operators on complex inner product spaces. We will then use the spectral theorem to help understand positive operators and unitary operators, which will lead to unitary matrices and matrix factorizations. The spectral theorem will also lead to the popular singular value decomposition, which will lead to the polar decomposition.

The most important results in the rest of this book are valid only in finite dimensions. Thus from now on we assume that V and W are finite-dimensional.

standing assumptions for this chapter

- $\mathbf{F}$ denotes $\mathbf{R}$ or $\mathbf{C}$.
- V and W are nonzero finite-dimensional inner product spaces over $\mathbf{F}$.

Market square in Lviv, a city that has had several names and has been in several countries because of changing international borders. From 1772 until 1918, the city was in Austria and was called Lemberg. Between World War I and World War II, the city was in Poland and was called Lwów. During this time, mathematicians in Lwów, particularly Stefan Banach (1892–1945) and his colleagues, developed the basic results of modern functional analysis, using tools of analysis to study infinite-dimensional vector spaces. Since the end of World War II, Lviv has been in Ukraine, which was part of the Soviet Union until Ukraine became an independent country in 1991.

7A *Self-Adjoint and Normal Operators*

Adjoints

7.1 definition: *adjoint,* T^*

Suppose $T \in \mathcal{L}(V, W)$. The *adjoint* of T is the function $T^* \colon W \to V$ such that

$$\langle Tv, w\rangle = \langle v, T^*w\rangle$$

for every $v \in V$ and every $w \in W$.

*The word **adjoint** has another meaning in linear algebra. In case you encounter the second meaning elsewhere, be warned that the two meanings for adjoint are unrelated to each other.*

To see why the definition above makes sense, suppose $T \in \mathcal{L}(V, W)$. Fix $w \in W$. Consider the linear functional

$$v \mapsto \langle Tv, w\rangle$$

on V that maps $v \in V$ to $\langle Tv, w\rangle$; this linear functional depends on T and w. By the Riesz representation theorem (6.42), there exists a unique vector in V such that this linear functional is given by taking the inner product with it. We call this unique vector T^*w. In other words, T^*w is the unique vector in V such that

$$\langle Tv, w\rangle = \langle v, T^*w\rangle$$

for every $v \in V$.

In the equation above, the inner product on the left takes place in W and the inner product on the right takes place in V. However, we use the same notation $\langle \cdot, \cdot \rangle$ for both inner products.

7.2 example: *adjoint of a linear map from* $\mathbf{R}^3$ *to* $\mathbf{R}^2$

Define $T \colon \mathbf{R}^3 \to \mathbf{R}^2$ by

$$T(x_1, x_2, x_3) = (x_2 + 3x_3, 2x_1).$$

To compute T^*, suppose $(x_1, x_2, x_3) \in \mathbf{R}^3$ and $(y_1, y_2) \in \mathbf{R}^2$. Then

$$\begin{aligned}\bigl\langle T(x_1, x_2, x_3), (y_1, y_2)\bigr\rangle &= \bigl\langle (x_2 + 3x_3, 2x_1), (y_1, y_2)\bigr\rangle \\ &= x_2y_1 + 3x_3y_1 + 2x_1y_2 \\ &= \bigl\langle (x_1, x_2, x_3), (2y_2, y_1, 3y_1)\bigr\rangle.\end{aligned}$$

The equation above and the definition of the adjoint imply that

$$T^*(y_1, y_2) = (2y_2, y_1, 3y_1).$$

7.3 example: *adjoint of a linear map with range of dimension at most* 1

Fix $u \in V$ and $x \in W$. Define $T \in \mathcal{L}(V, W)$ by

$$Tv = \langle v, u \rangle x$$

for each $v \in V$. To compute T^*, suppose $v \in V$ and $w \in W$. Then

$$\begin{aligned} \langle Tv, w \rangle &= \big\langle \langle v, u \rangle x, w \big\rangle \\ &= \langle v, u \rangle \langle x, w \rangle \\ &= \big\langle v, \langle w, x \rangle u \big\rangle. \end{aligned}$$

Thus

$$T^* w = \langle w, x \rangle u.$$

In the two examples above, T^* turned out to be not just a function from W to V but a linear map from W to V. This behavior is true in general, as shown by the next result.

The two examples above and the proof below use a common technique for computing T^*: start with a formula for $\langle Tv, w \rangle$ then manipulate it to get just v in the first slot; the entry in the second slot will then be $T^* w$.

7.4 *adjoint of a linear map is a linear map*

If $T \in \mathcal{L}(V, W)$, then $T^* \in \mathcal{L}(W, V)$.

Proof Suppose $T \in \mathcal{L}(V, W)$. If $v \in V$ and $w_1, w_2 \in W$, then

$$\begin{aligned} \langle Tv, w_1 + w_2 \rangle &= \langle Tv, w_1 \rangle + \langle Tv, w_2 \rangle \\ &= \big\langle v, T^* w_1 \big\rangle + \big\langle v, T^* w_2 \big\rangle \\ &= \big\langle v, T^* w_1 + T^* w_2 \big\rangle. \end{aligned}$$

The equation above shows that

$$T^*(w_1 + w_2) = T^* w_1 + T^* w_2.$$

If $v \in V$, $\lambda \in \mathbf{F}$, and $w \in W$, then

$$\begin{aligned} \langle Tv, \lambda w \rangle &= \overline{\lambda} \langle Tv, w \rangle \\ &= \overline{\lambda} \langle v, T^* w \rangle \\ &= \big\langle v, \lambda T^* w \big\rangle. \end{aligned}$$

The equation above shows that

$$T^*(\lambda w) = \lambda T^* w.$$

Thus T^* is a linear map, as desired. ■

7.5 *properties of the adjoint*

Suppose $T \in \mathcal{L}(V, W)$. Then

(a) $(S+T)^* = S^* + T^*$ for all $S \in \mathcal{L}(V, W)$;

(b) $(\lambda T)^* = \overline{\lambda} T^*$ for all $\lambda \in \mathbf{F}$;

(c) $(T^*)^* = T$;

(d) $(ST)^* = T^* S^*$ for all $S \in \mathcal{L}(W, U)$ (here U is a finite-dimensional inner product space over $\mathbf{F}$);

(e) $I^* = I$, where I is the identity operator on V;

(f) if T is invertible, then T^* is invertible and $(T^*)^{-1} = (T^{-1})^*$.

Proof Suppose $v \in V$ and $w \in W$.

(a) If $S \in \mathcal{L}(V, W)$, then

$$\begin{aligned} \langle (S+T)v, w \rangle &= \langle Sv, w \rangle + \langle Tv, w \rangle \\ &= \langle v, S^* w \rangle + \langle v, T^* w \rangle \\ &= \langle v, S^* w + T^* w \rangle. \end{aligned}$$

Thus $(S+T)^* w = S^* w + T^* w$, as desired.

(b) If $\lambda \in \mathbf{F}$, then

$$\langle (\lambda T)v, w \rangle = \lambda \langle Tv, w \rangle = \lambda \langle v, T^* w \rangle = \langle v, \overline{\lambda} T^* w \rangle.$$

Thus $(\lambda T)^* w = \overline{\lambda} T^* w$, as desired.

(c) We have

$$\langle T^* w, v \rangle = \overline{\langle v, T^* w \rangle} = \overline{\langle Tv, w \rangle} = \langle w, Tv \rangle.$$

Thus $(T^*)^* v = Tv$, as desired.

(d) Suppose $S \in \mathcal{L}(W, U)$ and $u \in U$. Then

$$\langle (ST)v, u \rangle = \langle S(Tv), u \rangle = \langle Tv, S^* u \rangle = \langle v, T^*(S^* u) \rangle.$$

Thus $(ST)^* u = T^*(S^* u)$, as desired.

(e) Suppose $u \in V$. Then

$$\langle Iu, v \rangle = \langle u, v \rangle.$$

Thus $I^* v = v$, as desired.

(f) Suppose T is invertible. Take adjoints of both sides of the equation $T^{-1}T = I$, then use (d) and (e) to show that $T^*(T^{-1})^* = I$. Similarly, the equation $TT^{-1} = I$ implies $(T^{-1})^* T^* = I$. Thus $(T^{-1})^*$ is the inverse of T^*, as desired. ■

If $\mathbf{F} = \mathbf{R}$, then the map $T \mapsto T^*$ is a linear map from $\mathcal{L}(V, W)$ to $\mathcal{L}(W, V)$, as follows from (a) and (b) of the result above. However, if $\mathbf{F} = \mathbf{C}$, then this map is not linear because of the complex conjugate that appears in (b).

The next result shows the relationship between the null space and the range of a linear map and its adjoint.

7.6 *null space and range of T^**

Suppose $T \in \mathcal{L}(V, W)$. Then

(a) $\operatorname{null} T^* = (\operatorname{range} T)^\perp$;

(b) $\operatorname{range} T^* = (\operatorname{null} T)^\perp$;

(c) $\operatorname{null} T = \bigl(\operatorname{range} T^*\bigr)^\perp$;

(d) $\operatorname{range} T = \bigl(\operatorname{null} T^*\bigr)^\perp$.

Proof We begin by proving (a). Let $w \in W$. Then

$$\begin{aligned} w \in \operatorname{null} T^* &\iff T^*w = 0 \\ &\iff \langle v, T^*w\rangle = 0 \text{ for all } v \in V \\ &\iff \langle Tv, w\rangle = 0 \text{ for all } v \in V \\ &\iff w \in (\operatorname{range} T)^\perp. \end{aligned}$$

Thus $\operatorname{null} T^* = (\operatorname{range} T)^\perp$, proving (a).

If we take the orthogonal complement of both sides of (a), we get (d), where we have used 6.52. Replacing T with T^* in (a) gives (c), where we have used 7.5(c). Finally, replacing T with T^* in (d) gives (b). ■

As we will soon see, the next definition is intimately connected to the matrix of the adjoint of a linear map.

7.7 definition: *conjugate transpose, A^**

The *conjugate transpose* of an m-by-n matrix A is the n-by-m matrix A^* obtained by interchanging the rows and columns and then taking the complex conjugate of each entry. In other words, if $j \in \{1, \dots, n\}$ and $k \in \{1, \dots, m\}$, then

$$(A^*)_{j,k} = \overline{A_{k,j}}.$$

7.8 example: *conjugate transpose of a 2-by-3 matrix*

The conjugate transpose of the 2-by-3 matrix $\begin{pmatrix} 2 & 3+4i & 7 \\ 6 & 5 & 8i \end{pmatrix}$ is the 3-by-2 matrix

$$\begin{pmatrix} 2 & 6 \\ 3-4i & 5 \\ 7 & -8i \end{pmatrix}.$$

If a matrix A has only real entries, then $A^* = A^{\mathrm{t}}$, where A^{t} denotes the transpose of A (the matrix obtained by interchanging the rows and the columns).

The next result shows how to compute the matrix of T^* from the matrix of T. **Caution:** With respect to nonorthonormal bases, the matrix of T^* does not necessarily equal the conjugate transpose of the matrix of T.

The adjoint of a linear map does not depend on a choice of basis. Thus we frequently emphasize adjoints of linear maps instead of transposes or conjugate transposes of matrices.

7.9 *matrix of T^* equals conjugate transpose of matrix of T*

Let $T \in \mathcal{L}(V, W)$. Suppose $e_1, \ldots, e_n$ is an orthonormal basis of V and $f_1, \ldots, f_m$ is an orthonormal basis of W. Then $\mathcal{M}\big(T^*, (f_1, \ldots, f_m), (e_1, \ldots, e_n)\big)$ is the conjugate transpose of $\mathcal{M}\big(T, (e_1, \ldots, e_n), (f_1, \ldots, f_m)\big)$. In other words,

$$\mathcal{M}(T^*) = \big(\mathcal{M}(T)\big)^*.$$

Proof In this proof, we will write $\mathcal{M}(T)$ and $\mathcal{M}(T^*)$ instead of the longer expressions $\mathcal{M}\big(T, (e_1, \ldots, e_n), (f_1, \ldots, f_m)\big)$ and $\mathcal{M}\big(T^*, (f_1, \ldots, f_m), (e_1, \ldots, e_n)\big)$.

Recall that we obtain the k^{th} column of $\mathcal{M}(T)$ by writing Te_k as a linear combination of the f_j's; the scalars used in this linear combination then become the k^{th} column of $\mathcal{M}(T)$. Because $f_1, \ldots, f_m$ is an orthonormal basis of W, we know how to write Te_k as a linear combination of the f_j's [see 6.30(a)]:

$$Te_k = \langle Te_k, f_1 \rangle f_1 + \cdots + \langle Te_k, f_m \rangle f_m.$$

Thus

$$\text{the entry in row } j \text{, column } k \text{, of } \mathcal{M}(T) \text{ is } \langle Te_k, f_j \rangle.$$

In the statement above, replace T with T^* and interchange $e_1, \ldots, e_n$ and $f_1, \ldots, f_m$. This shows that the entry in row j, column k, of $\mathcal{M}(T^*)$ is $\langle T^* f_k, e_j \rangle$, which equals $\langle f_k, Te_j \rangle$, which equals $\overline{\langle Te_j, f_k \rangle}$, which equals the complex conjugate of the entry in row k, column j, of $\mathcal{M}(T)$. Thus $\mathcal{M}(T^*) = \big(\mathcal{M}(T)\big)^*$. ■

The Riesz representation theorem as stated in 6.58 provides an identification of V with its dual space V' defined in 3.110. Under this identification, the orthogonal complement $U^\perp$ of a subset $U \subseteq V$ corresponds to the annihilator U^0 of U. If U is a subspace of V, then the formulas for the dimensions of $U^\perp$ and U^0 become identical under this identification—see 3.125 and 6.51.

Suppose $T\colon V \to W$ is a linear map. Under the identification of V with V' and the identification of W with W', the adjoint map $T^*\colon W \to V$ corresponds to the dual map $T'\colon W' \to V'$ defined in 3.118, as Exercise 32 asks you to verify. Under this identification, the formulas for null T^* and range T^* [7.6(a) and (b)] then become identical to the formulas for null T' and range T' [3.128(a) and 3.130(b)]. Furthermore, the theorem about the matrix of T^* (7.9) is analogous to the theorem about the matrix of T' (3.132).

Because orthogonal complements and adjoints are easier to deal with than annihilators and dual maps, there is no need to work with annihilators and dual maps in the context of inner product spaces.

Self-Adjoint Operators

Now we switch our attention to operators on inner product spaces. Instead of considering linear maps from V to W, we will focus on linear maps from V to V; recall that such linear maps are called operators.

7.10 definition: *self-adjoint*

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

If $T \in \mathcal{L}(V)$ and $e_1, \ldots, e_n$ is an orthonormal basis of V, then T is self-adjoint if and only if $\mathcal{M}\big(T, (e_1, \ldots, e_n)\big) = \mathcal{M}\big(T, (e_1, \ldots, e_n)\big)^*$, as follows from 7.9.

7.11 example: *determining whether T is self-adjoint from its matrix*

Suppose $c \in \mathbf{F}$ and T is the operator on $\mathbf{F}^2$ whose matrix (with respect to the standard basis) is

$$\mathcal{M}(T) = \begin{pmatrix} 2 & c \\ 3 & 7 \end{pmatrix}.$$

The matrix of T^* (with respect to the standard basis) is

$$\mathcal{M}(T^*) = \begin{pmatrix} 2 & 3 \\ \bar{c} & 7 \end{pmatrix}.$$

Thus $\mathcal{M}(T) = \mathcal{M}(T^*)$ if and only if $c = 3$. Hence the operator T is self-adjoint if and only if $c = 3$.

A good analogy to keep in mind is that the adjoint on $\mathcal{L}(V)$ plays a role similar to that of the complex conjugate on $\mathbf{C}$. A complex number z is real if and only if $z = \bar{z}$; thus a self-adjoint operator ($T = T^*$) is analogous to a real number.

We will see that the analogy discussed above is reflected in some important properties of self-adjoint operators, beginning with eigenvalues in the next result.

If $\mathbf{F} = \mathbf{R}$, then by definition every eigenvalue is real, so the next result is interesting only when $\mathbf{F} = \mathbf{C}$.

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

$$\langle Tv, w \rangle = \langle v, Tw \rangle$$

for all $v, w \in V$.

7.12 *eigenvalues of self-adjoint operators*

Every eigenvalue of a self-adjoint operator is real.

Proof Suppose T is a self-adjoint operator on V. Let λ be an eigenvalue of T, and let v be a nonzero vector in V such that $Tv = \lambda v$. Then

$$\lambda \|v\|^2 = \langle \lambda v, v \rangle = \langle Tv, v \rangle = \langle v, Tv \rangle = \langle v, \lambda v \rangle = \bar{\lambda} \|v\|^2.$$

Thus $\lambda = \bar{\lambda}$, which means that λ is real, as desired. ■

The next result is false for real inner product spaces. As an example, consider the operator $T \in \mathcal{L}(\mathbf{R}^2)$ that is a counterclockwise rotation of 90° around the origin; thus $T(x, y) = (-y, x)$. Notice that Tv is orthogonal to v for every $v \in \mathbf{R}^2$, even though $T \neq 0$.

7.13 *Tv is orthogonal to v for all v $\iff$ $T = 0$ (assuming $\mathbf{F} = \mathbf{C}$)*

Suppose V is a complex inner product space and $T \in \mathcal{L}(V)$. Then

$$\langle Tv, v\rangle = 0 \text{ for every } v \in V \iff T = 0.$$

Proof If $u, w \in V$, then

$$\begin{aligned}\langle Tu, w\rangle &= \frac{\langle T(u+w), u+w\rangle - \langle T(u-w), u-w\rangle}{4}\\ &\quad + \frac{\langle T(u+iw), u+iw\rangle - \langle T(u-iw), u-iw\rangle}{4} i,\end{aligned}$$

as can be verified by computing the right side. Note that each term on the right side is of the form $\langle Tv, v\rangle$ for appropriate $v \in V$.

Now suppose $\langle Tv, v\rangle = 0$ for every $v \in V$. Then the equation above implies that $\langle Tu, w\rangle = 0$ for all $u, w \in V$, which then implies that $Tu = 0$ for every $u \in V$ (take $w = Tu$). Hence $T = 0$, as desired. ∎

The next result is false for real inner product spaces, as shown by considering any operator on a real inner product space that is not self-adjoint.

The next result provides another good example of how self-adjoint operators behave like real numbers.

7.14 *$\langle Tv, v\rangle$ is real for all v $\iff$ T is self-adjoint (assuming $\mathbf{F} = \mathbf{C}$)*

Suppose V is a complex inner product space and $T \in \mathcal{L}(V)$. Then

$$T \text{ is self-adjoint} \iff \langle Tv, v\rangle \in \mathbf{R} \text{ for every } v \in V.$$

Proof If $v \in V$, then

7.15
$$\langle T^*v, v\rangle = \overline{\langle v, T^*v\rangle} = \overline{\langle Tv, v\rangle}.$$

Now

$$\begin{aligned} T \text{ is self-adjoint} &\iff T - T^* = 0\\ &\iff \langle (T - T^*)v, v\rangle = 0 \text{ for every } v \in V\\ &\iff \langle Tv, v\rangle - \overline{\langle Tv, v\rangle} = 0 \text{ for every } v \in V\\ &\iff \langle Tv, v\rangle \in \mathbf{R} \text{ for every } v \in V,\end{aligned}$$

where the second equivalence follows from 7.13 as applied to $T - T^*$ and the third equivalence follows from 7.15. ∎

On a real inner product space V, a nonzero operator T might satisfy $\langle Tv, v\rangle = 0$ for all $v \in V$. However, the next result shows that this cannot happen for a self-adjoint operator.

7.16 *T self-adjoint and $\langle Tv, v\rangle = 0$ for all v $\iff$ $T = 0$*

Suppose T is a self-adjoint operator on V. Then

$$\langle Tv, v\rangle = 0 \text{ for every } v \in V \iff T = 0.$$

Proof We have already proved this (without the hypothesis that T is self-adjoint) when V is a complex inner product space (see 7.13). Thus we can assume that V is a real inner product space. If $u, w \in V$, then

7.17
$$\langle Tu, w\rangle = \frac{\langle T(u+w), u+w\rangle - \langle T(u-w), u-w\rangle}{4},$$

as can be proved by computing the right side using the equation

$$\langle Tw, u\rangle = \langle w, Tu\rangle = \langle Tu, w\rangle,$$

where the first equality holds because T is self-adjoint and the second equality holds because we are working in a real inner product space.

Now suppose $\langle Tv, v\rangle = 0$ for every $v \in V$. Because each term on the right side of 7.17 is of the form $\langle Tv, v\rangle$ for appropriate v, this implies that $\langle Tu, w\rangle = 0$ for all $u, w \in V$. This implies that $Tu = 0$ for every $u \in V$ (take $w = Tu$). Hence $T = 0$, as desired. ■

Normal Operators

7.18 definition: *normal*

- An operator on an inner product space is called *normal* if it commutes with its adjoint.
- In other words, $T \in \mathcal{L}(V)$ is normal if $TT^* = T^*T$.

Every self-adjoint operator is normal, because if T is self-adjoint then $T^* = T$ and hence T commutes with T^*.

7.19 example: *an operator that is normal but not self-adjoint*

Let T be the operator on $\mathbf{F}^2$ whose matrix (with respect to the standard basis) is

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

Thus $T(w, z) = (2w - 3z, 3w + 2z)$.

This operator T is not self-adjoint because the entry in row 2, column 1 (which equals 3) does not equal the complex conjugate of the entry in row 1, column 2 (which equals -3).

The matrix of TT^* equals

$$\begin{pmatrix} 2 & -3 \\ 3 & 2 \end{pmatrix}\begin{pmatrix} 2 & 3 \\ -3 & 2 \end{pmatrix}, \text{ which equals } \begin{pmatrix} 13 & 0 \\ 0 & 13 \end{pmatrix}.$$

Similarly, the matrix of T^*T equals

$$\begin{pmatrix} 2 & 3 \\ -3 & 2 \end{pmatrix}\begin{pmatrix} 2 & -3 \\ 3 & 2 \end{pmatrix}, \text{ which equals } \begin{pmatrix} 13 & 0 \\ 0 & 13 \end{pmatrix}.$$

Because TT^* and T^*T have the same matrix, we see that $TT^* = T^*T$. Thus T is normal.

In the next section we will see why normal operators are worthy of special attention. The next result provides a useful characterization of normal operators.

7.20 *T is normal if and only if Tv and T^*v have the same norm*

Suppose $T \in \mathcal{L}(V)$. Then

$$T \text{ is normal} \iff \|Tv\| = \|T^*v\| \text{ for every } v \in V.$$

Proof We have

$$\begin{aligned} T \text{ is normal} &\iff T^*T - TT^* = 0 \\ &\iff \langle (T^*T - TT^*)v, v\rangle = 0 \text{ for every } v \in V \\ &\iff \langle T^*Tv, v\rangle = \langle TT^*v, v\rangle \text{ for every } v \in V \\ &\iff \langle Tv, Tv\rangle = \langle T^*v, T^*v\rangle \text{ for every } v \in V \\ &\iff \|Tv\|^2 = \|T^*v\|^2 \text{ for every } v \in V \\ &\iff \|Tv\| = \|T^*v\| \text{ for every } v \in V, \end{aligned}$$

where we used 7.16 to establish the second equivalence (note that the operator $T^*T - TT^*$ is self-adjoint). ∎

The next result presents several consequences of the result above. Compare (e) of the next result to Exercise 3. That exercise states that the eigenvalues of the adjoint of each operator are equal (as a set) to the complex conjugates of the eigenvalues of the operator. The exercise says nothing about eigenvectors, because an operator and its adjoint may have different eigenvectors. However, (e) of the next result implies that a normal operator and its adjoint have the same eigenvectors.

7.21 *range, null space, and eigenvectors of a normal operator*

Suppose $T \in \mathcal{L}(V)$ is normal. Then

(a) $\operatorname{null} T = \operatorname{null} T^*$;

(b) $\operatorname{range} T = \operatorname{range} T^*$;

(c) $V = \operatorname{null} T \oplus \operatorname{range} T$;

(d) $T - \lambda I$ is normal for every $\lambda \in \mathbf{F}$;

(e) if $v \in V$ and $\lambda \in \mathbf{F}$, then $Tv = \lambda v$ if and only if $T^* v = \overline{\lambda} v$.

Proof

(a) Suppose $v \in V$. Then

$$v \in \operatorname{null} T \iff \|Tv\| = 0 \iff \|T^* v\| = 0 \iff v \in \operatorname{null} T^*,$$

where the middle equivalence above follows from 7.20. Thus $\operatorname{null} T = \operatorname{null} T^*$.

(b) We have

$$\operatorname{range} T = \left(\operatorname{null} T^*\right)^{\perp} = (\operatorname{null} T)^{\perp} = \operatorname{range} T^*,$$

where the first equality comes from 7.6(d), the second equality comes from (a) in this result, and the third equality comes from 7.6(b).

(c) We have

$$V = (\operatorname{null} T) \oplus (\operatorname{null} T)^{\perp} = \operatorname{null} T \oplus \operatorname{range} T^* = \operatorname{null} T \oplus \operatorname{range} T,$$

where the first equality comes from 6.49, the second equality comes from 7.6(b), and the third equality comes from (b) in this result.

(d) Suppose $\lambda \in \mathbf{F}$. Then

$$\begin{aligned}
(T - \lambda I)(T - \lambda I)^* &= (T - \lambda I)\left(T^* - \overline{\lambda} I\right) \\
&= TT^* - \overline{\lambda} T - \lambda T^* + |\lambda|^2 I \\
&= T^* T - \overline{\lambda} T - \lambda T^* + |\lambda|^2 I \\
&= \left(T^* - \overline{\lambda} I\right)(T - \lambda I) \\
&= (T - \lambda I)^* (T - \lambda I).
\end{aligned}$$

Thus $T - \lambda I$ commutes with its adjoint. Hence $T - \lambda I$ is normal.

(e) Suppose $v \in V$ and $\lambda \in \mathbf{F}$. Then (d) and 7.20 imply that

$$\|(T - \lambda I) v\| = \|(T - \lambda I)^* v\| = \left\|\left(T^* - \overline{\lambda} I\right) v\right\|.$$

Thus $\|(T - \lambda I) v\| = 0$ if and only if $\left\|\left(T^* - \overline{\lambda} I\right) v\right\| = 0$. Hence $Tv = \lambda v$ if and only if $T^* v = \overline{\lambda} v$. ∎

Because every self-adjoint operator is normal, the next result applies in particular to self-adjoint operators.

7.22 *orthogonal eigenvectors for normal operators*

Suppose $T \in \mathcal{L}(V)$ is normal. Then eigenvectors of T corresponding to distinct eigenvalues are orthogonal.

Proof Suppose α, β are distinct eigenvalues of T, with corresponding eigenvectors u, v. Thus $Tu = \alpha u$ and $Tv = \beta v$. From 7.21(e) we have $T^*v = \overline{\beta} v$. Thus

$$\begin{aligned}(\alpha - \beta)\langle u, v\rangle &= \langle \alpha u, v\rangle - \langle u, \overline{\beta} v\rangle \\ &= \langle Tu, v\rangle - \langle u, T^*v\rangle \\ &= 0.\end{aligned}$$

Because $\alpha \neq \beta$, the equation above implies that $\langle u, v\rangle = 0$. Thus u and v are orthogonal, as desired. ■

As stated here, the next result makes sense only when $\mathbf{F} = \mathbf{C}$. However, see Exercise 12 for a version that makes sense when $\mathbf{F} = \mathbf{C}$ and when $\mathbf{F} = \mathbf{R}$.

Suppose $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$. Under the analogy between $\mathcal{L}(V)$ and $\mathbf{C}$, with the adjoint on $\mathcal{L}(V)$ playing a similar role to that of the complex conjugate on $\mathbf{C}$, the operators A and B as defined by 7.24 correspond to the real and imaginary parts of T. Thus the informal title of the result below should make sense.

7.23 *T is normal $\iff$ the real and imaginary parts of T commute*

Suppose $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$. Then T is normal if and only if there exist commuting self-adjoint operators A and B such that $T = A + iB$.

Proof First suppose T is normal. Let

7.24
$$A = \frac{T + T^*}{2} \quad \text{and} \quad B = \frac{T - T^*}{2i}.$$

Then A and B are self-adjoint and $T = A + iB$. A quick computation shows that

7.25
$$AB - BA = \frac{T^*T - TT^*}{2i}.$$

Because T is normal, the right side of the equation above equals 0. Thus the operators A and B commute, as desired.

To prove the implication in the other direction, now suppose there exist commuting self-adjoint operators A and B such that $T = A + iB$. Then $T^* = A - iB$. Adding the last two equations and then dividing by 2 produces the equation for A in 7.24. Subtracting the last two equations and then dividing by $2i$ produces the equation for B in 7.24. Now 7.24 implies 7.25. Because B and A commute, 7.25 implies that T is normal, as desired. ■

Exercises 7A

1 Suppose n is a positive integer. Define $T \in \mathcal{L}(\mathbf{F}^n)$ by

$$T(z_1, \dots, z_n) = (0, z_1, \dots, z_{n-1}).$$

Find a formula for $T^*(z_1, \dots, z_n)$.

2 Suppose $T \in \mathcal{L}(V, W)$. Prove that

$$T = 0 \iff T^* = 0 \iff T^*T = 0 \iff TT^* = 0.$$

3 Suppose $T \in \mathcal{L}(V)$ and $\lambda \in \mathbf{F}$. Prove that

$$\lambda \text{ is an eigenvalue of } T \iff \overline{\lambda} \text{ is an eigenvalue of } T^*.$$

4 Suppose $T \in \mathcal{L}(V)$ and U is a subspace of V. Prove that

$$U \text{ is invariant under } T \iff U^\perp \text{ is invariant under } T^*.$$

5 Suppose $T \in \mathcal{L}(V, W)$. Suppose $e_1, \dots, e_n$ is an orthonormal basis of V and $f_1, \dots, f_m$ is an orthonormal basis of W. Prove that

$$\|Te_1\|^2 + \cdots + \|Te_n\|^2 = \|T^*f_1\|^2 + \cdots + \|T^*f_m\|^2.$$

The numbers $\|Te_1\|^2, \dots, \|Te_n\|^2$ in the equation above depend on the orthonormal basis $e_1, \dots, e_n$, but the right side of the equation does not depend on $e_1, \dots, e_n$. Thus the equation above shows that the sum on the left side does not depend on which orthonormal basis $e_1, \dots, e_n$ is used.

6 Suppose $T \in \mathcal{L}(V, W)$. Prove that
(a) T is injective $\iff$ T^* is surjective;
(b) T is surjective $\iff$ T^* is injective.

7 Prove that if $T \in \mathcal{L}(V, W)$, then
(a) $\dim \operatorname{null} T^* = \dim \operatorname{null} T + \dim W - \dim V$;
(b) $\dim \operatorname{range} T^* = \dim \operatorname{range} T$.

8 Suppose A is an m-by-n matrix with entries in $\mathbf{F}$. Use (b) in Exercise 7 to prove that the row rank of A equals the column rank of A.

This exercise asks for yet another alternative proof of a result that was previously proved in 3.57 and 3.133.

9 Prove that the product of two self-adjoint operators on V is self-adjoint if and only if the two operators commute.

10 Suppose $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$. Prove that T is self-adjoint if and only if

$$\langle Tv, v \rangle = \langle T^*v, v \rangle$$

for all $v \in V$.

11 Define an operator $S\colon \mathbf{F}^2 \to \mathbf{F}^2$ by $S(w,z) = (-z,w)$.

(a) Find a formula for S^*.

(b) Show that S is normal but not self-adjoint.

(c) Find all eigenvalues of S.

If $\mathbf{F} = \mathbf{R}$, *then* S *is the operator on* $\mathbf{R}^2$ *of counterclockwise rotation by* 90°.

12 An operator $B \in \mathcal{L}(V)$ is called *skew* if

$$B^* = -B.$$

Suppose that $T \in \mathcal{L}(V)$. Prove that T is normal if and only if there exist commuting operators A and B such that A is self-adjoint, B is a skew operator, and $T = A + B$.

13 Suppose $\mathbf{F} = \mathbf{R}$. Define $\mathcal{A} \in \mathcal{L}\bigl(\mathcal{L}(V)\bigr)$ by $\mathcal{A}T = T^*$ for all $T \in \mathcal{L}(V)$.

(a) Find all eigenvalues of $\mathcal{A}$.

(b) Find the minimal polynomial of $\mathcal{A}$.

14 Define an inner product on $\mathcal{P}_2(\mathbf{R})$ by $\langle p,q\rangle = \int_0^1 pq$. Define an operator $T \in \mathcal{L}\bigl(\mathcal{P}_2(\mathbf{R})\bigr)$ by

$$T\bigl(ax^2 + bx + c\bigr) = bx.$$

(a) Show that with this inner product, the operator T is not self-adjoint.

(b) The matrix of T with respect to the basis $1, x, x^2$ is

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

This matrix equals its conjugate transpose, even though T is not self-adjoint. Explain why this is not a contradiction.

15 Suppose $T \in \mathcal{L}(V)$ is invertible. Prove that

(a) T is self-adjoint $\iff$ T^{-1} is self-adjoint;

(b) T is normal $\iff$ T^{-1} is normal.

16 Suppose $\mathbf{F} = \mathbf{R}$.

(a) Show that the set of self-adjoint operators on V is a subspace of $\mathcal{L}(V)$.

(b) What is the dimension of the subspace of $\mathcal{L}(V)$ in (a) [in terms of $\dim V$]?

17 Suppose $\mathbf{F} = \mathbf{C}$. Show that the set of self-adjoint operators on V is not a subspace of $\mathcal{L}(V)$.

18 Suppose $\dim V \geq 2$. Show that the set of normal operators on V is not a subspace of $\mathcal{L}(V)$.

19 Suppose $T \in \mathcal{L}(V)$ and $\|T^*v\| \le \|Tv\|$ for every $v \in V$. Prove that T is normal.

This exercise fails on infinite-dimensional inner product spaces, leading to what are called hyponormal operators, which have a well-developed theory.

20 Suppose $P \in \mathcal{L}(V)$ is such that $P^2 = P$. Prove that the following are equivalent.

(a) P is self-adjoint.

(b) P is normal.

(c) There is a subspace U of V such that $P = P_U$.

21 Suppose $D\colon \mathcal{P}_8(\mathbf{R}) \to \mathcal{P}_8(\mathbf{R})$ is the differentiation operator defined by $Dp = p'$. Prove that there does not exist an inner product on $\mathcal{P}_8(\mathbf{R})$ that makes D a normal operator.

22 Give an example of an operator $T \in \mathcal{L}(\mathbf{R}^3)$ such that T is normal but not self-adjoint.

23 Suppose T is a normal operator on V. Suppose also that $v, w \in V$ satisfy the equations

$$\|v\| = \|w\| = 2, \quad Tv = 3v, \quad Tw = 4w.$$

Show that $\|T(v + w)\| = 10$.

24 Suppose $T \in \mathcal{L}(V)$ and

$$a_0 + a_1 z + a_2 z^2 + \cdots + a_{m-1} z^{m-1} + z^m$$

is the minimal polynomial of T. Prove that the minimal polynomial of T^* is

$$\overline{a_0} + \overline{a_1}\, z + \overline{a_2}\, z^2 + \cdots + \overline{a_{m-1}}\, z^{m-1} + z^m.$$

This exercise shows that the minimal polynomial of T^* equals the minimal polynomial of T if $\mathbf{F} = \mathbf{R}$.

25 Suppose $T \in \mathcal{L}(V)$. Prove that T is diagonalizable if and only if T^* is diagonalizable.

26 Fix $u, x \in V$. Define $T \in \mathcal{L}(V)$ by $Tv = \langle v, u\rangle x$ for every $v \in V$.

(a) Prove that if V is a real vector space, then T is self-adjoint if and only if the list u, x is linearly dependent.

(b) Prove that T is normal if and only if the list u, x is linearly dependent.

27 Suppose $T \in \mathcal{L}(V)$ is normal. Prove that

$$\operatorname{null} T^k = \operatorname{null} T \quad \text{and} \quad \operatorname{range} T^k = \operatorname{range} T$$

for every positive integer k.

28 Suppose $T \in \mathcal{L}(V)$ is normal. Prove that if $\lambda \in \mathbf{F}$, then the minimal polynomial of T is not a polynomial multiple of $(x - \lambda)^2$.

29 Prove or give a counterexample: If $T \in \mathcal{L}(V)$ and there is an orthonormal basis $e_1, \ldots, e_n$ of V such that $\|Te_k\| = \|T^* e_k\|$ for each $k = 1, \ldots, n$, then T is normal.

30 Suppose that $T \in \mathcal{L}(\mathbf{F}^3)$ is normal and $T(1,1,1) = (2,2,2)$. Suppose $(z_1, z_2, z_3) \in \operatorname{null} T$. Prove that $z_1 + z_2 + z_3 = 0$.

31 Fix a positive integer n. In the inner product space of continuous real-valued functions on $[-\pi, \pi]$ with inner product $\langle f, g \rangle = \int_{-\pi}^{\pi} fg$, let

$$V = \operatorname{span}(1, \cos x, \cos 2x, \ldots, \cos nx, \sin x, \sin 2x, \ldots, \sin nx).$$

(a) Define $D \in \mathcal{L}(V)$ by $Df = f'$. Show that $D^* = -D$. Conclude that D is normal but not self-adjoint.

(b) Define $T \in \mathcal{L}(V)$ by $Tf = f''$. Show that T is self-adjoint.

32 Suppose $T \colon V \to W$ is a linear map. Show that under the standard identification of V with V' (see 6.58) and the corresponding identification of W with W', the adjoint map $T^* \colon W \to V$ corresponds to the dual map $T' \colon W' \to V'$. More precisely, show that

$$T'(\varphi_w) = \varphi_{T^* w}$$

for all $w \in W$, where φ_w and $\varphi_{T^* w}$ are defined as in 6.58.