7D Isometries, Unitary Operators, and Matrix Factorization

Isometries

Linear maps that preserve norms are sufficiently important to deserve a name.

7.44 definition: *isometry*

A linear map $S \in \mathcal{L}(V, W)$ is called an *isometry* if

$$\|Sv\| = \|v\|$$

for every $v \in V$. In other words, a linear map is an isometry if it preserves norms.

If $S \in \mathcal{L}(V, W)$ is an isometry and $v \in V$ is such that $Sv = 0$, then

$$\|v\| = \|Sv\| = \|0\| = 0,$$

which implies that $v = 0$. Thus every isometry is injective.

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

7.45 example: *orthonormal basis maps to orthonormal list* $\Longrightarrow$ *isometry*

Suppose $e_1, \dots, e_n$ is an orthonormal basis of V and $g_1, \dots, g_n$ is an orthonormal list in W. Let $S \in \mathcal{L}(V, W)$ be the linear map such that $Se_k = g_k$ for each $k = 1, \dots, n$. To show that S is an isometry, suppose $v \in V$. Then

$$v = \langle v, e_1 \rangle e_1 + \cdots + \langle v, e_n \rangle e_n \tag{7.46}$$

and

$$\|v\|^2 = \left|\langle v, e_1 \rangle\right|^2 + \cdots + \left|\langle v, e_n \rangle\right|^2, \tag{7.47}$$

where we have used 6.30(b). Applying S to both sides of 7.46 gives

$$Sv = \langle v, e_1 \rangle Se_1 + \cdots + \langle v, e_n \rangle Se_n = \langle v, e_1 \rangle g_1 + \cdots + \langle v, e_n \rangle g_n.$$

Thus

$$\|Sv\|^2 = \left|\langle v, e_1 \rangle\right|^2 + \cdots + |\langle v, e_n \rangle|^2. \tag{7.48}$$

Comparing 7.47 and 7.48 shows that $\|v\| = \|Sv\|$. Thus S is an isometry.

The next result gives conditions equivalent to being an isometry. The equivalence of (a) and (c) shows that a linear map is an isometry if and only if it preserves inner products. The equivalence of (a) and (d) shows that a linear map is an isometry if and only if it maps some orthonormal basis to an orthonormal list. Thus the isometries given by Example 7.45 include all isometries. Furthermore, a linear map is an isometry if and only if it maps every orthonormal basis to an orthonormal list [because whether or not (a) holds does not depend on the basis $e_1, \dots, e_n$].

The equivalence of (a) and (e) in the next result shows that a linear map is an isometry if and only if the columns of its matrix (with respect to any orthonormal bases) form an orthonormal list. Here we are identifying the columns of an m-by-n matrix with elements of $\mathbf{F}^m$ and then using the Euclidean inner product on $\mathbf{F}^m$.

7.49 *characterizations of isometries*

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

(a) S is an isometry.

(b) $S^*S = I$.

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

(d) $Se_1, \dots, Se_n$ is an orthonormal list in W.

(e) The columns of $\mathcal{M}\big(S, (e_1, \dots, e_n), (f_1, \dots, f_m)\big)$ form an orthonormal list in $\mathbf{F}^m$ with respect to the Euclidean inner product.

Proof First suppose (a) holds, so S is an isometry. If $v \in V$ then

$$\big\langle (I - S^*S)v, v \big\rangle = \langle v, v \rangle - \big\langle S^*Sv, v \big\rangle = \|v\|^2 - \langle Sv, Sv \rangle = \|v\|^2 - \|Sv\|^2 = 0.$$

Hence the self-adjoint operator $I - S^*S$ equals 0 (by 7.16). Thus $S^*S = I$, proving that (a) implies (b).

Now suppose (b) holds, so $S^*S = I$. If $u, v \in V$ then

$$\langle Su, Sv \rangle = \big\langle S^*Su, v \big\rangle = \langle Iu, v \rangle = \langle u, v \rangle,$$

proving that (b) implies (c).

Now suppose that (c) holds, so $\langle Su, Sv \rangle = \langle u, v \rangle$ for all $u, v \in V$. Thus if $j, k \in \{1, \dots, n\}$, then

$$\langle Se_j, Se_k \rangle = \langle e_j, e_k \rangle.$$

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

Now suppose that (d) holds, so $Se_1, \dots, Se_n$ is an orthonormal list in W. Let $A = \mathcal{M}\big(S, (e_1, \dots, e_n), (f_1, \dots, f_m)\big)$. If $k, r \in \{1, \dots, n\}$, then

$$\text{7.50} \qquad \sum_{j=1}^{m} A_{j,k}\overline{A_{j,r}} = \left\langle \sum_{j=1}^{m} A_{j,k} f_j, \sum_{j=1}^{m} A_{j,r} f_j \right\rangle = \langle Se_k, Se_r \rangle = \begin{cases} 1 & \text{if } k = r, \\ 0 & \text{if } k \neq r. \end{cases}$$

The left side of 7.50 is the inner product in $\mathbf{F}^m$ of columns k and r of A. Thus the columns of A form an orthonormal list in $\mathbf{F}^m$, proving that (d) implies (e).

Now suppose (e) holds, so the columns of the matrix A defined in the paragraph above form an orthonormal list in $\mathbf{F}^m$. Then 7.50 shows that $Se_1, \dots, Se_n$ is an orthonormal list in W. Thus Example 7.45, with $Se_1, \dots, Se_n$ playing the role of $g_1, \dots, g_n$, shows that S is an isometry, proving that (e) implies (a). ■

See Exercises 1 and 11 for additional conditions that are equivalent to being an isometry.

Unitary Operators

In this subsection, we confine our attention to linear maps from a vector space to itself. In other words, we will be working with operators.

7.51 definition: *unitary operator*

An operator $S \in \mathcal{L}(V)$ is called *unitary* if S is an invertible isometry.

Although the words "unitary" and "isometry" mean the same thing for operators on finite-dimensional inner product spaces, remember that a unitary operator maps a vector space to itself, while an isometry maps a vector space to another (possibly different) vector space.

As previously noted, every isometry is injective. Every injective operator on a finite-dimensional vector space is invertible (see 3.65). A standing assumption for this chapter is that V is a finite-dimensional inner product space. Thus we could delete the word "invertible" from the definition above without changing the meaning. The unnecessary word "invertible" has been retained in the definition above for consistency with the definition readers may encounter when learning about inner product spaces that are not necessarily finite-dimensional.

7.52 example: *rotation of* $\mathbf{R}^2$

Suppose $\theta \in \mathbf{R}$ and S is the operator on $\mathbf{F}^2$ whose matrix with respect to the standard basis of $\mathbf{F}^2$ is

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

The two columns of this matrix form an orthonormal list in $\mathbf{F}^2$; hence S is an isometry [by the equivalence of (a) and (e) in 7.49]. Thus S is a unitary operator.

If $\mathbf{F} = \mathbf{R}$, then S is the operator of counterclockwise rotation by θ radians around the origin of $\mathbf{R}^2$. This observation gives us another way to think about why S is an isometry, because each rotation around the origin of $\mathbf{R}^2$ preserves norms.

The next result (7.53) lists several conditions that are equivalent to being a unitary operator. All the conditions equivalent to being an isometry in 7.49 should be added to this list. The extra conditions in 7.53 arise because of limiting the context to linear maps from a vector space to itself. For example, 7.49 shows that a linear map $S \in \mathcal{L}(V, W)$ is an isometry if and only if $S^*S = I$, while 7.53 shows that an operator $S \in \mathcal{L}(V)$ is a unitary operator if and only if $S^*S = SS^* = I$.

Another difference is that 7.49(d) mentions an orthonormal list, while 7.53(d) mentions an orthonormal basis. Also, 7.49(e) mentions the columns of $\mathcal{M}(T)$, while 7.53(e) mentions the rows of $\mathcal{M}(T)$. Furthermore, $\mathcal{M}(T)$ in 7.49(e) is with respect to an orthonormal basis of V and an orthonormal basis of W, while $\mathcal{M}(T)$ in 7.53(e) is with respect to a single basis of V doing double duty.

7.53 *characterizations of unitary operators*

Suppose $S \in \mathcal{L}(V)$. Suppose $e_1, \ldots, e_n$ is an orthonormal basis of V. Then the following are equivalent.

(a) S is a unitary operator.

(b) $S^*S = SS^* = I$.

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

(d) $Se_1, \ldots, Se_n$ is an orthonormal basis of V.

(e) The rows of $\mathcal{M}\bigl(S, (e_1, \ldots, e_n)\bigr)$ form an orthonormal basis of $\mathbf{F}^n$ with respect to the Euclidean inner product.

(f) S^* is a unitary operator.

Proof First suppose (a) holds, so S is a unitary operator. Hence

$$S^*S = I$$

by the equivalence of (a) and (b) in 7.49. Multiply both sides of this equation by S^{-1} on the right, getting $S^* = S^{-1}$. Thus $SS^* = SS^{-1} = I$, as desired, proving that (a) implies (b).

The definitions of invertible and inverse show that (b) implies (c).

Now suppose (c) holds, so S is invertible and $S^{-1} = S^*$. Thus $S^*S = I$. Hence $Se_1, \ldots, Se_n$ is an orthonormal list in V, by the equivalence of (b) and (d) in 7.49. The length of this list equals $\dim V$. Thus $Se_1, \ldots, Se_n$ is an orthonormal basis of V, proving that (c) implies (d).

Now suppose (d) holds, so $Se_1, \ldots, Se_n$ is an orthonormal basis of V. The equivalence of (a) and (d) in 7.49 shows that S is a unitary operator. Thus

$$\bigl(S^*\bigr)^* S^* = SS^* = I,$$

where the last equation holds because we already showed that (a) implies (b) in this result. The equation above and the equivalence of (a) and (b) in 7.49 show that S^* is an isometry. Thus the columns of $\mathcal{M}\bigl(S^*, (e_1, \ldots, e_n)\bigr)$ form an orthonormal basis of $\mathbf{F}^n$ [by the equivalence of (a) and (e) of 7.49]. The rows of $\mathcal{M}\bigl(S, (e_1, \ldots, e_n)\bigr)$ are the complex conjugates of the columns of $\mathcal{M}\bigl(S^*, (e_1, \ldots, e_n)\bigr)$. Thus the rows of $\mathcal{M}\bigl(S, (e_1, \ldots, e_n)\bigr)$ form an orthonormal basis of $\mathbf{F}^n$, proving that (d) implies (e).

Now suppose (e) holds. Thus the columns of $\mathcal{M}\bigl(S^*, (e_1, \ldots, e_n)\bigr)$ form an orthonormal basis of $\mathbf{F}^n$. The equivalence of (a) and (e) in 7.49 shows that S^* is an isometry, proving that (e) implies (f).

Now suppose (f) holds, so S^* is a unitary operator. The chain of implications we have already proved in this result shows that (a) implies (f). Applying this result to S^* shows that $\bigl(S^*\bigr)^*$ is a unitary operator, proving that (f) implies (a).

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

Recall our analogy between **C** and $\mathcal{L}(V)$. Under this analogy, a complex number z corresponds to an operator $S \in \mathcal{L}(V)$, and $\bar{z}$ corresponds to S^*. The real numbers $(z = \bar{z})$ correspond to the self-adjoint operators $(S = S^*)$, 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 corresponds to the condition $S^*S = I$, which is equivalent to S being a unitary operator. Hence the analogy shows that the unit circle in **C** corresponds to the set of unitary operators. In the next two results, this analogy appears in the eigenvalues of unitary operators. Also see Exercise 15 for another example of this analogy.

7.54 *eigenvalues of unitary operators have absolute value* 1

Suppose λ is an eigenvalue of a unitary operator. Then $|\lambda| = 1$.

Proof Suppose $S \in \mathcal{L}(V)$ is a unitary operator and λ is an eigenvalue of S. Let $v \in V$ be such that $v \neq 0$ and $Sv = \lambda v$. Then

$$|\lambda|\,\|v\| = \|\lambda v\| = \|Sv\| = \|v\|.$$

Thus $|\lambda| = 1$, as desired. ■

The next result characterizes unitary operators on finite-dimensional complex inner product spaces, using the complex spectral theorem as the main tool.

7.55 *description of unitary operators on complex inner product spaces*

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

(a) S is a unitary operator.

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

Proof Suppose (a) holds, so S is a unitary operator. The equivalence of (a) and (b) in 7.53 shows that S is normal. Thus the complex spectral theorem (7.31) shows that there is an orthonormal basis $e_1, \ldots, e_n$ of V consisting of eigenvectors of S. Every eigenvalue of S has absolute value 1 (by 7.54), completing the proof that (a) implies (b).

Now suppose (b) holds. Let $e_1, \ldots, e_n$ be an orthonormal basis of V consisting of eigenvectors of S whose corresponding eigenvalues $\lambda_1, \ldots, \lambda_n$ all have absolute value 1. Then $Se_1, \ldots, Se_n$ is also an orthonormal basis of V because

$$\langle Se_j, Se_k\rangle = \langle \lambda_j e_j, \lambda_k e_k\rangle = \lambda_j \overline{\lambda_k}\langle e_j, e_k\rangle = \begin{cases} 0 & \text{if } j \neq k, \\ 1 & \text{if } j = k \end{cases}$$

for all $j, k = 1, \ldots, n$. Thus the equivalence of (a) and (d) in 7.53 shows that S is unitary, proving that (b) implies (a). ■

QR Factorization

In this subsection, we shift our attention from operators to matrices. This switch should give you good practice in identifying an operator with a square matrix (after picking a basis of the vector space on which the operator is defined). You should also become more comfortable with translating concepts and results back and forth between the context of operators and the context of square matrices.

When starting with n-by-n matrices instead of operators, unless otherwise specified assume that the associated operators live on $\mathbf{F}^n$ (with the Euclidean inner product) and that their matrices are computed with respect to the standard basis of $\mathbf{F}^n$.

We begin by making the following definition, transferring the notion of a unitary operator to a unitary matrix.

7.56 definition: ***unitary matrix***

An n-by-n matrix is called *unitary* if its columns form an orthonormal list in $\mathbf{F}^n$.

In the definition above, we could have replaced "orthonormal list in $\mathbf{F}^n$" with "orthonormal basis of $\mathbf{F}^n$" because every orthonormal list of length n in an n-dimensional inner product space is an orthonormal basis. If $S \in \mathcal{L}(V)$ and $e_1, \dots, e_n$ and $f_1, \dots, f_n$ are orthonormal bases of V, then S is a unitary operator if and only if $\mathcal{M}\big(S, (e_1, \dots, e_n), (f_1, \dots, f_n)\big)$ is a unitary matrix, as shown by the equivalence of (a) and (e) in 7.49. Also note that we could also have replaced "columns" in the definition above with "rows" by using the equivalence between conditions (a) and (e) in 7.53.

The next result, whose proof will be left as an exercise for the reader, gives some equivalent conditions for a square matrix to be unitary. In (c), Qv denotes the matrix product of Q and v, identifying elements of $\mathbf{F}^n$ with n-by-1 matrices (sometimes called column vectors). The norm in (c) below is the usual Euclidean norm on $\mathbf{F}^n$ that comes from the Euclidean inner product. In (d), Q^* denotes the conjugate transpose of the matrix Q, which corresponds to the adjoint of the associated operator.

7.57 ***characterizations of unitary matrices***

Suppose Q is an n-by-n matrix. Then the following are equivalent.

(a) Q is a unitary matrix.

(b) The rows of Q form an orthonormal list in $\mathbf{F}^n$.

(c) $\|Qv\| = \|v\|$ for every $v \in \mathbf{F}^n$.

(d) $Q^*Q = QQ^* = I$, the n-by-n matrix with 1's on the diagonal and 0's elsewhere.

The QR factorization stated and proved below is the main tool in the widely used QR algorithm (not discussed here) for finding good approximations to eigenvalues and eigenvectors of square matrices. In the result below, if the matrix A is in $\mathbf{F}^{n,n}$, then the matrices Q and R are also in $\mathbf{F}^{n,n}$.

7.58 *QR factorization*

Suppose A is a square matrix with linearly independent columns. Then there exist unique matrices Q and R such that Q is unitary, R is upper triangular with only positive numbers on its diagonal, and

$$A = QR.$$

Proof Let $v_1, \ldots, v_n$ denote the columns of A, thought of as elements of $\mathbf{F}^n$. Apply the Gram–Schmidt procedure (6.32) to the list $v_1, \ldots, v_n$, getting an orthonormal basis $e_1, \ldots, e_n$ of $\mathbf{F}^n$ such that

7.59 $$\operatorname{span}(v_1, \ldots, v_k) = \operatorname{span}(e_1, \ldots, e_k)$$

for each $k = 1, \ldots, n$. Let R be the n-by-n matrix defined by

$$R_{j,k} = \langle v_k, e_j \rangle,$$

where $R_{j,k}$ denotes the entry in row j, column k of R. If $j > k$, then e_j is orthogonal to $\operatorname{span}(e_1, \ldots, e_k)$ and hence e_j is orthogonal to v_k (by 7.59). In other words, if $j > k$ then $\langle v_k, e_j \rangle = 0$. Thus R is an upper-triangular matrix.

Let Q be the unitary matrix whose columns are $e_1, \ldots, e_n$. If $k \in \{1, \ldots, n\}$, then the k^{th} column of QR equals a linear combination of the columns of Q, with the coefficients for the linear combination coming from the k^{th} column of R—see 3.51(a). Hence the k^{th} column of QR equals

$$\langle v_k, e_1 \rangle e_1 + \cdots + \langle v_k, e_k \rangle e_k,$$

which equals v_k [by 6.30(a)], the k^{th} column of A. Thus $A = QR$, as desired.

The equations defining the Gram–Schmidt procedure (see 6.32) show that each v_k equals a positive multiple of e_k plus a linear combination of $e_1, \ldots, e_{k-1}$. Thus each $\langle v_k, e_k \rangle$ is a positive number. Hence all entries on the diagonal of R are positive numbers, as desired.

Finally, to show that Q and R are unique, suppose we also have $A = \widehat{Q}\widehat{R}$, where $\widehat{Q}$ is unitary and $\widehat{R}$ is upper triangular with only positive numbers on its diagonal. Let $q_1, \ldots, q_n$ denote the columns of $\widehat{Q}$. Thinking of matrix multiplication as above, we see that each v_k is a linear combination of $q_1, \ldots, q_k$, with the coefficients coming from the k^{th} column of $\widehat{R}$. This implies that $\operatorname{span}(v_1, \ldots, v_k) = \operatorname{span}(q_1, \ldots, q_k)$ and $\langle v_k, q_k \rangle > 0$. The uniqueness of the orthonormal lists satisfying these conditions (see Exercise 10 in Section 6B) now shows that $q_k = e_k$ for each $k = 1, \ldots, n$. Hence $\widehat{Q} = Q$, which then implies that $\widehat{R} = R$, completing the proof of uniqueness. ■

The proof of the QR factorization shows that the columns of the unitary matrix can be computed by applying the Gram–Schmidt procedure to the columns of the matrix to be factored. The next example illustrates the computation of the QR factorization based on the proof that we just completed.

7.60 example: *QR factorization of a* 3*-by-*3 *matrix*

To find the QR factorization of the matrix

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & -4 \\ 0 & 3 & 2 \end{pmatrix},$$

follow the proof of 7.58. Thus set v_1, v_2, v_3 equal to the columns of A:

$$v_1 = (1,0,0), \quad v_2 = (2,1,3), \quad v_3 = (1,-4,2).$$

Apply the Gram–Schmidt procedure to v_1, v_2, v_3, producing the orthonormal list

$$e_1 = (1,0,0), \quad e_2 = \Bigl(0, \tfrac{1}{\sqrt{10}}, \tfrac{3}{\sqrt{10}}\Bigr), \quad e_3 = \Bigl(0, -\tfrac{3}{\sqrt{10}}, \tfrac{1}{\sqrt{10}}\Bigr).$$

Still following the proof of 7.58, let Q be the unitary matrix whose columns are e_1, e_2, e_3:

$$Q = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{10}} & -\frac{3}{\sqrt{10}} \\ 0 & \frac{3}{\sqrt{10}} & \frac{1}{\sqrt{10}} \end{pmatrix}.$$

As in the proof of 7.58, let R be the 3-by-3 matrix whose entry in row j, column k is $\langle v_k, e_j \rangle$, which gives

$$R = \begin{pmatrix} 1 & 2 & 1 \\ 0 & \sqrt{10} & \frac{\sqrt{10}}{5} \\ 0 & 0 & \frac{7\sqrt{10}}{5} \end{pmatrix}.$$

Note that R is indeed an upper-triangular matrix with only positive numbers on the diagonal, as required by the QR factorization.

Now matrix multiplication can verify that $A = QR$ is the desired factorization of A:

$$QR = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{10}} & -\frac{3}{\sqrt{10}} \\ 0 & \frac{3}{\sqrt{10}} & \frac{1}{\sqrt{10}} \end{pmatrix} \begin{pmatrix} 1 & 2 & 1 \\ 0 & \sqrt{10} & \frac{\sqrt{10}}{5} \\ 0 & 0 & \frac{7\sqrt{10}}{5} \end{pmatrix} = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & -4 \\ 0 & 3 & 2 \end{pmatrix} = A.$$

Thus $A = QR$, as expected.

The QR factorization will be the major tool used in the proof of the Cholesky factorization (7.63) in the next subsection. For another nice application of the QR factorization, see the proof of Hadamard's inequality (9.66).

If a QR factorization is available, then it can be used to solve a corresponding system of linear equations without using Gaussian elimination. Specifically, suppose A is an n-by-n square matrix with linearly independent columns. Suppose that $b \in \mathbf{F}^n$ and we want to solve the equation $Ax = b$ for $x = (x_1, \dots, x_n) \in \mathbf{F}^n$ (as usual, we are identifying elements of $\mathbf{F}^n$ with n-by-1 column vectors).

Suppose $A = QR$, where Q is unitary and R is upper triangular with only positive numbers on its diagonal (Q and R are computable from A using just the Gram–Schmidt procedure, as shown in the proof of 7.58). The equation $Ax = b$ is equivalent to the equation $QRx = b$. Multiplying both sides of this last equation by Q^* on the left and using 7.57(d) gives the equation

$$Rx = Q^*b.$$

The matrix Q^* is the conjugate transpose of the matrix Q. Thus computing Q^*b is straightforward. Because R is an upper-triangular matrix with positive numbers on its diagonal, the system of linear equations represented by the equation above can quickly be solved by first solving for x_n, then for x_{n-1}, and so on.

Cholesky Factorization

We begin this subsection with a characterization of positive invertible operators in terms of inner products.

7.61 *positive invertible operator*

A self-adjoint operator $T \in \mathcal{L}(V)$ is a positive invertible operator if and only if $\langle Tv, v \rangle > 0$ for every nonzero $v \in V$.

Proof First suppose T is a positive invertible operator. If $v \in V$ and $v \neq 0$, then because T is invertible we have $Tv \neq 0$. This implies that $\langle Tv, v \rangle \neq 0$ (by 7.43). Hence $\langle Tv, v \rangle > 0$.

To prove the implication in the other direction, suppose now that $\langle Tv, v \rangle > 0$ for every nonzero $v \in V$. Thus $Tv \neq 0$ for every nonzero $v \in V$. Hence T is injective. Thus T is invertible, as desired. ■

The next definition transfers the result above to the language of matrices. Here we are using the usual Euclidean inner product on $\mathbf{F}^n$ and identifying elements of $\mathbf{F}^n$ with n-by-1 column vectors.

7.62 definition: *positive definite*

A matrix $B \in \mathbf{F}^{n,n}$ is called *positive definite* if $B^* = B$ and

$$\langle Bx, x \rangle > 0$$

for every nonzero $x \in \mathbf{F}^n$.

A matrix is upper triangular if and only if its conjugate transpose is lower triangular (meaning that all entries above the diagonal are 0). The factorization below, which has important consequences in computational linear algebra, writes a positive definite matrix as the product of a lower triangular matrix and its conjugate transpose.

Our next result is solely about matrices, although the proof makes use of the identification of results about operators with results about square matrices. In the result below, if the matrix B is in $\mathbf{F}^{n,n}$, then the matrix R is also in $\mathbf{F}^{n,n}$.

7.63 *Cholesky factorization*

Suppose B is a positive definite matrix. Then there exists a unique upper-triangular matrix R with only positive numbers on its diagonal such that

$$B = R^*R.$$

Proof Because B is positive definite, there exists an invertible square matrix A of the same size as B such that $B = A^*A$ [by the equivalence of (a) and (f) in 7.38].

Let $A = QR$ be the QR factorization of A (see 7.58), where Q is unitary and R is upper triangular with only positive numbers on its diagonal. Then $A^* = R^*Q^*$. Thus

$$B = A^*A = R^*Q^*QR = R^*R,$$

as desired.

André-Louis Cholesky (1875–1918) discovered this factorization, which was published posthumously in 1924.

To prove the uniqueness part of this result, suppose S is an upper-triangular matrix with only positive numbers on its diagonal and $B = S^*S$. The matrix S is invertible because B is invertible (see Exercise 11 in Section 3D). Multiplying both sides of the equation $B = S^*S$ by S^{-1} on the right gives the equation $BS^{-1} = S^*$.

Let A be the matrix from the first paragraph of this proof. Then

$$\begin{aligned}
\left(AS^{-1}\right)^*\left(AS^{-1}\right) &= (S^*)^{-1}A^*AS^{-1} \\
&= (S^*)^{-1}BS^{-1} \\
&= (S^*)^{-1}S^* \\
&= I.
\end{aligned}$$

Thus AS^{-1} is unitary.

Hence $A = \left(AS^{-1}\right)S$ is a factorization of A as the product of a unitary matrix and an upper-triangular matrix with only positive numbers on its diagonal. The uniqueness of the QR factorization, as stated in 7.58, now implies that $S = R$. ■

In the first paragraph of the proof above, we could have chosen A to be the unique positive definite matrix that is a square root of B (see 7.39). However, the proof was presented with the more general choice of A because for specific positive definite matrices B, it may be easier to find a different choice of A.

Exercises 7D

1 Suppose $\dim V \geq 2$ and $S \in \mathcal{L}(V, W)$. Prove that S is an isometry if and only if Se_1, Se_2 is an orthonormal list in W for every orthonormal list e_1, e_2 of length two in V.

2 Suppose $T \in \mathcal{L}(V, W)$. Prove that T is a scalar multiple of an isometry if and only if T preserves orthogonality.

The phrase "T preserves orthogonality" means that $\langle Tu, Tv\rangle = 0$ for all $u, v \in V$ such that $\langle u, v\rangle = 0$.

3 (a) Show that the product of two unitary operators on V is a unitary operator.
(b) Show that the inverse of a unitary operator on V is a unitary operator.

This exercise shows that the set of unitary operators on V is a group, where the group operation is the usual product of two operators.

4 Suppose $\mathbf{F} = \mathbf{C}$ and $A, B \in \mathcal{L}(V)$ are self-adjoint. Show that $A + iB$ is unitary if and only if $AB = BA$ and $A^2 + B^2 = I$.

5 Suppose $S \in \mathcal{L}(V)$. Prove that the following are equivalent.
(a) S is a self-adjoint unitary operator.
(b) $S = 2P - I$ for some orthogonal projection P on V.
(c) There exists a subspace U of V such that $Su = u$ for every $u \in U$ and $Sw = -w$ for every $w \in U^{\perp}$.

6 Suppose T_1, T_2 are both normal operators on $\mathbf{F}^3$ with $2, 5, 7$ as eigenvalues. Prove that there exists a unitary operator $S \in \mathcal{L}(\mathbf{F}^3)$ such that $T_1 = S^* T_2 S$.

7 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 a unitary operator $S \in \mathcal{L}(\mathbf{F}^4)$ such that $T_1 = S^* T_2 S$. Be sure to explain why there is no unitary operator with the required property.

8 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 $\|Se_k\| = 1$ for each e_k, then S is a unitary operator.

9 Suppose $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$. Suppose every eigenvalue of T has absolute value 1 and $\|Tv\| \leq \|v\|$ for every $v \in V$. Prove that T is a unitary operator.

10 Suppose $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$ is a self-adjoint operator such that $\|Tv\| \leq \|v\|$ for all $v \in V$.
(a) Show that $I - T^2$ is a positive operator.
(b) Show that $T + i\sqrt{I - T^2}$ is a unitary operator.

11 Suppose $S \in \mathcal{L}(V)$. Prove that S is a unitary operator if and only if

$$\{Sv : v \in V \text{ and } \|v\| \leq 1\} = \{v \in V : \|v\| \leq 1\}.$$

12 Prove or give a counterexample: If $S \in \mathcal{L}(V)$ is invertible and $\|S^{-1}v\| = \|Sv\|$ for every $v \in V$, then S is unitary.

13 Explain why the columns of a square matrix of complex numbers form an orthonormal list in $\mathbf{C}^n$ if and only if the rows of the matrix form an orthonormal list in $\mathbf{C}^n$.

14 Suppose $v \in V$ with $\|v\| = 1$ and $b \in \mathbf{F}$. Also suppose $\dim V \geq 2$. Prove that there exists a unitary operator $S \in \mathcal{L}(V)$ such that $\langle Sv, v \rangle = b$ if and only if $|b| \leq 1$.

15 Suppose T is a unitary operator on V such that $T - I$ is invertible.

(a) Prove that $(T + I)(T - I)^{-1}$ is a skew operator (meaning that it equals the negative of its adjoint).

(b) Prove that if $\mathbf{F} = \mathbf{C}$, then $i(T + I)(T - I)^{-1}$ is a self-adjoint operator.

The function $z \mapsto i(z+1)(z-1)^{-1}$ maps the unit circle in $\mathbf{C}$ (except for the point 1) to $\mathbf{R}$. Thus (b) illustrates the analogy between the unitary operators and the unit circle in $\mathbf{C}$, along with the analogy between the self-adjoint operators and $\mathbf{R}$.

16 Suppose $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$ is self-adjoint. Prove that $(T + iI)(T - iI)^{-1}$ is a unitary operator and 1 is not an eigenvalue of this operator.

17 Explain why the characterizations of unitary matrices given by 7.57 hold.

18 A square matrix A is called *symmetric* if it equals its transpose. Prove that if A is a symmetric matrix with real entries, then there exists a unitary matrix Q with real entries such that Q^*AQ is a diagonal matrix.

19 Suppose n is a positive integer. For this exercise, we adopt the notation that a typical element z of $\mathbf{C}^n$ is denoted by $z = (z_0, z_1, \dots, z_{n-1})$. Define linear functionals $\omega_0, \omega_1, \dots, \omega_{n-1}$ on $\mathbf{C}^n$ by

$$\omega_j(z_0, z_1, \dots, z_{n-1}) = \frac{1}{\sqrt{n}} \sum_{m=0}^{n-1} z_m e^{-2\pi i j m/n}.$$

The *discrete Fourier transform* is the operator $\mathcal{F}\colon \mathbf{C}^n \to \mathbf{C}^n$ defined by

$$\mathcal{F}z = \bigl(\omega_0(z), \omega_1(z), \dots, \omega_{n-1}(z)\bigr).$$

(a) Show that $\mathcal{F}$ is a unitary operator on $\mathbf{C}^n$.

(b) Show that if $(z_0, \dots, z_{n-1}) \in \mathbf{C}^n$ and z_n is defined to equal z_0, then

$$\mathcal{F}^{-1}(z_0, z_1, \dots, z_{n-1}) = \mathcal{F}(z_n, z_{n-1}, \dots, z_1).$$

(c) Show that $\mathcal{F}^4 = I$.

The discrete Fourier transform has many important applications in data analysis. The usual Fourier transform involves expressions of the form $\int_{-\infty}^{\infty} f(x) e^{-2\pi i t x} dx$ for complex-valued integrable functions f defined on $\mathbf{R}$.

20 Suppose A is a square matrix with linearly independent columns. Prove that there exist unique matrices R and Q such that R is lower triangular with only positive numbers on its diagonal, Q is unitary, and $A = RQ$.