CHAPTER

7

Isaac Newton (1642–1727), as envisioned by British poet and artist William Blake in this 1795 painting.

Operators on Inner Product Spaces

The deepest results related to inner product spaces deal with the subject to which we now turn—operators on inner product spaces. By exploiting properties of the adjoint, we will develop a detailed description of several important classes of operators on inner product spaces.

A new assumption for this chapter is listed in the second bullet point below:

7.1 Notation $\mathbf{F}$, V

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

LEARNING OBJECTIVES FOR THIS CHAPTER

- adjoint
- Spectral Theorem
- positive operators
- isometries
- Polar Decomposition
- Singular Value Decomposition

S. Axler, *Linear Algebra Done Right*, Undergraduate Texts in Mathematics,
DOI 10.1007/978-3-319-11080-6_7

7.A Self-Adjoint and Normal Operators

Adjoints

7.2 Definition ***adjoint***, T^*

Suppose $T \in \mathcal{L}(V, W)$. The ***adjoint*** of T is the function $T^*: W \to V$ 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. We do not need the second meaning in this book. In case you encounter the second meaning for adjoint 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 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 $v \in V$.

7.3 Example Define $T: \mathbf{R}^3 \to \mathbf{R}^2$ by

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

Find a formula for T^*.

Solution Here T^* will be a function from $\mathbf{R}^2$ to $\mathbf{R}^3$. To compute T^*, fix a point $(y_1, y_2) \in \mathbf{R}^2$. Then for every $(x_1, x_2, x_3) \in \mathbf{R}^3$ we have

$$\begin{aligned}
\langle (x_1, x_2, x_3), T^*(y_1, y_2) \rangle &= \langle T(x_1, x_2, x_3), (y_1, y_2) \rangle \\
&= \langle (x_2 + 3x_3, 2x_1), (y_1, y_2) \rangle \\
&= x_2 y_1 + 3x_3 y_1 + 2x_1 y_2 \\
&= \langle (x_1, x_2, x_3), (2y_2, y_1, 3y_1) \rangle.
\end{aligned}$$

Thus

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

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

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

for every $v \in V$. Find a formula for T^*.

Solution Fix $w \in W$. Then for every $v \in V$ we have

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

Thus

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

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

The proofs of the next two results use a common technique: flip T^* from one side of an inner product to become T on the other side.

7.5 The adjoint is a linear map

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

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

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

which shows that $T^*(w_1 + w_2) = T^* w_1 + T^* w_2$.

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

$$\begin{aligned} \langle v, T^*(\lambda w) \rangle &= \langle Tv, \lambda w \rangle \\ &= \bar{\lambda} \langle Tv, w \rangle \\ &= \bar{\lambda} \langle v, T^* w \rangle \\ &= \langle v, \lambda T^* w \rangle, \end{aligned}$$

which shows that $T^*(\lambda w) = \lambda T^* w$.

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

7.6 Properties of the adjoint

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

(b) $(\lambda T)^* = \bar{\lambda} T^*$ for all $\lambda \in \mathbf{F}$ and $T \in \mathcal{L}(V, W)$;

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

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

(e) $(ST)^* = T^* S^*$ for all $T \in \mathcal{L}(V, W)$ and $S \in \mathcal{L}(W, U)$ (here U is an inner product space over $\mathbf{F}$).

Proof

(a) Suppose $S, T \in \mathcal{L}(V, W)$. If $v \in V$ and $w \in W$, then

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

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

(b) Suppose $\lambda \in \mathbf{F}$ and $T \in \mathcal{L}(V, W)$. If $v \in V$ and $w \in W$, then

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

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

(c) Suppose $T \in \mathcal{L}(V, W)$. If $v \in V$ and $w \in W$, then

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

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

(d) If $v, u \in V$, then

$$\langle v, I^* u\rangle = \langle Iv, u\rangle = \langle v, u\rangle.$$

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

(e) Suppose $T \in \mathcal{L}(V, W)$ and $S \in \mathcal{L}(W, U)$. If $v \in V$ and $u \in U$, then

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

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

The next result shows the relationship between the null space and the range of a linear map and its adjoint. The symbol $\Longleftrightarrow$ used in the proof means "if and only if"; this symbol could also be read to mean "is equivalent to".

7.7 Null space and range of T^*

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

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

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

(c) null $T = (\text{range } T^*)^\perp$;

(d) range $T = (\text{null } T^*)^\perp$.

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

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

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

If we take the orthogonal complement of both sides of (a), we get (d), where we have used 6.51. Replacing T with T^* in (a) gives (c), where we have used 7.6(c). Finally, replacing T with T^* in (d) gives (b). ■

7.8 Definition *conjugate transpose*

The ***conjugate transpose*** of an m-by-n matrix is the n-by-m matrix obtained by interchanging the rows and columns and then taking the complex conjugate of each entry.

7.9 Example

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

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

If $\mathbf{F} = \mathbf{R}$, *then the conjugate transpose of a matrix is the same as its* ***transpose***, *which is the matrix obtained by interchanging the rows and columns.*

The adjoint of a linear map does not depend on a choice of basis. This explains why this book emphasizes adjoints of linear maps instead of conjugate transposes of matrices.

The next result shows how to compute the matrix of T^* from the matrix of T.

Caution: Remember that the result below applies only when we are dealing with orthonormal bases. With respect to nonorthonormal bases, the matrix of T^* does not necessarily equal the conjugate transpose of the matrix of T.

7.10 The matrix of T^*

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

$$\mathcal{M}\big(T^*, (f_1, \dots, f_m), (e_1, \dots, e_n)\big)$$

is the conjugate transpose of

$$\mathcal{M}\big(T, (e_1, \dots, e_n), (f_1, \dots, f_m)\big).$$

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

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 $f_1, \dots, f_m$ 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):

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

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

Replacing T with T^* and interchanging the roles played by the e's and f's, we see 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 $\mathcal{M}(T)$. In other words, $\mathcal{M}(T^*)$ is the conjugate transpose of $\mathcal{M}(T)$. ■

Self-Adjoint Operators

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

7.11 Definition ***self-adjoint***

An operator $T \in \mathcal{L}(V)$ is called ***self-adjoint*** if $T = T^*$. In other words, $T \in \mathcal{L}(V)$ is self-adjoint if and only if

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

for all $v, w \in V$.

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

$$\begin{pmatrix} 2 & b \\ 3 & 7 \end{pmatrix}.$$

Find all numbers b such that T is self-adjoint.

Solution The operator T is self-adjoint if and only if $b = 3$ (because $\mathcal{M}(T) = \mathcal{M}(T^*)$ if and only if $b = 3$; recall that $\mathcal{M}(T^*)$ is the conjugate transpose of $\mathcal{M}(T)$—see 7.10).

You should verify that the sum of two self-adjoint operators is self-adjoint and that the product of a real scalar and a self-adjoint operator is self-adjoint.

A good analogy to keep in mind (especially when $\mathbf{F} = \mathbf{C}$) is that the adjoint on $\mathcal{L}(V)$ plays a role similar to complex conjugation 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.

Some mathematicians use the term ***Hermitian*** *instead of self-adjoint, honoring French mathematician Charles Hermite, who in 1873 published the first proof that e is not a zero of any polynomial with integer coefficients.*

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 $\mathbf{F} = \mathbf{C}$.

7.13 Eigenvalues of self-adjoint operators are real

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 $T(x, y) = (-y, x)$. Obviously Tv is orthogonal to v for every $v \in \mathbf{R}^2$, even though $T \neq 0$.

7.14 Over $\mathbf{C}$, Tv is orthogonal to v for all v only for the 0 operator

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

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

for all $v \in V$. Then $T = 0$.

Proof We have

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

for all $u, w \in V$, 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 $v \in V$. Thus our hypothesis implies that $\langle Tu, w\rangle = 0$ for all $u, w \in V$. This implies that $T = 0$ (take $w = Tu$). ■

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

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.

7.15 Over $\mathbf{C}$, $\langle Tv, v\rangle$ is real for all v only for self-adjoint operators

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

$$\langle Tv, v\rangle \in \mathbf{R}$$

for every $v \in V$.

Proof Let $v \in V$. Then

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

If $\langle Tv, v\rangle \in \mathbf{R}$ for every $v \in V$, then the left side of the equation above equals 0, so $\langle (T-T^*)v, v\rangle = 0$ for every $v \in V$. This implies that $T - T^* = 0$ (by 7.14). Hence T is self-adjoint.

Conversely, if T is self-adjoint, then the right side of the equation above equals 0, so $\langle Tv, v\rangle = \overline{\langle Tv, v\rangle}$ for every $v \in V$. This implies that $\langle Tv, v\rangle \in \mathbf{R}$ for every $v \in V$, as desired. ■

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 If $T = T^*$ and $\langle Tv, v\rangle = 0$ for all v, then $T = 0$

Suppose T is a self-adjoint operator on V such that

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

for all $v \in V$. Then $T = 0$.

Proof We have already proved this (without the hypothesis that T is self-adjoint) when V is a complex inner product space (see 7.14). Thus we can assume that V is a real inner product space. If $u, w \in V$, then

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

this is 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.

Each term on the right side of 7.17 is of the form $\langle Tv, v\rangle$ for appropriate v. Thus $\langle Tu, w\rangle = 0$ for all $u, w \in V$. This implies that $T = 0$ (take $w = Tu$). ■

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
$$TT^* = T^*T.$$

Obviously every self-adjoint operator is normal, because if T is self-adjoint then $T^* = T$.

7.19 Example Let T be the operator on $\mathbf{F}^2$ whose matrix (with respect to the standard basis) is
$$\begin{pmatrix} 2 & -3 \\ 3 & 2 \end{pmatrix}.$$
Show that T is not self-adjoint and that T is normal.

Solution This operator 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.

The next result implies that null T = null T^* *for every normal operator* T.

In the next section we will see why normal operators are worthy of special attention.

The next result provides a simple characterization of normal operators.

7.20 T is normal if and only if $\|Tv\| = \|T^*v\|$ for all v

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

$$\|Tv\| = \|T^*v\|$$

for all $v \in V$.

Proof Let $T \in \mathcal{L}(V)$. We will prove both directions of this result at the same time. Note that

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

where we used 7.16 to establish the second equivalence (note that the operator $T^*T - TT^*$ is self-adjoint). The equivalence of the first and last conditions above gives the desired result. ■

Compare the next corollary to Exercise 2. 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, the next corollary implies that a normal operator and its adjoint have the same eigenvectors.

7.21 For T normal, T and T^* have the same eigenvectors

Suppose $T \in \mathcal{L}(V)$ is normal and $v \in V$ is an eigenvector of T with eigenvalue λ. Then v is also an eigenvector of T^* with eigenvalue $\bar{\lambda}$.

Proof Because T is normal, so is $T - \lambda I$, as you should verify. Using 7.20, we have

$$0 = \|(T - \lambda I)v\| = \|(T - \lambda I)^*v\| = \|(T^* - \bar{\lambda} I)v\|.$$

Hence v is an eigenvector of T^* with eigenvalue $\bar{\lambda}$, as desired. ■

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 we have $T^*v = \bar{\beta}v$. Thus

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

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

EXERCISES 7.A

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 $T \in \mathcal{L}(V)$ and $\lambda \in \mathbf{F}$. Prove that λ is an eigenvalue of T if and only if $\bar{\lambda}$ is an eigenvalue of T^*.

3 Suppose $T \in \mathcal{L}(V)$ and U is a subspace of V. Prove that U is invariant under T if and only if $U^\perp$ is invariant under T^*.

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

(a) T is injective if and only if T^* is surjective;

(b) T is surjective if and only if T^* is injective.

5 Prove that

$$\dim \operatorname{null} T^* = \dim \operatorname{null} T + \dim W - \dim V$$

and

$$\dim \operatorname{range} T^* = \dim \operatorname{range} T$$

for every $T \in \mathcal{L}(V, W)$.

6 Make $\mathcal{P}_2(\mathbf{R})$ into an inner product space by defining

$$\langle p, q \rangle = \int_0^1 p(x)q(x)\,dx.$$

Define $T \in \mathcal{L}\big(\mathcal{P}_2(\mathbf{R})\big)$ by $T(a_0 + a_1 x + a_2 x^2) = a_1 x$.

(a) Show that 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.

7 Suppose $S, T \in \mathcal{L}(V)$ are self-adjoint. Prove that ST is self-adjoint if and only if $ST = TS$.

8 Suppose V is a real inner product space. Show that the set of self-adjoint operators on V is a subspace of $\mathcal{L}(V)$.

9 Suppose V is a complex inner product space with $V \neq \{0\}$. Show that the set of self-adjoint operators on V is not a subspace of $\mathcal{L}(V)$.

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

11 Suppose $P \in \mathcal{L}(V)$ is such that $P^2 = P$. Prove that there is a subspace U of V such that $P = P_U$ if and only if P is self-adjoint.

12 Suppose that T is a normal operator on V and that 3 and 4 are eigenvalues of T. Prove that there exists a vector $v \in V$ such that $\|v\| = \sqrt{2}$ and $\|Tv\| = 5$.

13 Give an example of an operator $T \in \mathcal{L}(\mathbf{C}^4)$ such that T is normal but not self-adjoint.

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

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

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

for every $v \in V$.

(a) Suppose $\mathbf{F} = \mathbf{R}$. Prove that T is self-adjoint if and only if u, x is linearly dependent.

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

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

$$\operatorname{range} T = \operatorname{range} T^*.$$

17 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.

18 Prove or give a counterexample: If $T \in \mathcal{L}(V)$ and there exists an orthonormal basis $e_1, \dots, e_n$ of V such that $\|Te_j\| = \|T^*e_j\|$ for each j, then T is normal.

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

20 Suppose $T \in \mathcal{L}(V, W)$ and $\mathbf{F} = \mathbf{R}$. Let Φ_V be the isomorphism from V onto the dual space V' given by Exercise 17 in Section 6.B, and let Φ_W be the corresponding isomorphism from W onto W'. Show that if Φ_V and Φ_W are used to identify V and W with V' and W', then T^* is identified with the dual map T'. More precisely, show that $\Phi_V \circ T^* = T' \circ \Phi_W$.

21 Fix a positive integer n. In the inner product space of continuous real-valued functions on $[-\pi, \pi]$ with inner product

$$\langle f, g \rangle = \int_{-\pi}^{\pi} f(x)g(x)\,dx,$$

let

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

(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.

7.B *The Spectral Theorem*

Recall that a diagonal matrix is a square matrix that is 0 everywhere except possibly along the diagonal. Recall also that an operator on V has a diagonal matrix with respect to a basis if and only if the basis consists of eigenvectors of the operator (see 5.41).

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 normal operators when $\mathbf{F} = \mathbf{C}$ and as the self-adjoint operators when $\mathbf{F} = \mathbf{R}$. The Spectral Theorem is probably the most useful tool in the study of operators on inner product spaces.

Because the conclusion of the Spectral Theorem depends on $\mathbf{F}$, we will break the Spectral Theorem into two pieces, called the Complex Spectral Theorem and the Real Spectral Theorem. As is often the case in linear algebra, complex vector spaces are easier to deal with than real vector spaces. Thus we present the Complex Spectral Theorem first.

The Complex Spectral Theorem

The key part of the Complex Spectral Theorem (7.24) states 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. The next example illustrates this conclusion.

7.23 **Example** Consider the normal operator $T \in \mathcal{L}(\mathbf{C}^2)$ from Example 7.19, whose matrix (with respect to the standard basis) is

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

As you can verify, $\frac{(i,1)}{\sqrt{2}}, \frac{(-i,1)}{\sqrt{2}}$ 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

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

In the next result, the equivalence of (b) and (c) is easy (see 5.41). Thus we prove only that (c) implies (a) and that (a) implies (c).

7.24 Complex Spectral Theorem

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

(a) T is normal.

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

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

Proof First suppose (c) 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.

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

7.25
$$\mathcal{M}\big(T, (e_1, \dots, e_n)\big) = \begin{pmatrix} a_{1,1} & \dots & a_{1,n} \\ & \ddots & \vdots \\ 0 & & a_{n,n} \end{pmatrix}.$$

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

We see from the matrix above that
$$\|Te_1\|^2 = |a_{1,1}|^2$$
and
$$\|T^*e_1\|^2 = |a_{1,1}|^2 + |a_{1,2}|^2 + \cdots + |a_{1,n}|^2.$$
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 7.25, except possibly the first entry $a_{1,1}$, equal 0.

Now from 7.25 we see that
$$\|Te_2\|^2 = |a_{2,2}|^2$$
(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 7.25, except possibly the diagonal entry $a_{2,2}$, equal 0.

Continuing in this fashion, we see that all the nondiagonal entries in the matrix 7.25 equal 0. Thus (c) holds. ■

The Real Spectral Theorem

We will need a few preliminary results, which apply to both real and complex inner product spaces, for our proof of the Real Spectral Theorem.

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

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

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

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), we are led to the result below.

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 invertible.

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 (6.15). 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 3.69). ■

We know that every operator, self-adjoint or not, on a finite-dimensional nonzero complex vector space has an eigenvalue (see 5.21). Thus the next result tells us something new only for real inner product spaces.

7.27 Self-adjoint operators have eigenvalues

Suppose $V \neq \{0\}$ and $T \in \mathcal{L}(V)$ is a self-adjoint operator. Then T has an eigenvalue.

Proof We can assume that V is a real inner product space, as we have already noted. Let $n = \dim V$ and choose $v \in V$ with $v \neq 0$. Then

$$v, Tv, T^2v, \ldots, T^nv$$

cannot be linearly independent, because V has dimension n and we have $n+1$ vectors. Thus there exist real numbers $a_0, \ldots, a_n$, not all 0, such that

$$0 = a_0v + a_1Tv + \cdots + a_nT^nv.$$

Make the a's the coefficients of a polynomial, which can be written in factored form (see 4.17) as

$$\begin{aligned} &a_0 + a_1x + \cdots + a_nx^n \\ &\quad = c(x^2 + b_1x + c_1)\cdots(x^2 + b_Mx + c_M)(x - \lambda_1)\cdots(x - \lambda_m), \end{aligned}$$

where c is a nonzero real number, each b_j, c_j, and λ_j is real, each ${b_j}^2$ is less than $4c_j$, $m + M \geq 1$, and the equation holds for all real x. We then have

$$\begin{aligned} 0 &= a_0v + a_1Tv + \cdots + a_nT^nv \\ &= (a_0I + a_1T + \cdots + a_nT^n)v \\ &= c(T^2 + b_1T + c_1I)\cdots(T^2 + b_MT + c_MI)(T - \lambda_1I)\cdots(T - \lambda_mI)v. \end{aligned}$$

By 7.26, each $T^2 + b_jT + c_jI$ is invertible. Recall also that $c \neq 0$. Thus the equation above implies that $m > 0$ and

$$0 = (T - \lambda_1I)\cdots(T - \lambda_mI)v.$$

Hence $T - \lambda_jI$ is not injective for at least one j. In other words, T has an eigenvalue. ■

The next result shows that if U is a subspace of V that is invariant under a self-adjoint operator T, then $U^\perp$ is also invariant under T. Later we will show that the hypothesis that T is self-adjoint can be replaced with the weaker hypothesis that T is normal (see 9.30).

7.28 Self-adjoint operators and invariant subspaces

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

(a) $U^{\perp}$ is invariant under T;

(b) $T|_U \in \mathcal{L}(U)$ is self-adjoint;

(c) $T|_{U^{\perp}} \in \mathcal{L}(U^{\perp})$ is self-adjoint.

Proof To prove (a), suppose $v \in U^{\perp}$. Let $u \in U$. Then

$$\langle Tv, u \rangle = \langle v, Tu \rangle = 0,$$

where the first equality above holds because T is self-adjoint and the second equality above holds because U is invariant under T (and hence $Tu \in U$) and because $v \in U^{\perp}$. Because the equation above holds for each $u \in U$, we conclude that $Tv \in U^{\perp}$. Thus $U^{\perp}$ is invariant under T, completing the proof of (a).

To prove (b), note that if $u, v \in U$, then

$$\langle (T|_U)u, v \rangle = \langle Tu, v \rangle = \langle u, Tv \rangle = \langle u, (T|_U)v \rangle.$$

Thus $T|_U$ is self-adjoint.

Now (c) follows from replacing U with $U^{\perp}$ in (b), which makes sense by (a). ■

We can now prove the next result, which 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) V has an orthonormal basis consisting of eigenvectors of T.

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

Proof First suppose (c) holds, so T has a diagonal matrix with respect to some orthonormal basis of V. A diagonal matrix equals its transpose. Hence $T = T^*$, and thus T is self-adjoint. In other words, (a) holds.

We will prove that (a) implies (b) by induction on $\dim V$. To get started, note that if $\dim V = 1$, then (a) implies (b). Now assume that $\dim V > 1$ and that (a) implies (b) for all real inner product spaces of smaller dimension.

Suppose (a) holds, so $T \in \mathcal{L}(V)$ is self-adjoint. Let u be an eigenvector of T with $\|u\| = 1$ (7.27 guarantees that T has an eigenvector, which can then be divided by its norm to produce an eigenvector with norm 1). Let $U = \operatorname{span}(u)$. Then U is a 1-dimensional subspace of V that is invariant under T. By 7.28(c), the operator $T|_{U^\perp} \in \mathcal{L}(U^\perp)$ is self-adjoint.

By our induction hypothesis, there is an orthonormal basis of $U^\perp$ consisting of eigenvectors of $T|_{U^\perp}$. Adjoining u to this orthonormal basis of $U^\perp$ gives an orthonormal basis of V consisting of eigenvectors of T, completing the proof that (a) implies (b).

We have proved that (c) implies (a) and that (a) implies (b). Clearly (b) implies (c), completing the proof. ■

7.30 **Example** Consider the self-adjoint operator T on $\mathbf{R}^3$ whose matrix (with respect to the standard basis) is

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

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, and with respect to this basis, the matrix of T is the diagonal matrix

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

If $\mathbf{F} = \mathbf{C}$, then the Complex Spectral Theorem gives a complete description of the normal operators on V. A complete description of the self-adjoint operators on V then easily follows (they are the normal operators on V whose eigenvalues all are real; see Exercise 6).

If $\mathbf{F} = \mathbf{R}$, then the Real Spectral Theorem gives a complete description of the self-adjoint operators on V. In Chapter 9, we will give a complete description of the normal operators on V (see 9.34).

EXERCISES 7.B

1 True or false (and give a proof of your answer): There exists $T \in \mathcal{L}(\mathbf{R}^3)$ such that T is not self-adjoint (with respect to the usual inner product) and such that there is a basis of $\mathbf{R}^3$ consisting of eigenvectors of T.

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

3 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 $T^2 - 5T + 6I \neq 0$.

4 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.

5 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.

6 Prove that a normal operator on a complex inner product space is self-adjoint if and only if all its eigenvalues are real.
[*The exercise above strengthens the analogy (for normal operators) between self-adjoint operators and real numbers.*]

7 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 $T^2 = T$.

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

9 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 $S^2 = T$.)

10 Give an example of a real inner product space V and $T \in \mathcal{L}(V)$ and real numbers b, c with $b^2 < 4c$ such that $T^2 + bT + cI$ is not invertible.
[*The exercise above shows that the hypothesis that T is self-adjoint is needed in 7.26, even for real vector spaces.*]

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

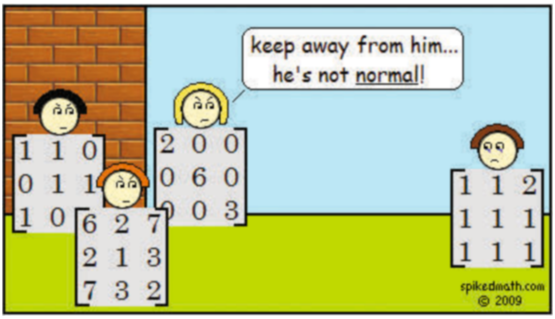
$$\|Tv - \lambda v\| < \epsilon.$$

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

13 Give an alternative proof of the Complex Spectral Theorem that avoids Schur's Theorem and instead follows the pattern of the proof of the Real Spectral Theorem.

14 Suppose U is a finite-dimensional real vector space and $T \in \mathcal{L}(U)$. Prove that U has a basis consisting of eigenvectors of T if and only if there is an inner product on U that makes T into a self-adjoint operator.

15 Find the matrix entry below that is covered up.



7.C *Positive Operators and Isometries*

Positive Operators

7.31 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 is self-adjoint can be dropped from the definition above (by 7.15).

7.32 Example ***positive operators***

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

(b) 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.

7.33 Definition ***square root***

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

7.34 Example 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.

The positive operators correspond to the numbers $[0, \infty)$, so better terminology would use the term nonnegative instead of positive. However, operator theorists consistently call these the positive operators, so we will follow that custom.

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

7.35 Characterization of positive operators

Let $T \in \mathcal{L}(V)$. Then the following are equivalent:

(a) T is positive;

(b) T is self-adjoint and all the eigenvalues of T are nonnegative;

(c) T has a positive square root;

(d) T has a self-adjoint square root;

(e) there exists an operator $R \in \mathcal{L}(V)$ such that $T = R^* R$.

Proof We will prove that (a) $\Rightarrow$ (b) $\Rightarrow$ (c) $\Rightarrow$ (d) $\Rightarrow$ (e) $\Rightarrow$ (a).

First suppose (a) holds, so that T is positive. Obviously T is self-adjoint (by the definition of a positive operator). To prove the other condition in (b), suppose λ is an eigenvalue of T. Let v be an eigenvector of T corresponding to λ. Then

$$0 \leq \langle Tv, v \rangle = \langle \lambda v, v \rangle = \lambda \langle v, v \rangle.$$

Thus λ is a nonnegative number. Hence (b) holds.

Now suppose (b) holds, so that T is self-adjoint and all the eigenvalues of T are nonnegative. By the Spectral Theorem (7.24 and 7.29), there is an orthonormal basis $e_1, \ldots, e_n$ of V consisting of eigenvectors of T. Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of T corresponding to $e_1, \ldots, e_n$,; thus each λ_j is a nonnegative number. Let R be the linear map from V to V such that

$$Re_j = \sqrt{\lambda_j} e_j$$

for $j = 1, \ldots, n$ (see 3.5). Then R is a positive operator, as you should verify. Furthermore, $R^2 e_j = \lambda_j e_j = Te_j$ for each j, which implies that $R^2 = T$. Thus R is a positive square root of T. Hence (c) holds.

Clearly (c) implies (d) (because, by definition, every positive operator is self-adjoint).

Now suppose (d) holds, meaning that there exists a self-adjoint operator R on V such that $T = R^2$. Then $T = R^* R$ (because $R^* = R$). Hence (e) holds.

Finally, suppose (e) holds. Let $R \in \mathcal{L}(V)$ be such that $T = R^* R$. Then $T^* = (R^* R)^* = R^* (R^*)^* = R^* R = T$. Hence T is self-adjoint. To complete the proof that (a) holds, note that

$$\langle Tv, v \rangle = \langle R^* Rv, v \rangle = \langle Rv, Rv \rangle \geq 0$$

for every $v \in V$. Thus T is positive. ■

Each nonnegative number has a unique nonnegative square root. The next result shows that positive operators enjoy a similar property.

Some mathematicians also use the term ***positive semidefinite operator****, which means the same as positive operator.*

7.36 Each positive operator has only one positive square root

Every positive operator on V has a unique positive square root.

Proof Suppose $T \in \mathcal{L}(V)$ is positive. Suppose $v \in V$ is an eigenvector of T. Thus there exists $\lambda \geq 0$ such that $Tv = \lambda v$.

A positive operator can have infinitely many square roots (although only one of them can be positive). For example, the identity operator on V has infinitely many square roots if $\dim V > 1$.

Let R be a positive square root of T. We will prove that $Rv = \sqrt{\lambda} v$. This will imply that the behavior of R on the eigenvectors of T is uniquely determined. Because there is a basis of V consisting of eigenvectors of T (by the Spectral Theorem), this will imply that R is uniquely determined.

To prove that $Rv = \sqrt{\lambda} v$, note that the Spectral Theorem asserts that there is an orthonormal basis $e_1, \dots, e_n$ of V consisting of eigenvectors of R. Because R is a positive operator, all its eigenvalues are nonnegative. Thus there exist nonnegative numbers $\lambda_1, \dots, \lambda_n$ such that $Re_j = \sqrt{\lambda_j} e_j$ for $j = 1, \dots, n$.

Because $e_1, \dots, e_n$ is a basis of V, we can write

$$v = a_1 e_1 + \cdots + a_n e_n$$

for some numbers $a_1, \dots, a_n \in \mathbf{F}$. Thus

$$Rv = a_1 \sqrt{\lambda_1} e_1 + \cdots + a_n \sqrt{\lambda_n} e_n$$

and hence

$$R^2 v = a_1 \lambda_1 e_1 + \cdots + a_n \lambda_n e_n.$$

But $R^2 = T$, and $Tv = \lambda v$. Thus the equation above implies

$$a_1 \lambda e_1 + \cdots + a_n \lambda e_n = a_1 \lambda_1 e_1 + \cdots + a_n \lambda_n e_n.$$

The equation above implies that $a_j(\lambda - \lambda_j) = 0$ for $j = 1, \dots, n$. Hence

$$v = \sum_{\{j : \lambda_j = \lambda\}} a_j e_j,$$

and thus

$$Rv = \sum_{\{j : \lambda_j = \lambda\}} a_j \sqrt{\lambda} e_j = \sqrt{\lambda} v,$$

as desired. ∎

Isometries

Operators that preserve norms are sufficiently important to deserve a name:

7.37 Definition ***isometry***

- An operator $S \in \mathcal{L}(V)$ is called an ***isometry*** if
$$\|Sv\| = \|v\|$$
for all $v \in V$.
- In other words, an operator is an isometry if it preserves norms.

*The Greek word **isos** means equal; the Greek word **metron** means measure. Thus **isometry** literally means equal measure.*

For example, λI is an isometry whenever $\lambda \in \mathbf{F}$ satisfies $|\lambda| = 1$. We will see soon that if $\mathbf{F} = \mathbf{C}$, then the next example includes all isometries.

7.38 Example Suppose $\lambda_1, \dots, \lambda_n$ are scalars with absolute value 1 and $S \in \mathcal{L}(V)$ satisfies $Se_j = \lambda_j e_j$ for some orthonormal basis $e_1, \dots, e_n$ of V. Show that S is an isometry.

Solution Suppose $v \in V$. Then

7.39 $$v = \langle v, e_1 \rangle e_1 + \cdots + \langle v, e_n \rangle e_n$$

and

7.40 $$\|v\|^2 = |\langle v, e_1 \rangle|^2 + \cdots + |\langle v, e_n \rangle|^2,$$

where we have used 6.30. Applying S to both sides of 7.39 gives

$$\begin{aligned} Sv &= \langle v, e_1 \rangle Se_1 + \cdots + \langle v, e_n \rangle Se_n \\ &= \lambda_1 \langle v, e_1 \rangle e_1 + \cdots + \lambda_n \langle v, e_n \rangle e_n. \end{aligned}$$

The last equation, along with the equation $|\lambda_j| = 1$, shows that

7.41 $$\|Sv\|^2 = |\langle v, e_1 \rangle|^2 + \cdots + |\langle v, e_n \rangle|^2.$$

Comparing 7.40 and 7.41 shows that $\|v\| = \|Sv\|$. In other words, S is an isometry.

The next result provides several conditions that are equivalent to being an isometry. The equivalence of (a) and (b) shows that an operator is an isometry if and only if it preserves inner products. The equivalence of (a) and (c) [or (d)] shows that an operator is an isometry if and only if the list of columns of its matrix with respect to every [or some] basis is orthonormal. Exercise 10 implies that in the previous sentence we can replace "columns" with "rows".

*An isometry on a real inner product space is often called an **orthogonal** operator. An isometry on a complex inner product space is often called a **unitary** operator. We use the term isometry so that our results can apply to both real and complex inner product spaces.*

7.42 Characterization of isometries

Suppose $S \in \mathcal{L}(V)$. Then the following are equivalent:

(a) S is an isometry;

(b) $\langle Su, Sv\rangle = \langle u, v\rangle$ for all $u, v \in V$;

(c) $Se_1, \dots, Se_n$ is orthonormal for every orthonormal list of vectors $e_1, \dots, e_n$ in V;

(d) there exists an orthonormal basis $e_1, \dots, e_n$ of V such that $Se_1, \dots, Se_n$ is orthonormal;

(e) $S^*S = I$;

(f) $SS^* = I$;

(g) S^* is an isometry;

(h) S is invertible and $S^{-1} = S^*$.

Proof First suppose (a) holds, so S is an isometry. Exercises 19 and 20 in Section 6.A show that inner products can be computed from norms. Because S preserves norms, this implies that S preserves inner products, and hence (b) holds. More precisely, if V is a real inner product space, then for every $u, v \in V$ we have

$$\begin{aligned}\langle Su, Sv\rangle &= (\|Su + Sv\|^2 - \|Su - Sv\|^2)/4\\ &= (\|S(u + v)\|^2 - \|S(u - v)\|^2)/4\\ &= (\|u + v\|^2 - \|u - v\|^2)/4\\ &= \langle u, v\rangle,\end{aligned}$$

where the first equality comes from Exercise 19 in Section 6.A, the second equality comes from the linearity of S, the third equality holds because S is an isometry, and the last equality again comes from Exercise 19 in Section 6.A. If V is a complex inner product space, then use Exercise 20 in Section 6.A instead of Exercise 19 to obtain the same conclusion. In either case, we see that (b) holds.

Now suppose (b) holds, so S preserves inner products. Suppose that $e_1, \dots, e_n$ is an orthonormal list of vectors in V. Then we see that the list $Se_1, \dots, Se_n$ is orthonormal because $\langle Se_j, Se_k \rangle = \langle e_j, e_k \rangle$. Thus (c) holds.

Clearly (c) implies (d).

Now suppose (d) holds. Let $e_1, \dots, e_n$ be an orthonormal basis of V such that $Se_1, \dots, Se_n$ is orthonormal. Thus

$$\langle S^*Se_j, e_k \rangle = \langle e_j, e_k \rangle$$

for $j, k = 1, \dots, n$ [because the term on the left equals $\langle Se_j, Se_k \rangle$ and $(Se_1, \dots, Se_n)$ is orthonormal]. All vectors $u, v \in V$ can be written as linear combinations of $e_1, \dots, e_n$, and thus the equation above implies that $\langle S^*Su, v \rangle = \langle u, v \rangle$. Hence $S^*S = I$; in other words, (e) holds.

Now suppose (e) holds, so that $S^*S = I$. In general, an operator S need not commute with S^*. However, $S^*S = I$ if and only if $SS^* = I$; this is a special case of Exercise 10 in Section 3.D. Thus $SS^* = I$, showing that (f) holds.

Now suppose (f) holds, so $SS^* = I$. If $v \in V$, then

$$\|S^*v\|^2 = \langle S^*v, S^*v \rangle = \langle SS^*v, v \rangle = \langle v, v \rangle = \|v\|^2.$$

Thus S^* is an isometry, showing that (g) holds.

Now suppose (g) holds, so S^* is an isometry. We know that (a) $\Rightarrow$ (e) and (a) $\Rightarrow$ (f) because we have shown (a) $\Rightarrow$ (b) $\Rightarrow$ (c) $\Rightarrow$ (d) $\Rightarrow$ (e) $\Rightarrow$ (f). Using the implications (a) $\Rightarrow$ (e) and (a) $\Rightarrow$ (f) but with S replaced with S^* [and using the equation $(S^*)^* = S$], we conclude that $SS^* = I$ and $S^*S = I$. Thus S is invertible and $S^{-1} = S^*$; in other words, (h) holds.

Now suppose (h) holds, so S is invertible and $S^{-1} = S^*$. Thus $S^*S = I$. If $v \in V$, then

$$\|Sv\|^2 = \langle Sv, Sv \rangle = \langle S^*Sv, v \rangle = \langle v, v \rangle = \|v\|^2.$$

Thus S is an isometry, showing that (a) holds.

We have shown (a) $\Rightarrow$ (b) $\Rightarrow$ (c) $\Rightarrow$ (d) $\Rightarrow$ (e) $\Rightarrow$ (f) $\Rightarrow$ (g) $\Rightarrow$ (h) $\Rightarrow$ (a), completing the proof. ■

The previous result shows that every isometry is normal [see (a), (e), and (f) of 7.42]. Thus characterizations of normal operators can be used to give descriptions of isometries. We do this in the next result in the complex case and in Chapter 9 in the real case (see 9.36).

7.43 Description of isometries when $\mathbf{F} = \mathbf{C}$

Suppose V is a complex inner product space and $S \in \mathcal{L}(V)$. Then the following are equivalent:

(a) S is an isometry.

(b) There is an orthonormal basis of V consisting of eigenvectors of S whose corresponding eigenvalues all have absolute value 1.

Proof We have already shown (see Example 7.38) that (b) implies (a).

To prove the other direction, suppose (a) holds, so S is an isometry. By the Complex Spectral Theorem (7.24), there is an orthonormal basis $e_1, \dots, e_n$ of V consisting of eigenvectors of S. For $j \in \{1, \dots, n\}$, let λ_j be the eigenvalue corresponding to e_j. Then

$$|\lambda_j| = \|\lambda_j e_j\| = \|Se_j\| = \|e_j\| = 1.$$

Thus each eigenvalue of S has absolute value 1, completing the proof. ■

EXERCISES 7.C

1 Prove or give a counterexample: If $T \in \mathcal{L}(V)$ is self-adjoint and there exists an orthonormal basis $e_1, \dots, e_n$ of V such that $\langle Te_j, e_j \rangle \geq 0$ for each j, then T is a positive operator.

2 Suppose T is a positive operator on V. Suppose $v, w \in V$ are such that

$$Tv = w \quad \text{and} \quad Tw = v.$$

Prove that $v = w$.

3 Suppose T is a positive operator on V and U is a subspace of V invariant under T. Prove that $T|_U \in \mathcal{L}(U)$ is a positive operator on U.

4 Suppose $T \in \mathcal{L}(V, W)$. Prove that T^*T is a positive operator on V and TT^* is a positive operator on W.

5 Prove that the sum of two positive operators on V is positive.

6 Suppose $T \in \mathcal{L}(V)$ is positive. Prove that T^k is positive for every positive integer k.

7 Suppose T is a positive operator on V. Prove that T is invertible if and only if

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

for every $v \in V$ with $v \neq 0$.

8 Suppose $T \in \mathcal{L}(V)$. For $u, v \in V$, define $\langle u, v \rangle_T$ by

$$\langle u, v \rangle_T = \langle Tu, v \rangle.$$

Prove that $\langle \cdot, \cdot \rangle_T$ is an inner product on V if and only if T is an invertible positive operator (with respect to the original inner product $\langle \cdot, \cdot \rangle$).

9 Prove or disprove: the identity operator on $\mathbf{F}^2$ has infinitely many self-adjoint square roots.

10 Suppose $S \in \mathcal{L}(V)$. Prove that the following are equivalent:

(a) S is an isometry;

(b) $\langle S^*u, S^*v \rangle = \langle u, v \rangle$ for all $u, v \in V$;

(c) $S^*e_1, \dots, S^*e_m$ is an orthonormal list for every orthonormal list of vectors $e_1, \dots, e_m$ in V;

(d) $S^*e_1, \dots, S^*e_n$ is an orthonormal basis for some orthonormal basis $e_1, \dots, e_n$ of V.

11 Suppose T_1, T_2 are normal operators on $\mathcal{L}(\mathbf{F}^3)$ and both operators have $2, 5, 7$ as eigenvalues. Prove that there exists an isometry $S \in \mathcal{L}(\mathbf{F}^3)$ such that $T_1 = S^*T_2S$.

12 Give an example of two self-adjoint operators $T_1, T_2 \in \mathcal{L}(\mathbf{F}^4)$ such that the eigenvalues of both operators are $2, 5, 7$ but there does not exist an isometry $S \in \mathcal{L}(\mathbf{F}^4)$ such that $T_1 = S^*T_2S$. Be sure to explain why there is no isometry with the required property.

13 Prove or give a counterexample: if $S \in \mathcal{L}(V)$ and there exists an orthonormal basis $e_1, \dots, e_n$ of V such that $\|Se_j\| = 1$ for each e_j, then S is an isometry.

14 Let T be the second derivative operator in Exercise 21 in Section 7.A. Show that $-T$ is a positive operator.

7.D Polar Decomposition and Singular Value Decomposition

Polar Decomposition

Recall our analogy between $\mathbf{C}$ and $\mathcal{L}(V)$. Under this analogy, a complex number z corresponds to an operator T, and $\bar{z}$ corresponds to T^*. The real numbers ($z = \bar{z}$) correspond to the self-adjoint operators ($T = T^*$), and the nonnegative numbers correspond to the (badly named) positive operators.

Another distinguished subset of $\mathbf{C}$ is the unit circle, which consists of the complex numbers z such that $|z| = 1$. The condition $|z| = 1$ is equivalent to the condition $\bar{z}z = 1$. Under our analogy, this would correspond to the condition $T^*T = I$, which is equivalent to T being an isometry (see 7.42). In other words, the unit circle in $\mathbf{C}$ corresponds to the isometries.

Continuing with our analogy, note that each complex number z except 0 can be written in the form

$$z = \left(\frac{z}{|z|}\right)|z| = \left(\frac{z}{|z|}\right)\sqrt{\bar{z}z},$$

where the first factor, namely, $z/|z|$, is an element of the unit circle. Our analogy leads us to guess that each operator $T \in \mathcal{L}(V)$ can be written as an isometry times $\sqrt{T^*T}$. That guess is indeed correct, as we now prove after defining the obvious notation, which is justified by 7.36.

7.44 Notation $\sqrt{T}$

If T is a positive operator, then $\sqrt{T}$ denotes the unique positive square root of T.

Now we can state and prove the Polar Decomposition, which gives a beautiful description of an arbitrary operator on V. Note that T^*T is a positive operator for every $T \in \mathcal{L}(V)$, and thus $\sqrt{T^*T}$ is well defined.

7.45 Polar Decomposition

Suppose $T \in \mathcal{L}(V)$. Then there exists an isometry $S \in \mathcal{L}(V)$ such that

$$T = S\sqrt{T^*T}.$$

Proof If $v \in V$, then

$$\begin{aligned}\|Tv\|^2 = \langle Tv, Tv\rangle = \langle T^*Tv, v\rangle &= \langle \sqrt{T^*T}\sqrt{T^*T}v, v\rangle \\ &= \langle \sqrt{T^*T}v, \sqrt{T^*T}v\rangle \\ &= \|\sqrt{T^*T}v\|^2.\end{aligned}$$

Thus

7.46 $$\|Tv\| = \|\sqrt{T^*T}v\|$$

for all $v \in V$.

Define a linear map $S_1\colon \operatorname{range}\sqrt{T^*T} \to \operatorname{range} T$ by

7.47 $$S_1(\sqrt{T^*T}v) = Tv.$$

The idea of the proof is to extend S_1 to an isometry $S \in \mathcal{L}(V)$ such that $T = S\sqrt{T^*T}$. Now for the details.

First we must check that S_1 is well defined. To do this, suppose $v_1, v_2 \in V$ are such that $\sqrt{T^*T}v_1 = \sqrt{T^*T}v_2$. For the definition given by 7.47 to make sense, we must show that $Tv_1 = Tv_2$. Note that

$$\begin{aligned}\|Tv_1 - Tv_2\| &= \|T(v_1 - v_2)\| \\ &= \|\sqrt{T^*T}(v_1 - v_2)\| \\ &= \|\sqrt{T^*T}v_1 - \sqrt{T^*T}v_2\| \\ &= 0,\end{aligned}$$

where the second equality holds by 7.46. The equation above shows that $Tv_1 = Tv_2$, so S_1 is indeed well defined. You should verify that S_1 is a linear map.

We see from 7.47 that S_1 maps range $\sqrt{T^*T}$ onto range T. Clearly 7.46 and 7.47 imply that

$$\|S_1 u\| = \|u\|$$

for all $u \in \operatorname{range}\sqrt{T^*T}$.

The rest of the proof extends S_1 to an isometry S on all of V.

In particular, S_1 is injective. Thus from the Fundamental Theorem of Linear Maps (3.22), applied to S_1, we have

$$\dim\operatorname{range}\sqrt{T^*T} = \dim\operatorname{range} T.$$

This implies that $\dim(\operatorname{range}\sqrt{T^*T})^\perp = \dim(\operatorname{range} T)^\perp$ (see 6.50). Thus orthonormal bases $e_1, \dots, e_m$ of $(\operatorname{range}\sqrt{T^*T})^\perp$ and $f_1, \dots, f_m$ of $(\operatorname{range} T)^\perp$ can be chosen; the key point here is that these two orthonormal bases have the same length (denoted m). Now define a linear map $S_2 \colon (\operatorname{range}\sqrt{T^*T})^\perp \to (\operatorname{range} T)^\perp$ by

$$S_2(a_1 e_1 + \cdots + a_m e_m) = a_1 f_1 + \cdots + a_m f_m.$$

For all $w \in (\operatorname{range}\sqrt{T^*T})^\perp$, we have $\|S_2 w\| = \|w\|$ (from 6.25).

Now let S be the operator on V that equals S_1 on $\operatorname{range}\sqrt{T^*T}$ and equals S_2 on $(\operatorname{range}\sqrt{T^*T})^\perp$. More precisely, recall that each $v \in V$ can be written uniquely in the form

7.48
$$v = u + w,$$

where $u \in \operatorname{range}\sqrt{T^*T}$ and $w \in (\operatorname{range}\sqrt{T^*T})^\perp$ (see 6.47). For $v \in V$ with decomposition as above, define Sv by

$$Sv = S_1 u + S_2 w.$$

For each $v \in V$ we have

$$S(\sqrt{T^*T}v) = S_1(\sqrt{T^*T}v) = Tv,$$

so $T = S\sqrt{T^*T}$, as desired. All that remains is to show that S is an isometry. However, this follows easily from two uses of the Pythagorean Theorem: if $v \in V$ has decomposition as in 7.48, then

$$\|Sv\|^2 = \|S_1 u + S_2 w\|^2 = \|S_1 u\|^2 + \|S_2 w\|^2 = \|u\|^2 + \|w\|^2 = \|v\|^2;$$

the second equality holds because $S_1 u \in \operatorname{range} T$ and $S_2 w \in (\operatorname{range} T)^\perp$. ■

The Polar Decomposition (7.45) states that each operator on V is the product of an isometry and a positive operator. Thus we can write each operator on V as the product of two operators, each of which comes from a class that we can completely describe and that we understand reasonably well. The isometries are described by 7.43 and 9.36; the positive operators are described by the Spectral Theorem (7.24 and 7.29).

Specifically, consider the case $\mathbf{F} = \mathbf{C}$, and suppose $T = S\sqrt{T^*T}$ is a Polar Decomposition of an operator $T \in \mathcal{L}(V)$, where S is an isometry. Then there is an orthonormal basis of V with respect to which S has a diagonal matrix, and there is an orthonormal basis of V with respect to which $\sqrt{T^*T}$ has a diagonal matrix. **Warning:** there may not exist an orthonormal basis that simultaneously puts the matrices of both S and $\sqrt{T^*T}$ into these nice diagonal forms. In other words, S may require one orthonormal basis and $\sqrt{T^*T}$ may require a different orthonormal basis.

Singular Value Decomposition

The eigenvalues of an operator tell us something about the behavior of the operator. Another collection of numbers, called the singular values, is also useful. Recall that eigenspaces and the notation E are defined in 5.36.

7.49 Definition ***singular values***

Suppose $T \in \mathcal{L}(V)$. The ***singular values*** of T are the eigenvalues of $\sqrt{T^*T}$, with each eigenvalue λ repeated $\dim E(\lambda, \sqrt{T^*T})$ times.

The singular values of T are all nonnegative, because they are the eigenvalues of the positive operator $\sqrt{T^*T}$.

7.50 Example Define $T \in \mathcal{L}(\mathbf{F}^4)$ by

$$T(z_1, z_2, z_3, z_4) = (0, 3z_1, 2z_2, -3z_4).$$

Find the singular values of T.

Solution A calculation shows $T^*T(z_1, z_2, z_3, z_4) = (9z_1, 4z_2, 0, 9z_4)$, as you should verify. Thus

$$\sqrt{T^*T}(z_1, z_2, z_3, z_4) = (3z_1, 2z_2, 0, 3z_4),$$

and we see that the eigenvalues of $\sqrt{T^*T}$ are $3, 2, 0$ and

$$\dim E(3, \sqrt{T^*T}) = 2, \ \dim E(2, \sqrt{T^*T}) = 1, \ \dim E(0, \sqrt{T^*T}) = 1.$$

Hence the singular values of T are $3, 3, 2, 0$.

Note that -3 and 0 are the only eigenvalues of T. Thus in this case, the collection of eigenvalues did not pick up the number 2 that appears in the definition (and hence the behavior) of T, but the collection of singular values does include 2.

Each $T \in \mathcal{L}(V)$ has $\dim V$ singular values, as can be seen by applying the Spectral Theorem and 5.41 [see especially part (e)] to the positive (hence self-adjoint) operator $\sqrt{T^*T}$. For example, the operator T defined in Example 7.50 on the four-dimensional vector space $\mathbf{F}^4$ has four singular values (they are $3, 3, 2, 0$), as we saw above.

The next result shows that every operator on V has a clean description in terms of its singular values and two orthonormal bases of V.

7.51 Singular Value Decomposition

Suppose $T \in \mathcal{L}(V)$ has singular values $s_1, \ldots, s_n$. Then there exist orthonormal bases $e_1, \ldots, e_n$ and $f_1, \ldots, f_n$ of V such that

$$Tv = s_1 \langle v, e_1 \rangle f_1 + \cdots + s_n \langle v, e_n \rangle f_n$$

for every $v \in V$.

Proof By the Spectral Theorem applied to $\sqrt{T^*T}$, there is an orthonormal basis $e_1, \ldots, e_n$ of V such that $\sqrt{T^*T} e_j = s_j e_j$ for $j = 1, \ldots, n$.

We have

$$v = \langle v, e_1 \rangle e_1 + \cdots + \langle v, e_n \rangle e_n$$

for every $v \in V$ (see 6.30). Apply $\sqrt{T^*T}$ to both sides of this equation, getting

$$\sqrt{T^*T} v = s_1 \langle v, e_1 \rangle e_1 + \cdots + s_n \langle v, e_n \rangle e_n$$

for every $v \in V$. By the Polar Decomposition (see 7.45), there is an isometry $S \in \mathcal{L}(V)$ such that $T = S\sqrt{T^*T}$. Apply S to both sides of the equation above, getting

$$Tv = s_1 \langle v, e_1 \rangle S e_1 + \cdots + s_n \langle v, e_n \rangle S e_n$$

for every $v \in V$. For each j, let $f_j = S e_j$. Because S is an isometry, $f_1, \ldots, f_n$ is an orthonormal basis of V (see 7.42). The equation above now becomes

$$Tv = s_1 \langle v, e_1 \rangle f_1 + \cdots + s_n \langle v, e_n \rangle f_n$$

for every $v \in V$, completing the proof. ■

When we worked with linear maps from one vector space to a second vector space, we considered the matrix of a linear map with respect to a basis of the first vector space and a basis of the second vector space. When dealing with operators, which are linear maps from a vector space to itself, we almost always use only one basis, making it play both roles.

The Singular Value Decomposition allows us a rare opportunity to make good use of two different bases for the matrix of an operator. To do this, suppose $T \in \mathcal{L}(V)$. Let $s_1, \ldots, s_n$ denote the singular values of T, and let $e_1, \ldots, e_n$ and $f_1, \ldots, f_n$ be orthonormal bases of V such that the Singular Value Decomposition 7.51 holds. Because $T e_j = s_j f_j$ for each j, we have

$$\mathcal{M}\big(T, (e_1, \ldots, e_n), (f_1, \ldots, f_n)\big) = \begin{pmatrix} s_1 & & 0 \\ & \ddots & \\ 0 & & s_n \end{pmatrix}.$$

In other words, every operator on V has a diagonal matrix with respect to some orthonormal bases of V, provided that we are permitted to use two different bases rather than a single basis as customary when working with operators.

Singular values and the Singular Value Decomposition have many applications (some are given in the exercises), including applications in computational linear algebra. To compute numeric approximations to the singular values of an operator T, first compute T^*T and then compute approximations to the eigenvalues of T^*T (good techniques exist for approximating eigenvalues of positive operators). The nonnegative square roots of these (approximate) eigenvalues of T^*T will be the (approximate) singular values of T. In other words, the singular values of T can be approximated without computing the square root of T^*T. The next result helps justify working with T^*T instead of $\sqrt{T^*T}$.

7.52 Singular values without taking square root of an operator

Suppose $T \in \mathcal{L}(V)$. Then the singular values of T are the nonnegative square roots of the eigenvalues of T^*T, with each eigenvalue λ repeated $\dim E(\lambda, T^*T)$ times.

Proof The Spectral Theorem implies that there are an orthonormal basis $e_1, \dots, e_n$ and nonnegative numbers $\lambda_1, \dots, \lambda_n$ such that $T^*Te_j = \lambda_j e_j$ for $j = 1, \dots, n$. It is easy to see that $\sqrt{T^*T}e_j = \sqrt{\lambda_j}e_j$ for $j = 1, \dots, n$, which implies the desired result. ■

EXERCISES 7.D

1 Fix $u, x \in V$ with $u \neq 0$. Define $T \in \mathcal{L}(V)$ by

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

for every $v \in V$. Prove that

$$\sqrt{T^*T}v = \frac{\|x\|}{\|u\|}\langle v, u \rangle u$$

for every $v \in V$.

2 Give an example of $T \in \mathcal{L}(\mathbf{C}^2)$ such that 0 is the only eigenvalue of T and the singular values of T are $5, 0$.

3 Suppose $T \in \mathcal{L}(V)$. Prove that there exists an isometry $S \in \mathcal{L}(V)$ such that
$$T = \sqrt{TT^*}\, S.$$

4 Suppose $T \in \mathcal{L}(V)$ and s is a singular value of T. Prove that there exists a vector $v \in V$ such that $\|v\| = 1$ and $\|Tv\| = s$.

5 Suppose $T \in \mathcal{L}(\mathbf{C}^2)$ is defined by $T(x, y) = (-4y, x)$. Find the singular values of T.

6 Find the singular values of the differentiation operator $D \in \mathcal{P}(\mathbf{R}^2)$ defined by $Dp = p'$, where the inner product on $\mathcal{P}(\mathbf{R}^2)$ is as in Example 6.33.

7 Define $T \in \mathcal{L}(\mathbf{F}^3)$ by
$$T(z_1, z_2, z_3) = (z_3, 2z_1, 3z_2).$$
Find (explicitly) an isometry $S \in \mathcal{L}(\mathbf{F}^3)$ such that $T = S\sqrt{T^*T}$.

8 Suppose $T \in \mathcal{L}(V)$, $S \in \mathcal{L}(V)$ is an isometry, and $R \in \mathcal{L}(V)$ is a positive operator such that $T = SR$. Prove that $R = \sqrt{T^*T}$.
[*The exercise above shows that if we write* T *as the product of an isometry and a positive operator* (*as in the Polar Decomposition 7.45*), *then the positive operator equals* $\sqrt{T^*T}$.]

9 Suppose $T \in \mathcal{L}(V)$. Prove that T is invertible if and only if there exists a unique isometry $S \in \mathcal{L}(V)$ such that $T = S\sqrt{T^*T}$.

10 Suppose $T \in \mathcal{L}(V)$ is self-adjoint. Prove that the singular values of T equal the absolute values of the eigenvalues of T, repeated appropriately.

11 Suppose $T \in \mathcal{L}(V)$. Prove that T and T^* have the same singular values.

12 Prove or give a counterexample: if $T \in \mathcal{L}(V)$, then the singular values of T^2 equal the squares of the singular values of T.

13 Suppose $T \in \mathcal{L}(V)$. Prove that T is invertible if and only if 0 is not a singular value of T.

14 Suppose $T \in \mathcal{L}(V)$. Prove that $\dim \operatorname{range} T$ equals the number of nonzero singular values of T.

15 Suppose $S \in \mathcal{L}(V)$. Prove that S is an isometry if and only if all the singular values of S equal 1.

16 Suppose $T_1, T_2 \in \mathcal{L}(V)$. Prove that T_1 and T_2 have the same singular values if and only if there exist isometries $S_1, S_2 \in \mathcal{L}(V)$ such that $T_1 = S_1 T_2 S_2$.

17 Suppose $T \in \mathcal{L}(V)$ has singular value decomposition given by

$$Tv = s_1 \langle v, e_1 \rangle f_1 + \cdots + s_n \langle v, e_n \rangle f_n$$

for every $v \in V$, where $s_1, \ldots, s_n$ are the singular values of T and $e_1, \ldots, e_n$ and $f_1, \ldots, f_n$ are orthonormal bases of V.

(a) Prove that if $v \in V$, then

$$T^* v = s_1 \langle v, f_1 \rangle e_1 + \cdots + s_n \langle v, f_n \rangle e_n.$$

(b) Prove that if $v \in V$, then

$$T^* T v = {s_1}^2 \langle v, e_1 \rangle e_1 + \cdots + {s_n}^2 \langle v, e_n \rangle e_n.$$

(c) Prove that if $v \in V$, then

$$\sqrt{T^* T} v = s_1 \langle v, e_1 \rangle e_1 + \cdots + s_n \langle v, e_n \rangle e_n.$$

(d) Suppose T is invertible. Prove that if $v \in V$, then

$$T^{-1} v = \frac{\langle v, f_1 \rangle e_1}{s_1} + \cdots + \frac{\langle v, f_n \rangle e_n}{s_n}$$

for every $v \in V$.

18 Suppose $T \in \mathcal{L}(V)$. Let $\hat{s}$ denote the smallest singular value of T, and let s denote the largest singular value of T.

(a) Prove that $\hat{s}\|v\| \leq \|Tv\| \leq s\|v\|$ for every $v \in V$.

(b) Suppose λ is an eigenvalue of T. Prove that $\hat{s} \leq |\lambda| \leq s$.

19 Suppose $T \in \mathcal{L}(V)$. Show that T is uniformly continuous with respect to the metric d on V defined by $d(u, v) = \|u - v\|$.

20 Suppose $S, T \in \mathcal{L}(V)$. Let s denote the largest singular value of S, let t denote the largest singular value of T, and let r denote the largest singular value of $S + T$. Prove that $r \leq s + t$.