CHAPTER 7

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.

Recall that **F** denotes **R** or **C**.
Let's agree that for this chapter
V is a finite-dimensional, nonzero, inner-product space over **F**.

Self-Adjoint and Normal Operators

Instead of self-adjoint, some mathematicians use the term ***Hermitian*** *(in honor of the French mathematician Charles Hermite, who in 1873 published the first proof that e is not the root of any polynomial with integer coefficients).*

An operator $T \in \mathcal{L}(V)$ is called ***self-adjoint*** if $T = T^*$. For example, if T is the operator on $\mathbf{F}^2$ whose matrix (with respect to the standard basis) is

$$\begin{bmatrix} 2 & b \\ 3 & 7 \end{bmatrix},$$

then 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 6.47).

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. We will see that this analogy is reflected in some important properties of self-adjoint operators, beginning with eigenvalues.

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

7.1 Proposition: *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

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

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

The next proposition 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 is not 0.

7.2 Proposition: *If V is a complex inner-product space and T is an operator on V such that*

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

for all $v \in V$, then $T = 0$.

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

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

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 $v \in V$. If $\langle Tv, v\rangle = 0$ for all $v \in V$, then the equation above implies that $\langle Tu, w\rangle = 0$ for all $u, w \in V$. This implies that $T = 0$ (take $w = Tu$). ■

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

7.3 Corollary: *Let V be a complex inner-product space and let $T \in \mathcal{L}(V)$. Then T is self-adjoint if and only if*

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

for every $v \in V$.

This corollary provides another example of how self-adjoint operators behave like real numbers.

PROOF: Let $v \in V$. Then

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

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 7.2), and 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 may satisfy $\langle Tv, v\rangle = 0$ for all $v \in V$. However, the next proposition shows that this cannot happen for a self-adjoint operator.

7.4 Proposition: *If 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 7.2). Thus we can assume that V is a real inner-product space and that T is a self-adjoint operator on V. For $u, w \in V$, we have

7.5 $$\langle Tu, w\rangle = \frac{\langle T(u+w), u+w\rangle - \langle T(u-w), u-w\rangle}{4};$$

this is proved by computing the right side, using

$$\begin{aligned}\langle Tw, u\rangle &= \langle w, Tu\rangle \\ &= \langle Tu, w\rangle,\end{aligned}$$

where the first equality holds because T is self-adjoint and the second equality holds because we are working on a real inner-product space. If $\langle Tv, v\rangle = 0$ for all $v \in V$, then 7.5 implies that $\langle Tu, w\rangle = 0$ for all $u, w \in V$. This implies that $T = 0$ (take $w = Tu$). ■

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. For an example of a normal operator that is not self-adjoint, consider the operator on $\mathbf{F}^2$ whose matrix (with respect to the standard basis) is

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

Clearly this operator is not self-adjoint, but an easy calculation (which you should do) shows that it is normal.

We will soon see why normal operators are worthy of special attention. The next proposition provides a simple characterization of normal operators.

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

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

for all $v \in V$.

Note that this proposition implies that $\operatorname{null} T = \operatorname{null} T^*$ for every normal operator T.

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 7.4 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 28 in the previous chapter. That exercise implies that the eigenvalues of the adjoint of any 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.

7.7 Corollary: *Suppose $T \in \mathcal{L}(V)$ is normal. If $v \in V$ is an eigenvector of T with eigenvalue $\lambda \in \mathbf{F}$, then v is also an eigenvector of T^* with eigenvalue $\bar{\lambda}$.*

PROOF: Suppose $v \in V$ is an eigenvector of T with eigenvalue λ. Thus $(T - \lambda I)v = 0$. Because T is normal, so is $T - \lambda I$, as you should verify. Using 7.6, we have

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

and 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.

7.8 Corollary: *If $T \in \mathcal{L}(V)$ is normal, then eigenvectors of T corresponding to distinct eigenvalues are orthogonal.*

PROOF: Suppose $T \in \mathcal{L}(V)$ is normal and α, β are distinct eigenvalues of T, with corresponding eigenvectors u, v. Thus $Tu = \alpha u$ and $Tv = \beta v$. From 7.7 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. ■

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 some basis if and only if there is a basis of V consisting of eigenvectors of the operator (see 5.21).

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, so we present the complex spectral theorem first.

As an illustration of the complex spectral theorem, consider the normal operator $T \in \mathcal{L}(\mathbf{C}^2)$ whose matrix (with respect to the standard basis) is

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

You should verify that

$$\left(\frac{(i,1)}{\sqrt{2}}, \frac{(-i,1)}{\sqrt{2}}\right)$$

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

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

7.9 Complex Spectral Theorem: *Suppose that V is a complex inner-product space and $T \in \mathcal{L}(V)$. Then V has an orthonormal basis consisting of eigenvectors of T if and only if T is normal.*

Because every self-adjoint operator is normal, the complex spectral theorem implies that every self-adjoint operator on a finite-dimensional complex inner-product space has a diagonal matrix with respect to some orthonormal basis.

PROOF: First suppose that V has an orthonormal basis consisting of eigenvectors of T. With respect to this basis, T has a diagonal matrix. 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 must be normal, as desired.

To prove the other direction, now suppose that T is normal. There is an orthonormal basis $(e_1, \dots, e_n)$ of V with respect to which T has an upper-triangular matrix (by 6.28). Thus we can write

7.10 $$\mathcal{M}\big(T, (e_1, \dots, e_n)\big) = \begin{bmatrix} a_{1,1} & \dots & a_{1,n} \\ & \ddots & \vdots \\ 0 & & a_{n,n} \end{bmatrix}.$$

We will show that this matrix is actually a diagonal matrix, which means that $(e_1, \dots, e_n)$ is an orthonormal basis of V consisting of eigenvectors of T.

We see from the matrix above that

$$\|Te_1\|^2 = |a_{1,1}|^2$$

and

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

Because T is normal, $\|Te_1\| = \|T^*e_1\|$ (see 7.6). Thus the two equations above imply that all entries in the first row of the matrix in 7.10, except possibly the first entry $a_{1,1}$, equal 0.

Now from 7.10 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 7.10, except possibly the diagonal entry $a_{2,2}$, equal 0.

Continuing in this fashion, we see that all the nondiagonal entries in the matrix 7.10 equal 0, as desired. ■

We will need two lemmas for our proof of the real spectral theorem. You could guess that the next lemma is true and even discover its proof by thinking about quadratic polynomials with real coefficients. Specifically, suppose $\alpha, \beta \in \mathbf{R}$ and $\alpha^2 < 4\beta$. Let x be a real number. Then

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

$$\begin{aligned} x^2 + \alpha x + \beta &= (x + \frac{\alpha}{2})^2 + (\beta - \frac{\alpha^2}{4}) \\ &> 0. \end{aligned}$$

In particular, $x^2 + \alpha x + \beta$ 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 lemma below.

7.11 Lemma: *Suppose $T \in \mathcal{L}(V)$ is self-adjoint. If $\alpha, \beta \in \mathbf{R}$ are such that $\alpha^2 < 4\beta$, then*

$$T^2 + \alpha T + \beta I$$

is invertible.

PROOF: Suppose $\alpha, \beta \in \mathbf{R}$ are such that $\alpha^2 < 4\beta$. Let v be a nonzero vector in V. Then

$$\begin{aligned} \langle (T^2 + \alpha T + \beta I)v, v\rangle &= \langle T^2 v, v\rangle + \alpha\langle Tv, v\rangle + \beta\langle v, v\rangle \\ &= \langle Tv, Tv\rangle + \alpha\langle Tv, v\rangle + \beta\|v\|^2 \\ &\geq \|Tv\|^2 - |\alpha|\|Tv\|\|v\| + \beta\|v\|^2 \\ &= (\|Tv\| - \frac{|\alpha|\|v\|}{2})^2 + (\beta - \frac{\alpha^2}{4})\|v\|^2 \\ &> 0, \end{aligned}$$

where the first inequality holds by the Cauchy-Schwarz inequality (6.6). The last inequality implies that $(T^2 + \alpha T + \beta I)v \neq 0$. Thus $T^2 + \alpha T + \beta I$ is injective, which implies that it is invertible (see 3.21). ■

We have proved that every operator, self-adjoint or not, on a finite-dimensional complex vector space has an eigenvalue (see 5.10), so the next lemma tells us something new only for real inner-product spaces.

7.12 Lemma: *Suppose $T \in \mathcal{L}(V)$ is self-adjoint. Then T has an eigenvalue.*

Proof: As noted above, we can assume that V is a real inner-product space. Let $n = \dim V$ and choose $v \in V$ with $v \neq 0$. Then

$$(v, Tv, T^2 v, \dots, T^n v)$$

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

$$0 = a_0 v + a_1 Tv + \cdots + a_n T^n v.$$

Here we are imitating the proof that T has an invariant subspace of dimension 1 or 2 (see 5.24).

Make the a's the coefficients of a polynomial, which can be written in factored form (see 4.14) as

$$\begin{aligned} &a_0 + a_1 x + \cdots + a_n x^n \\ &\quad = c(x^2 + \alpha_1 x + \beta_1) \dots (x^2 + \alpha_M x + \beta_M)(x - \lambda_1) \dots (x - \lambda_m), \end{aligned}$$

where c is a nonzero real number, each α_j, β_j, and λ_j is real, each ${\alpha_j}^2 < 4\beta_j$, $m + M \geq 1$, and the equation holds for all real x. We then have

$$\begin{aligned} 0 &= a_0 v + a_1 Tv + \cdots + a_n T^n v \\ &= (a_0 I + a_1 T + \cdots + a_n T^n) v \\ &= c(T^2 + \alpha_1 T + \beta_1 I) \dots (T^2 + \alpha_M T + \beta_M I)(T - \lambda_1 I) \dots (T - \lambda_m I) v. \end{aligned}$$

Each $T^2 + \alpha_j T + \beta_j I$ is invertible because T is self-adjoint and each ${\alpha_j}^2 < 4\beta_j$ (see 7.11). Recall also that $c \neq 0$. Thus the equation above implies that

$$0 = (T - \lambda_1 I) \dots (T - \lambda_m I) v.$$

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

As an illustration of the real spectral theorem, consider the self-adjoint operator T on $\mathbf{R}^3$ whose matrix (with respect to the standard basis) is

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

You should verify that

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

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

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

Combining the complex spectral theorem and the real spectral theorem, we conclude that every self-adjoint operator on V has a diagonal matrix with respect to some orthonormal basis. This statement, which is the most useful part of the spectral theorem, holds regardless of whether $\mathbf{F} = \mathbf{C}$ or $\mathbf{F} = \mathbf{R}$.

7.13 Real Spectral Theorem: *Suppose that V is a real inner-product space and $T \in \mathcal{L}(V)$. Then V has an orthonormal basis consisting of eigenvectors of T if and only if T is self-adjoint.*

PROOF: First suppose that V has an orthonormal basis consisting of eigenvectors of T. With respect to this basis, T has a diagonal matrix. This matrix equals its conjugate transpose. Hence $T = T^*$ and so T is self-adjoint, as desired.

To prove the other direction, now suppose that T is self-adjoint. We will prove that V has an orthonormal basis consisting of eigenvectors of T by induction on the dimension of V. To get started, note that our desired result clearly holds if $\dim V = 1$. Now assume that $\dim V > 1$ and that the desired result holds on vector spaces of smaller dimension.

The idea of the proof is to take any eigenvector u of T with norm 1, then adjoin to it an orthonormal basis of eigenvectors of $T|_{\{u\}^\perp}$. Now for the details, the most important of which is verifying that $T|_{\{u\}^\perp}$ is self-adjoint (this allows us to apply our induction hypothesis).

Let λ be any eigenvalue of T (because T is self-adjoint, we know from the previous lemma that it has an eigenvalue) and let $u \in V$ denote a corresponding eigenvector with $\|u\| = 1$. Let U denote the one-dimensional subspace of V consisting of all scalar multiples of u. Note that a vector $v \in V$ is in $U^\perp$ if and only if $\langle u, v\rangle = 0$.

To get an eigenvector of norm 1, take any nonzero eigenvector and divide it by its norm.

Suppose $v \in U^\perp$. Then because T is self-adjoint, we have

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

and hence $Tv \in U^\perp$. Thus $Tv \in U^\perp$ whenever $v \in U^\perp$. In other words, $U^\perp$ is invariant under T. Thus we can define an operator $S \in \mathcal{L}(U^\perp)$ by $S = T|_{U^\perp}$. If $v, w \in U^\perp$, then

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

which shows that S is self-adjoint (note that in the middle equality above we used the self-adjointness of T). Thus, by our induction hypothesis, there is an orthonormal basis of $U^\perp$ consisting of eigenvectors of S. Clearly every eigenvector of S is an eigenvector of T (because $Sv = Tv$ for every $v \in U^\perp$). Thus adjoining u to an orthonormal basis of $U^\perp$ consisting of eigenvectors of S gives an orthonormal basis of V consisting of eigenvectors of T, as desired. ∎

For $T \in \mathcal{L}(V)$ self-adjoint (or, more generally, $T \in \mathcal{L}(V)$ normal when $\mathbf{F} = \mathbf{C}$), the corollary below provides the nicest possible decomposition of V into subspaces invariant under T. On each $\operatorname{null}(T - \lambda_j I)$, the operator T is just multiplication by λ_j.

7.14 Corollary: *Suppose that $T \in \mathcal{L}(V)$ is self-adjoint (or that $\mathbf{F} = \mathbf{C}$ and that $T \in \mathcal{L}(V)$ is normal). Let $\lambda_1, \dots, \lambda_m$ denote the distinct eigenvalues of T. Then*

$$V = \operatorname{null}(T - \lambda_1 I) \oplus \cdots \oplus \operatorname{null}(T - \lambda_m I).$$

Furthermore, each vector in each $\operatorname{null}(T - \lambda_j I)$ is orthogonal to all vectors in the other subspaces of this decomposition.

Proof: The spectral theorem (7.9 and 7.13) implies that V has a basis consisting of eigenvectors of T. The desired decomposition of V now follows from 5.21.

The orthogonality statement follows from 7.8. ∎

Normal Operators on Real Inner-Product Spaces

The complex spectral theorem (7.9) gives a complete description of normal operators on complex inner-product spaces. In this section we will give a complete description of normal operators on real inner-product spaces. Along the way, we will encounter a proposition (7.18) and a technique (block diagonal matrices) that are useful for both real and complex inner-product spaces.

We begin with a description of the operators on a two-dimensional real inner-product space that are normal but not self-adjoint.

7.15 Lemma: *Suppose V is a two-dimensional real inner-product space and $T \in \mathcal{L}(V)$. Then the following are equivalent:*

(a) *T is normal but not self-adjoint;*

(b) *the matrix of T with respect to every orthonormal basis of V has the form*
$$\begin{bmatrix} a & -b \\ b & a \end{bmatrix},$$
with $b \neq 0$;

(c) *the matrix of T with respect to some orthonormal basis of V has the form*
$$\begin{bmatrix} a & -b \\ b & a \end{bmatrix},$$
with $b > 0$.

PROOF: First suppose that (a) holds, so that T is normal but not self-adjoint. Let (e_1, e_2) be an orthonormal basis of V. Suppose

7.16
$$\mathcal{M}(T, (e_1, e_2)) = \begin{bmatrix} a & c \\ b & d \end{bmatrix}.$$

Then $\|Te_1\|^2 = a^2 + b^2$ and $\|T^*e_1\|^2 = a^2 + c^2$. Because T is normal, $\|Te_1\| = \|T^*e_1\|$ (see 7.6); thus these equations imply that $b^2 = c^2$. Thus $c = b$ or $c = -b$. But $c \neq b$ because otherwise T would be self-adjoint, as can be seen from the matrix in 7.16. Hence $c = -b$, so

7.17
$$\mathcal{M}(T, (e_1, e_2)) = \begin{bmatrix} a & -b \\ b & d \end{bmatrix}.$$

Of course, the matrix of T^* is the transpose of the matrix above. Use matrix multiplication to compute the matrices of TT^* and T^*T (do it now). Because T is normal, these two matrices must be equal. Equating the entries in the upper-right corner of the two matrices you computed, you will discover that $bd = ab$. Now $b \neq 0$ because otherwise T would be self-adjoint, as can be seen from the matrix in 7.17. Thus $d = a$, completing the proof that (a) implies (b).

Now suppose that (b) holds. We want to prove that (c) holds. Choose any orthonormal basis (e_1, e_2) of V. We know that the matrix of T with respect to this basis has the form given by (b), with $b \neq 0$. If $b > 0$, then (c) holds and we have proved that (b) implies (c). If $b < 0$, then, as you should verify, the matrix of T with respect to the orthonormal basis $(e_1, -e_2)$ equals $\begin{bmatrix} a & b \\ -b & a \end{bmatrix}$, where $-b > 0$; thus in this case we also see that (b) implies (c).

Now suppose that (c) holds, so that the matrix of T with respect to some orthonormal basis has the form given in (c) with $b > 0$. Clearly the matrix of T is not equal to its transpose (because $b \neq 0$), and hence T is not self-adjoint. Now use matrix multiplication to verify that the matrices of TT^* and T^*T are equal. We conclude that $TT^* = T^*T$, and hence T is normal. Thus (c) implies (a), completing the proof. ■

As an example of the notation we will use to write a matrix as a matrix of smaller matrices, consider the matrix

$$D = \begin{bmatrix} 1 & 1 & 2 & 2 & 2 \\ 1 & 1 & 2 & 2 & 2 \\ 0 & 0 & 3 & 3 & 3 \\ 0 & 0 & 3 & 3 & 3 \\ 0 & 0 & 3 & 3 & 3 \end{bmatrix}.$$

We can write this matrix in the form

Often we can understand a matrix better by thinking of it as composed of smaller matrices. We will use this technique in the next proposition and in later chapters.

$$D = \begin{bmatrix} A & B \\ 0 & C \end{bmatrix},$$

where

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & 2 & 2 \\ 2 & 2 & 2 \end{bmatrix}, \quad C = \begin{bmatrix} 3 & 3 & 3 \\ 3 & 3 & 3 \\ 3 & 3 & 3 \end{bmatrix},$$

and 0 denotes the 3-by-2 matrix consisting of all 0's.

The next result will play a key role in our characterization of the normal operators on a real inner-product space.

Without normality, an easier result also holds: if $T \in \mathcal{L}(V)$ and U invariant under T, then $U^\perp$ is invariant under T^*; see Exercise 29 in Chapter 6.

7.18 Proposition: *Suppose $T \in \mathcal{L}(V)$ is normal and U is a subspace of V that is invariant under T. Then*

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

(b) *U is invariant under T^*;*

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

(d) *$T|_U$ is a normal operator on U;*

(e) *$T|_{U^\perp}$ is a normal operator on $U^\perp$.*

PROOF: First we will prove (a). Let $(e_1, \dots, e_m)$ be an orthonormal basis of U. Extend to an orthonormal basis $(e_1, \dots, e_m, f_1, \dots, f_n)$ of V (this is possible by 6.25). Because U is invariant under T, each Te_j is a linear combination of $(e_1, \dots, e_m)$. Thus the matrix of T with respect to the basis $(e_1, \dots, e_m, f_1, \dots, f_n)$ is of the form

$$\mathcal{M}(T) = \begin{array}{c c} & \begin{array}{cccccc} e_1 & \dots & e_m & f_1 & \dots & f_n \end{array} \\ \begin{array}{c} e_1 \\ \vdots \\ e_m \\ f_1 \\ \vdots \\ f_n \end{array} & \left[\begin{array}{cc} A & B \\ & \\ 0 & C \end{array} \right] \end{array};$$

here A denotes an m-by-m matrix, 0 denotes the n-by-m matrix consisting of all 0's, B denotes an m-by-n matrix, C denotes an n-by-n matrix, and for convenience the basis has been listed along the top and left sides of the matrix.

For each $j \in \{1, \dots, m\}$, $\|Te_j\|^2$ equals the sum of the squares of the absolute values of the entries in the j^{th} column of A (see 6.17). Hence

7.19 $$\sum_{j=1}^{m} \|Te_j\|^2 = \text{the sum of the squares of the absolute values of the entries of } A.$$

For each $j \in \{1, \dots, m\}$, $\|T^*e_j\|^2$ equals the sum of the squares of the absolute values of the entries in the j^{th} rows of A and B. Hence

7.20 $$\sum_{j=1}^{m} \|T^* e_j\|^2 = \text{the sum of the squares of the absolute values of the entries of } A \text{ and } B.$$

Because T is normal, $\|Te_j\| = \|T^* e_j\|$ for each j (see 7.6); thus

$$\sum_{j=1}^{m} \|Te_j\|^2 = \sum_{j=1}^{m} \|T^* e_j\|^2.$$

This equation, along with 7.19 and 7.20, implies that the sum of the squares of the absolute values of the entries of B must equal 0. In other words, B must be the matrix consisting of all 0's. Thus

7.21 $$\mathcal{M}(T) = \begin{array}{c} \\ e_1 \\ \vdots \\ e_m \\ f_1 \\ \vdots \\ f_n \end{array} \begin{array}{c} \begin{array}{cccccc} e_1 & \dots & e_m & f_1 & \dots & f_n \end{array} \\ \left[\begin{array}{cc} A & 0 \\ & \\ 0 & C \end{array} \right] \end{array}.$$

This representation shows that Tf_k is in the span of $(f_1, \dots, f_n)$ for each k. Because $(f_1, \dots, f_n)$ is a basis of $U^\perp$, this implies that $Tv \in U^\perp$ whenever $v \in U^\perp$. In other words, $U^\perp$ is invariant under T, completing the proof of (a).

To prove (b), note that $\mathcal{M}(T^*)$ has a block of 0's in the lower left corner (because $\mathcal{M}(T)$, as given above, has a block of 0's in the upper right corner). In other words, each $T^* e_j$ can be written as a linear combination of $(e_1, \dots, e_m)$. Thus U is invariant under T^*, completing the proof of (b).

To prove (c), let $S = T|_U$. Fix $v \in U$. Then

$$\begin{aligned} \langle Su, v \rangle &= \langle Tu, v \rangle \\ &= \langle u, T^* v \rangle \end{aligned}$$

for all $u \in U$. Because $T^* v \in U$ (by (b)), the equation above shows that $S^* v = T^* v$. In other words, $(T|_U)^* = (T^*)|_U$, completing the proof of (c).

To prove (d), note that T commutes with T^* (because T is normal) and that $(T|_U)^* = (T^*)|_U$ (by (c)). Thus $T|_U$ commutes with its adjoint and hence is normal, completing the proof of (d).

To prove (e), note that in (d) we showed that the restriction of T to any invariant subspace is normal. However, $U^\perp$ is invariant under T (by (a)), and hence $T|_{U^\perp}$ is normal. ■

The key step in the proof of the last proposition was showing that $\mathcal{M}(T)$ is an appropriate block diagonal matrix; see 7.21.

In proving 7.18 we thought of a matrix as composed of smaller matrices. Now we need to make additional use of that idea. A ***block diagonal matrix*** is a square matrix of the form

$$\begin{bmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_m \end{bmatrix},$$

where $A_1, \dots, A_m$ are square matrices lying along the diagonal and all the other entries of the matrix equal 0. For example, the matrix

7.22
$$A = \begin{bmatrix} 4 & 0 & 0 & 0 & 0 \\ 0 & 2 & -3 & 0 & 0 \\ 0 & 3 & 2 & 0 & 0 \\ 0 & 0 & 0 & 1 & -7 \\ 0 & 0 & 0 & 7 & 1 \end{bmatrix}$$

is a block diagonal matrix with

$$A = \begin{bmatrix} A_1 & & 0 \\ & A_2 & \\ 0 & & A_3 \end{bmatrix},$$

where

7.23
$$A_1 = \begin{bmatrix} 4 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 2 & -3 \\ 3 & 2 \end{bmatrix}, \quad A_3 = \begin{bmatrix} 1 & -7 \\ 7 & 1 \end{bmatrix}.$$

If A and B are block diagonal matrices of the form

$$A = \begin{bmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_m \end{bmatrix}, \quad B = \begin{bmatrix} B_1 & & 0 \\ & \ddots & \\ 0 & & B_m \end{bmatrix},$$

where A_j has the same size as B_j for $j = 1, \dots, m$, then AB is a block diagonal matrix of the form

7.24
$$AB = \begin{bmatrix} A_1B_1 & & 0 \\ & \ddots & \\ 0 & & A_mB_m \end{bmatrix},$$

as you should verify. In other words, to multiply together two block diagonal matrices (with the same size blocks), just multiply together the corresponding entries on the diagonal, as with diagonal matrices.

A diagonal matrix is a special case of a block diagonal matrix where each block has size 1-by-1. At the other extreme, every square matrix is a block diagonal matrix because we can take the first (and only) block to be the entire matrix. Thus to say that an operator has a block diagonal matrix with respect to some basis tells us nothing unless we know something about the size of the blocks. The smaller the blocks, the nicer the operator (in the vague sense that the matrix then contains more 0's). The nicest situation is to have an orthonormal basis that gives a diagonal matrix. We have shown that this happens on a complex inner-product space precisely for the normal operators (see 7.9) and on a real inner-product space precisely for the self-adjoint operators (see 7.13).

Note that if an operator T has a block diagonal matrix with respect to some basis, then the entry in any 1-by-1 block on the diagonal of this matrix must be an eigenvalue of T.

Our next result states that each normal operator on a real inner-product space comes close to having a diagonal matrix—specifically, we get a block diagonal matrix with respect to some orthonormal basis, with each block having size at most 2-by-2. We cannot expect to do better than that because on a real inner-product space there exist normal operators that do not have a diagonal matrix with respect to any basis. For example, the operator $T \in \mathcal{L}(\mathbf{R}^2)$ defined by $T(x, y) = (-y, x)$ is normal (as you should verify) but has no eigenvalues; thus this particular T does not have even an upper-triangular matrix with respect to any basis of $\mathbf{R}^2$.

Note that the matrix in 7.22 is the type of matrix promised by the theorem below. In particular, each block of 7.22 (see 7.23) has size at most 2-by-2 and each of the 2-by-2 blocks has the required form (upper left entry equals lower right entry, lower left entry is positive, and upper right entry equals the negative of lower left entry).

7.25 Theorem: *Suppose that V is a real inner-product space and $T \in \mathcal{L}(V)$. Then T is normal if and only if there is an orthonormal basis of V with respect to which T has a block diagonal matrix where each block is a 1-by-1 matrix or a 2-by-2 matrix of the form*

7.26
$$\begin{bmatrix} a & -b \\ b & a \end{bmatrix},$$

with $b > 0$.

PROOF: To prove the easy direction, first suppose that there is an orthonormal basis of V such that the matrix of T is a block diagonal matrix where each block is a 1-by-1 matrix or a 2-by-2 matrix of the form 7.26. With respect to this basis, the matrix of T commutes with the matrix of T^* (which is the conjugate of the matrix of T), as you should verify (use formula 7.24 for the product of two block diagonal matrices). Thus T commutes with T^*, which means that T is normal.

To prove the other direction, now suppose that T is normal. We will prove our desired result by induction on the dimension of V. To get started, note that our desired result clearly holds if $\dim V = 1$ (trivially) or if $\dim V = 2$ (if T is self-adjoint, use the real spectral theorem 7.13; if T is not self-adjoint, use 7.15).

Now assume that $\dim V > 2$ and that the desired result holds on vector spaces of smaller dimension. Let U be a subspace of V of dimension 1 that is invariant under T if such a subspace exists (in other words, if T has a nonzero eigenvector, let U be the span of this eigenvector). If no such subspace exists, let U be a subspace of V of dimension 2 that is invariant under T (an invariant subspace of dimension 1 or 2 always exists by 5.24).

In a real vector space with dimension 1, there are precisely two vectors with norm 1.

If $\dim U = 1$, choose a vector in U with norm 1; this vector will be an orthonormal basis of U, and of course the matrix of $T|_U$ is a 1-by-1 matrix. If $\dim U = 2$, then $T|_U$ is normal (by 7.18) but not self-adjoint (otherwise $T|_U$, and hence T, would have a nonzero eigenvector; see 7.12), and thus we can choose an orthonormal basis of U with respect to which the matrix of $T|_U$ has the form 7.26 (see 7.15).

Now $U^\perp$ is invariant under T and $T|_{U^\perp}$ is a normal operator on $U^\perp$ (see 7.18). Thus by our induction hypothesis, there is an orthonormal basis of $U^\perp$ with respect to which the matrix of $T|_{U^\perp}$ has the desired form. Adjoining this basis to the basis of U gives an orthonormal basis of V with respect to which the matrix of T has the desired form. ■

Positive Operators

*Many mathematicians also use the term **positive semidefinite operator**, which means the same as 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 $v \in V$. Note that if V is a complex vector space, then the condition that T be self-adjoint can be dropped from this definition (by 7.3).

You should verify that every orthogonal projection is positive. For another set of examples, look at the proof of 7.11, where we showed that if $T \in \mathcal{L}(V)$ is self-adjoint and $\alpha, \beta \in \mathbf{R}$ are such that $\alpha^2 < 4\beta$, then $T^2 + \alpha T + \beta I$ is positive.

An operator S is called a ***square root*** of an operator T if $S^2 = T$. For 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 $S \in \mathcal{L}(\mathbf{F}^3)$ defined by $S(z_1, z_2, z_3) = (z_2, z_3, 0)$ is a square root of T.

The following theorem is the main result about positive operators. Note that its characterizations of the positive operators 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) below. Also, z is nonnegative if and only if it has a real square root, corresponding to condition (d) below. Finally, z is nonnegative if and only if there exists a complex number w such that $z = \bar{w}w$, corresponding to condition (e) below.

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

7.27 Theorem: *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 $S \in \mathcal{L}(V)$ such that $T = S^*S$.*

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

First suppose that (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 that λ is an eigenvalue of T. Let v be a nonzero eigenvector of T corresponding to λ. Then

$$\begin{aligned} 0 \le & \langle Tv, v \rangle \\ &= \langle \lambda v, v \rangle \\ &= \lambda \langle v, v \rangle, \end{aligned}$$

and thus λ is a nonnegative number. Hence (b) holds.

Now suppose that (b) holds, so that T is self-adjoint and all the eigenvalues of T are nonnegative. By the spectral theorem (7.9 and 7.13), 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 $e_1, \ldots, e_n$, so that each λ_j is a nonnegative number. Define $S \in \mathcal{L}(V)$ by

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

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

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

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

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

$$\begin{aligned} \langle Tv, v\rangle &= \langle S^*Sv, v\rangle \\ &= \langle Sv, Sv\rangle \\ &\geq 0 \end{aligned}$$

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

Each nonnegative number has a unique nonnegative square root. The next proposition shows that positive operators enjoy a similar property. Because of this proposition, we can use the notation $\sqrt{T}$ to denote the unique positive square root of a positive operator T, just as $\sqrt{\lambda}$ denotes the unique nonnegative square root of a nonnegative number λ.

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

7.28 Proposition: *Every positive operator on V has a unique positive square root.*

PROOF: Suppose $T \in \mathcal{L}(V)$ is positive. Let $\lambda_1, \ldots, \lambda_m$ denote the distinct eigenvalues of T; because T is positive, all these numbers are nonnegative (by 7.27). Because T is self-adjoint, we have

7.29 $$V = \operatorname{null}(T - \lambda_1 I) \oplus \cdots \oplus \operatorname{null}(T - \lambda_m I);$$

see 7.14.

Now suppose $S \in \mathcal{L}(V)$ is a positive square root of T. Suppose α is an eigenvalue of S. If $v \in \operatorname{null}(S - \alpha I)$, then $Sv = \alpha v$, which implies that

7.30 $$Tv = S^2 v = \alpha^2 v,$$

so $v \in \operatorname{null}(T - \alpha^2 I)$. Thus α^2 is an eigenvalue of T, which means that α^2 must equal some λ_j. In other words, $\alpha = \sqrt{\lambda_j}$ for some j. Furthermore, 7.30 implies that

7.31 $$\operatorname{null}(S - \sqrt{\lambda_j} I) \subset \operatorname{null}(T - \lambda_j I).$$

In the paragraph above, we showed that the only possible eigenvalues for S are $\sqrt{\lambda_1}, \dots, \sqrt{\lambda_m}$. Because S is self-adjoint, this implies that

7.32 $$V = \operatorname{null}(S - \sqrt{\lambda_1} I) \oplus \cdots \oplus \operatorname{null}(S - \sqrt{\lambda_m} I);$$

see 7.14. Now 7.29, 7.32, and 7.31 imply that

$$\operatorname{null}(S - \sqrt{\lambda_j} I) = \operatorname{null}(T - \lambda_j I)$$

for each j. In other words, on $\operatorname{null}(T - \lambda_j I)$, the operator S is just multiplication by $\sqrt{\lambda_j}$. Thus S, the positive square root of T, is uniquely determined by T. ■

Isometries

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

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. For example, λI is an isometry whenever $\lambda \in \mathbf{F}$ satisfies $|\lambda| = 1$. More generally, suppose $\lambda_1, \dots, \lambda_n$ are scalars with absolute value 1 and $S \in \mathcal{L}(V)$ satisfies $S(e_j) = \lambda_j e_j$ for some orthonormal basis $(e_1, \dots, e_n)$ of V. Suppose $v \in V$. Then

7.33 $$v = \langle v, e_1 \rangle e_1 + \cdots + \langle v, e_n \rangle e_n$$

and

7.34 $$\|v\|^2 = |\langle v, e_1 \rangle|^2 + \cdots + |\langle v, e_n \rangle|^2,$$

where we have used 6.17. Applying S to both sides of 7.33 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

7.35 $$\|Sv\|^2 = |\langle v, e_1\rangle|^2 + \cdots + |\langle v, e_n\rangle|^2.$$

Comparing 7.34 and 7.35 shows that $\|v\| = \|Sv\|$. In other words, S is an isometry.

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 will use the term isometry so that our results can apply to both real and complex inner-product spaces.*

For another example, let $\theta \in \mathbf{R}$. Then the operator on $\mathbf{R}^2$ of counterclockwise rotation (centered at the origin) by an angle of θ is an isometry (you should find the matrix of this operator with respect to the standard basis of $\mathbf{R}^2$).

If $S \in \mathcal{L}(V)$ is an isometry, then S is injective (because if $Sv = 0$, then $\|v\| = \|Sv\| = 0$, and hence $v = 0$). Thus every isometry is invertible (by 3.21).

The next theorem provides several conditions that are equivalent to being an isometry. These equivalences have several important interpretations. In particular, the equivalence of (a) and (b) shows that an isometry preserves inner products. Because (a) implies (d), we see that if S is an isometry and $(e_1, \dots, e_n)$ is an orthonormal basis of V, then the columns of the matrix of S (with respect to this basis) are orthonormal; because (e) implies (a), we see that the converse also holds. Because (a) is equivalent to conditions (i) and (j), we see that in the last sentence we can replace "columns" with "rows".

7.36 Theorem: *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) *$S^*S = I$;*

(d) *$(Se_1, \dots, Se_n)$ is orthonormal whenever $(e_1, \dots, e_n)$ is an orthonormal list of vectors in V;*

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

(f) *S^* is an isometry;*

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

(h) $SS^* = I$*;*

(i) $(S^*e_1, \dots, S^*e_n)$ *is orthonormal whenever* $(e_1, \dots, e_n)$ *is an orthonormal list of vectors in* V*;*

(j) *there exists an orthonormal basis* $(e_1, \dots, e_n)$ *of* V *such that* $(S^*e_1, \dots, S^*e_n)$ *is orthonormal.*

PROOF: First suppose that (a) holds. 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 6 in Chapter 6, 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 6 in Chapter 6. If V is a complex inner-product space, then use Exercise 7 in Chapter 6 instead of Exercise 6 to obtain the same conclusion. In either case, we see that (a) implies (b).

Now suppose that (b) holds. Then

$$\begin{aligned}\langle (S^*S - I)u, v\rangle &= \langle Su, Sv\rangle - \langle u, v\rangle\\ &= 0\end{aligned}$$

for every $u, v \in V$. Taking $v = (S^*S - I)u$, we see that $S^*S - I = 0$. Hence $S^*S = I$, proving that (b) implies (c).

Now suppose that (c) holds. Suppose $(e_1, \dots, e_n)$ is an orthonormal list of vectors in V. Then

$$\begin{aligned}\langle Se_j, Se_k\rangle &= \langle S^*Se_j, e_k\rangle\\ &= \langle e_j, e_k\rangle.\end{aligned}$$

Hence $(Se_1, \dots, Se_n)$ is orthonormal, proving that (c) implies (d).

Obviously (d) implies (e).

Now suppose (e) holds. Let $(e_1, \dots, e_n)$ be an orthonormal basis of V such that $(Se_1, \dots, Se_n)$ is orthonormal. If $v \in V$, then

$$\begin{aligned}\|Sv\|^2 &= \|S(\langle v, e_1\rangle e_1 + \cdots + \langle v, e_n\rangle e_n)\|^2\\ &= \|\langle v, e_1\rangle Se_1 + \cdots + \langle v, e_n\rangle Se_n\|^2\\ &= |\langle v, e_1\rangle|^2 + \cdots + |\langle v, e_n\rangle|^2\\ &= \|v\|^2,\end{aligned}$$

where the first and last equalities come from 6.17. Taking square roots, we see that S is an isometry, proving that (e) implies (a).

Having shown that (a) $\Rightarrow$ (b) $\Rightarrow$ (c) $\Rightarrow$ (d) $\Rightarrow$ (e) $\Rightarrow$ (a), we know at this stage that (a) through (e) are all equivalent to each other. Replacing S with S^*, we see that (f) through (j) are all equivalent to each other. Thus to complete the proof, we need only show that one of the conditions in the group (a) through (e) is equivalent to one of the conditions in the group (f) through (j). The easiest way to connect the two groups of conditions is to show that (c) is equivalent to (h). In general, of course, S need not commute with S^*. However, $S^*S = I$ if and only if $SS^* = I$; this is a special case of Exercise 23 in Chapter 3. Thus (c) is equivalent to (h), completing the proof. ■

The last theorem shows that every isometry is normal (see (a), (c), and (h) of 7.36). Thus the characterizations of normal operators can be used to give complete descriptions of isometries. We do this in the next two theorems.

7.37 Theorem: *Suppose V is a complex inner-product space and $S \in \mathcal{L}(V)$. Then S is an isometry if and only if there is an orthonormal basis of V consisting of eigenvectors of S all of whose corresponding eigenvalues have absolute value* 1.

PROOF: We already proved (see the first paragraph of this section) that if there is an orthonormal basis of V consisting of eigenvectors of S all of whose eigenvalues have absolute value 1, then S is an isometry.

To prove the other direction, suppose S is an isometry. By the complex spectral theorem (7.9), there is an orthonormal basis $(e_1, \ldots, e_n)$ of V consisting of eigenvectors of S. For $j \in \{1, \ldots, n\}$, 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. ■

If $\theta \in \mathbf{R}$, then the operator on $\mathbf{R}^2$ of counterclockwise rotation (centered at the origin) by an angle of θ has matrix 7.39 with respect to the standard basis, as you should verify. The next result states that every isometry on a real inner-product space is composed of pieces that look like rotations on two-dimensional subspaces, pieces that equal the identity operator, and pieces that equal multiplication by -1.

7.38 Theorem: *Suppose that V is a real inner-product space and $S \in \mathcal{L}(V)$. Then S is an isometry if and only if there is an orthonormal basis of V with respect to which S has a block diagonal matrix where each block on the diagonal is a 1-by-1 matrix containing 1 or -1 or a 2-by-2 matrix of the form*

This theorem implies that an isometry on an odd-dimensional real inner-product space must have 1 or -1 as an eigenvalue.

7.39
$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix},$$

with $\theta \in (0, \pi)$.

Proof: First suppose that S is an isometry. Because S is normal, there is an orthonormal basis of V such that with respect to this basis S has a block diagonal matrix, where each block is a 1-by-1 matrix or a 2-by-2 matrix of the form

7.40
$$\begin{bmatrix} a & -b \\ b & a \end{bmatrix},$$

with $b > 0$ (see 7.25).

If λ is an entry in a 1-by-1 along the diagonal of the matrix of S (with respect to the basis mentioned above), then there is a basis vector e_j such that $Se_j = \lambda e_j$. Because S is an isometry, this implies that $|\lambda| = 1$. Thus $\lambda = 1$ or $\lambda = -1$ because these are the only real numbers with absolute value 1.

Now consider a 2-by-2 matrix of the form 7.40 along the diagonal of the matrix of S. There are basis vectors e_j, e_{j+1} such that

$$Se_j = ae_j + be_{j+1}.$$

Thus

$$1 = \|e_j\|^2 = \|Se_j\|^2 = a^2 + b^2.$$

The equation above, along with the condition $b > 0$, implies that there exists a number $\theta \in (0, \pi)$ such that $a = \cos\theta$ and $b = \sin\theta$. Thus the matrix 7.40 has the required form 7.39, completing the proof in this direction.

Conversely, now suppose that there is an orthonormal basis of V with respect to which the matrix of S has the form required by the theorem. Thus there is a direct sum decomposition

$$V = U_1 \oplus \cdots \oplus U_m,$$

where each U_j is a subspace of V of dimension 1 or 2. Furthermore, any two vectors belonging to distinct U's are orthogonal, and each $S|_{U_j}$ is an isometry mapping U_j into U_j. If $v \in V$, we can write

$$v = u_1 + \cdots + u_m,$$

where each $u_j \in U_j$. Applying S to the equation above and then taking norms gives

$$\begin{aligned} \|Sv\|^2 &= \|Su_1 + \cdots + Su_m\|^2 \\ &= \|Su_1\|^2 + \cdots + \|Su_m\|^2 \\ &= \|u_1\|^2 + \cdots + \|u_m\|^2 \\ &= \|v\|^2. \end{aligned}$$

Thus S is an isometry, as desired. ■

Polar and Singular-Value Decompositions

Recall our analogy between **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 correspond to the self-adjoint operators, and the nonnegative numbers correspond to the (badly named) positive operators. Another distinguished subset of **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 7.36). In other words, the unit circle in **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 any 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.

7.41 Polar Decomposition: *If $T \in \mathcal{L}(V)$, then there exists an isometry $S \in \mathcal{L}(V)$ such that*

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

If you know a bit of complex analysis, you will recognize the analogy to polar coordinates for complex numbers: every complex number can be written in the form $e^{\theta i}r$, where $\theta \in [0, 2\pi)$ and $r \geq 0$. Note that $e^{\theta i}$ is in the unit circle, corresponding to S being an isometry, and r is nonnegative, corresponding to $\sqrt{T^*T}$ being a positive operator.

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

7.42 $$\|Tv\| = \|\sqrt{T^*T}v\|$$

for all $v \in V$.

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

7.43 $$S_1(\sqrt{T^*T}v) = Tv.$$

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 7.43 to make sense, we must show that $Tv_1 = Tv_2$. However,

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

where the second equality holds by 7.42. 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.

In the rest of the proof all we are doing is extending S_1 to an isometry S on all of V.

We see from 7.43 that S_1 maps range $\sqrt{T^*T}$ onto range T. Clearly 7.42 and 7.43 imply that $\|S_1 u\| = \|u\|$ for all $u \in$ range $\sqrt{T^*T}$. In particular, S_1 is injective. Thus from 3.4, 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 Exercise 15 in Chapter 6). 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. Define a linear map $S_2 \colon (\operatorname{range} \sqrt{T^*T})^\perp \to (\operatorname{range} T)^\perp$ by

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

Obviously $\|S_2 w\| = \|w\|$ for all $w \in (\operatorname{range} \sqrt{T^*T})^\perp$.

Now let S be the operator on V that equals S_1 on 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

7.44 $$v = u + w,$$

where $u \in$ range $\sqrt{T^*T}$ and $w \in (\operatorname{range} \sqrt{T^*T})^\perp$ (see 6.29). 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 the two uses of the Pythagorean theorem: if $v \in V$ has decomposition as in 7.44, then

$$\begin{aligned} \|Sv\|^2 &= \|S_1 u + S_2 w\|^2 \\ &= \|S_1 u\|^2 + \|S_2 w\|^2 \\ &= \|u\|^2 + \|w\|^2 \\ &= \|v\|^2, \end{aligned}$$

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

The polar decomposition (7.41) 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 have completely described and that we understand reasonably well. The isometries are described by 7.37 and 7.38; the positive operators (which are all self-adjoint) are described by the spectral theorem (7.9 and 7.13).

Specifically, suppose $T = S\sqrt{T^*T}$ is the polar decomposition of $T \in \mathcal{L}(V)$, where S is an isometry. Then there is an orthonormal basis of V with respect to which S has a diagonal matrix (if $\mathbf{F} = \mathbf{C}$) or a block diagonal matrix with blocks of size at most 2-by-2 (if $\mathbf{F} = \mathbf{R}$), 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 forms (diagonal or block diagonal with small blocks). In other words, S may require one orthonormal basis and $\sqrt{T^*T}$ may require a different orthonormal basis.

Suppose $T \in \mathcal{L}(V)$. The ***singular values*** of T are the eigenvalues of $\sqrt{T^*T}$, with each eigenvalue λ repeated $\dim \operatorname{null}(\sqrt{T^*T} - \lambda I)$ times. The singular values of T are all nonnegative because they are the eigenvalues of the positive operator $\sqrt{T^*T}$.

For example, if $T \in \mathcal{L}(\mathbf{F}^4)$ is defined by

7.45 $$T(z_1, z_2, z_3, z_4) = (0, 3z_1, 2z_2, -3z_4),$$

then $T^*T(z_1, z_2, z_3, z_4) = (9z_1, 4z_2, 0, 9z_4)$, 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 $3, 2, 0$. Clearly

$$\dim \operatorname{null}(\sqrt{T^*T} - 3I) = 2, \ \dim \operatorname{null}(\sqrt{T^*T} - 2I) = 1, \ \dim \operatorname{null} \sqrt{T^*T} = 1.$$

Hence the singular values of T are $3, 3, 2, 0$. In this example -3 and 0 are the only eigenvalues of T, as you should verify.

Each $T \in \mathcal{L}(V)$ has $\dim V$ singular values, as can be seen by applying the spectral theorem and 5.21 (see especially part (e)) to the positive (hence self-adjoint) operator $\sqrt{T^*T}$. For example, the operator T defined by 7.45 on the four-dimensional vector space $\mathbf{F}^4$ has four singular values (they are $3, 3, 2, 0$), as we saw in the previous paragraph.

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

7.46 Singular-Value Decomposition: *Suppose $T \in \mathcal{L}(V)$ has singular values $s_1, \dots, s_n$. Then there exist orthonormal bases $(e_1, \dots, e_n)$ and $(f_1, \dots, f_n)$ of V such that*

7.47 $$Tv = s_1\langle v, e_1\rangle f_1 + \dots + s_n\langle v, e_n\rangle f_n$$

for every $v \in V$.

PROOF: By the spectral theorem (also see 7.14) applied to $\sqrt{T^*T}$, there is an orthonormal basis $(e_1, \dots, e_n)$ of V such that $\sqrt{T^*T}e_j = s_j e_j$ for $j = 1, \dots, n$. We have

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

for every $v \in V$ (see 6.17). Apply $\sqrt{T^*T}$ to both sides of this equation, getting

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

This proof illustrates the usefulness of the polar decomposition.

for every $v \in V$. By the polar decomposition (see 7.41), 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

$$Tv = s_1\langle v, e_1\rangle Se_1 + \dots + s_n\langle v, e_n\rangle Se_n$$

for every $v \in V$. For each j, let $f_j = Se_j$. Because S is an isometry, $(f_1, \dots, f_n)$ is an orthonormal basis of V (see 7.36). The equation above now becomes

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

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 for the first vector space and a basis for 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 use two different bases for the matrix of an operator. To do this, suppose $T \in \mathcal{L}(V)$. Let $s_1, \dots, s_n$ denote the singular values of T, and let $(e_1, \dots, e_n)$ and $(f_1, \dots, f_n)$ be orthonormal bases of V such that the singular-value decomposition 7.47 holds. Then clearly

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

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 (as can be seen from the proof of 7.28). In other words, the singular values of T can be approximated without computing the square root of T^*T.

Exercises

1. 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 $T \in \mathcal{L}(\mathcal{P}_2(\mathbf{R}))$ by $T(a_0 + a_1x + a_2x^2) = a_1x$.
   (a) Show that T is not self-adjoint.
   (b) The matrix of T with respect to the basis $(1, x, x^2)$ is
$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$
This matrix equals its conjugate transpose, even though T is not self-adjoint. Explain why this is not a contradiction.

2. Prove or give a counterexample: the product of any two self-adjoint operators on a finite-dimensional inner-product space is self-adjoint.

3. (a) Show that if V is a real inner-product space, then the set of self-adjoint operators on V is a subspace of $\mathcal{L}(V)$.
   (b) Show that if V is a complex inner-product space, then the set of self-adjoint operators on V is not a subspace of $\mathcal{L}(V)$.

4. Suppose $P \in \mathcal{L}(V)$ is such that $P^2 = P$. Prove that P is an orthogonal projection if and only if P is self-adjoint.

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

6. Prove that if $T \in \mathcal{L}(V)$ is normal, then
$$\operatorname{range} T = \operatorname{range} T^*.$$

7. Prove that if $T \in \mathcal{L}(V)$ is normal, then
$$\operatorname{null} T^k = \operatorname{null} T \quad \text{and} \quad \operatorname{range} T^k = \operatorname{range} T$$
for every positive integer k.

8. Prove that there does not exist a self-adjoint operator $T \in \mathcal{L}(\mathbf{R}^3)$ such that $T(1,2,3) = (0,0,0)$ and $T(2,5,7) = (2,5,7)$.

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

Exercise 9 strengthens the analogy (for normal operators) between self-adjoint operators and real numbers.

10. 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 $T^2 = T$.

11. 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 $S^2 = T$.)

12. Give an example of a real inner-product space V and $T \in \mathcal{L}(V)$ and real numbers α, β with $\alpha^2 < 4\beta$ such that $T^2 + \alpha T + \beta I$ is not invertible.

This exercise shows that the hypothesis that T is self-adjoint is needed in 7.11, even for real vector spaces.

13. 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 $S^3 = T$.)

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

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

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

15. 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.

16. Give an example of an operator T on an inner product space such that T has an invariant subspace whose orthogonal complement is not invariant under T.

This exercise shows that 7.18 can fail without the hypothesis that T is normal.

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

18. Prove that if $T \in \mathcal{L}(V)$ is positive, then so is T^k for every positive integer k.

19. Suppose that T is a positive operator on V. Prove that T is invertible if and only if
$$\langle Tv, v\rangle > 0$$
for every $v \in V \setminus \{0\}$.

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

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

22. Prove that if $S \in \mathcal{L}(\mathbf{R}^3)$ is an isometry, then there exists a nonzero vector $x \in \mathbf{R}^3$ such that $S^2x = x$.

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

Exercise 24 shows that if we write T as the product of an isometry and a positive operator (as in the polar decomposition), then the positive operator must equal $\sqrt{T^*T}$.

24. 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 $R = \sqrt{T^*T}$.

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

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

27. 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.

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

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

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

31. 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 $T_1 = S_1 T_2 S_2$.

32. 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 $s_1, \dots, s_n$ are the singular values of T and $(e_1, \dots, e_n)$ and $(f_1, \dots, f_n)$ are orthonormal bases of V.

(a) Prove that

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

for every $v \in V$.

(b) Prove that if T is invertible, 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 $v \in V$.

33. 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. Prove that

$$\hat{s}\|v\| \le \|Tv\| \le s\|v\|$$

for every $v \in V$.

34. Suppose $T', T'' \in \mathcal{L}(V)$. Let s' denote the largest singular value of T', let s'' denote the largest singular value of T'', and let s denote the largest singular value of $T' + T''$. Prove that $s \le s' + s''$.