8. Inner Product Spaces

8.1. Inner Products

Throughout this chapter we consider only real or complex vector spaces, that is, vector spaces over the field of real numbers or the field of complex numbers. Our main object is to study vector spaces in which it makes sense to speak of the 'length' of a vector and of the 'angle' between two vectors. We shall do this by studying a certain type of scalar-valued function on pairs of vectors, known as an inner product. One example of an inner product is the scalar or dot product of vectors in R^3. The scalar product of the vectors

$$\alpha = (x_1, x_2, x_3) \quad \text{and} \quad \beta = (y_1, y_2, y_3)$$

in R^3 is the real number

$$(\alpha|\beta) = x_1y_1 + x_2y_2 + x_3y_3.$$

Geometrically, this dot product is the product of the length of α, the length of β, and the cosine of the angle between α and β. It is therefore possible to define the geometric concepts of 'length' and 'angle' in R^3 by means of the algebraically defined scalar product.

An inner product on a vector space is a function with properties similar to the dot product in R^3, and in terms of such an inner product one can also define 'length' and 'angle.' Our comments about the general notion of angle will be restricted to the concept of perpendicularity (or orthogonality) of vectors. In this first section we shall say what an inner product is, consider some particular examples, and establish a few basic

properties of inner products. Then we turn to the task of discussing length and orthogonality.

Definition. *Let F be the field of real numbers or the field of complex numbers, and V a vector space over F. An* **inner product** *on V is a function which assigns to each ordered pair of vectors α, β in V a scalar $(\alpha|\beta)$ in F in such a way that for all α, β, γ in V and all scalars c*

(a) $(\alpha + \beta|\gamma) = (\alpha|\gamma) + (\beta|\gamma)$;
(b) $(c\alpha|\beta) = c(\alpha|\beta)$;
(c) $(\beta|\alpha) = \overline{(\alpha|\beta)}$, *the bar denoting complex conjugation;*
(d) $(\alpha|\alpha) > 0$ *if* $\alpha \neq 0$.

It should be observed that conditions (a), (b), and (c) imply that

(e) $$(\alpha|c\beta + \gamma) = \bar{c}(\alpha|\beta) + (\alpha|\gamma).$$

One other point should be made. When F is the field R of real numbers, the complex conjugates appearing in (c) and (e) are superfluous; however, in the complex case they are necessary for the consistency of the conditions. Without these complex conjugates, we would have the contradiction:

$$(\alpha|\alpha) > 0 \quad \text{and} \quad (i\alpha|i\alpha) = -1(\alpha|\alpha) > 0.$$

In the examples that follow and throughout the chapter, F is either the field of real numbers or the field of complex numbers.

Example 1. On F^n there is an inner product which we call the **standard inner product.** It is defined on $\alpha = (x_1, \ldots, x_n)$ and $\beta = (y_1, \ldots, y_n)$ by

(8-1) $$(\alpha|\beta) = \sum_j x_j\bar{y}_j.$$

When $F = R$, this may also be written

$$(\alpha|\beta) = \sum_j x_jy_j.$$

In the real case, the standard inner product is often called the dot or scalar product and denoted by $\alpha \cdot \beta$.

Example 2. For $\alpha = (x_1, x_2)$ and $\beta = (y_1, y_2)$ in R^2, let

$$(\alpha|\beta) = x_1y_1 - x_2y_1 - x_1y_2 + 4x_2y_2.$$

Since $(\alpha|\alpha) = (x_1 - x_2)^2 + 3x_2^2$, it follows that $(\alpha|\alpha) > 0$ if $\alpha \neq 0$. Conditions (a), (b), and (c) of the definition are easily verified.

Example 3. Let V be $F^{n\times n}$, the space of all $n \times n$ matrices over F. Then V is isomorphic to F^{n^2} in a natural way. It therefore follows from Example 1 that the equation

$$(A|B) = \sum_{j,k} A_{jk}\bar{B}_{jk}$$

defines an inner product on V. Furthermore, if we introduce the **conjugate transpose** matrix B^*, where $B^*_{kj} = \overline{B}_{jk}$, we may express this inner product on $F^{n\times n}$ in terms of the trace function:

$$(A|B) = \operatorname{tr}(AB^*) = \operatorname{tr}(B^*A).$$

For

$$\begin{aligned}\operatorname{tr}(AB^*) &= \sum_j (AB^*)_{jj}\\ &= \sum_j \sum_k A_{jk}B^*_{kj}\\ &= \sum_j \sum_k A_{jk}\overline{B}_{jk}.\end{aligned}$$

EXAMPLE 4. Let $F^{n\times 1}$ be the space of $n \times 1$ (column) matrices over F, and let Q be an $n \times n$ invertible matrix over F. For X, Y in $F^{n\times 1}$ set

$$(X|Y) = Y^*Q^*QX.$$

We are identifying the 1×1 matrix on the right with its single entry. When Q is the identity matrix, this inner product is essentially the same as that in Example 1; we call it the **standard inner product** on $F^{n\times 1}$. The reader should note that the terminology 'standard inner product' is used in two special contexts. For a general finite-dimensional vector space over F, there is no obvious inner product that one may call standard.

EXAMPLE 5. Let V be the vector space of all continuous complex-valued functions on the unit interval, $0 \leq t \leq 1$. Let

$$(f|g) = \int_0^1 f(t)\overline{g(t)}\,dt.$$

The reader is probably more familiar with the space of real-valued continuous functions on the unit interval, and for this space the complex conjugate on g may be omitted.

EXAMPLE 6. This is really a whole class of examples. One may construct new inner products from a given one by the following method. Let V and W be vector spaces over F and suppose $(\ |\ )$ is an inner product on W. If T is a non-singular linear transformation from V into W, then the equation

$$p_T(\alpha, \beta) = (T\alpha|T\beta)$$

defines an inner product p_T on V. The inner product in Example 4 is a special case of this situation. The following are also special cases.

(a) Let V be a finite-dimensional vector space, and let

$$\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$$

be an ordered basis for V. Let $\epsilon_1, \ldots, \epsilon_n$ be the standard basis vectors in F^n, and let T be the linear transformation from V into F^n such that $T\alpha_j = \epsilon_j$, $j = 1, \ldots, n$. In other words, let T be the 'natural' isomorphism of V onto F^n that is determined by $\mathcal{B}$. If we take the standard inner product on F^n, then

$$p_T(\sum_j x_j\alpha_j, \sum_k y_k\alpha_k) = \sum_{j=1}^n x_j\bar{y}_j.$$

Thus, for any basis for V there is an inner product on V with the property $(\alpha_j|\alpha_k) = \delta_{jk}$; in fact, it is easy to show that there is exactly one such inner product. Later we shall show that every inner product on V is determined by some basis $\mathcal{B}$ in the above manner.

(b) We look again at Example 5 and take $V = W$, the space of continuous functions on the unit interval. Let T be the linear operator 'multiplication by t,' that is, $(Tf)(t) = tf(t)$, $0 \leq t \leq 1$. It is easy to see that T is linear. Also T is non-singular; for suppose $Tf = 0$. Then $tf(t) = 0$ for $0 \leq t \leq 1$; hence $f(t) = 0$ for $t > 0$. Since f is continuous, we have $f(0) = 0$ as well, or $f = 0$. Now using the inner product of Example 5, we construct a new inner product on V by setting

$$\begin{aligned}p_T(f, g) &= \int_0^1 (Tf)(t)\overline{(Tg)(t)}\,dt\\ &= \int_0^1 f(t)\overline{g(t)}t^2\,dt.\end{aligned}$$

We turn now to some general observations about inner products. Suppose V is a complex vector space with an inner product. Then for all α, β in V

$$(\alpha|\beta) = \operatorname{Re}(\alpha|\beta) + i\operatorname{Im}(\alpha|\beta)$$

where $\operatorname{Re}(\alpha|\beta)$ and $\operatorname{Im}(\alpha|\beta)$ are the real and imaginary parts of the complex number $(\alpha|\beta)$. If z is a complex number, then $\operatorname{Im}(z) = \operatorname{Re}(-iz)$. It follows that

$$\operatorname{Im}(\alpha|\beta) = \operatorname{Re}[-i(\alpha|\beta)] = \operatorname{Re}(\alpha|i\beta).$$

Thus the inner product is completely determined by its 'real part' in accordance with

$$(\alpha|\beta) = \operatorname{Re}(\alpha|\beta) + i\operatorname{Re}(\alpha|i\beta). \tag{8-2}$$

Occasionally it is very useful to know that an inner product on a real or complex vector space is determined by another function, the so-called quadratic form determined by the inner product. To define it, we first denote the positive square root of $(\alpha|\alpha)$ by $\|\alpha\|$; $\|\alpha\|$ is called the **norm** of α with respect to the inner product. By looking at the standard inner products in R^1, C^1, R^2, and R^3, the reader should be able to convince himself that it is appropriate to think of the norm of α as the 'length' or 'magnitude' of α. The **quadratic form** determined by the inner product is the function that assigns to each vector α the scalar $||\alpha||^2$. It follows from the properties of the inner product that

$$||\alpha \pm \beta||^2 = ||\alpha||^2 \pm 2 \operatorname{Re} (\alpha|\beta) + ||\beta||^2$$

for all vectors α and β. Thus in the real case

$$(\alpha|\beta) = \frac{1}{4}||\alpha + \beta||^2 - \frac{1}{4}||\alpha - \beta||^2. \tag{8-3}$$

In the complex case we use (8-2) to obtain the more complicated expression

$$(\alpha|\beta) = \frac{1}{4}||\alpha + \beta||^2 - \frac{1}{4}||\alpha - \beta||^2 + \frac{i}{4}||\alpha + i\beta||^2 - \frac{i}{4}||\alpha - i\beta||^2. \tag{8-4}$$

Equations (8-3) and (8-4) are called the **polarization identities.** Note that (8-4) may also be written as follows:

$$(\alpha|\beta) = \frac{1}{4}\sum_{n=1}^{4} i^n ||\alpha + i^n\beta||^2.$$

The properties obtained above hold for any inner product on a real or complex vector space V, regardless of its dimension. We turn now to the case in which V is finite-dimensional. As one might guess, an inner product on a finite-dimensional space may always be described in terms of an ordered basis by means of a matrix.

Suppose that V is finite-dimensional, that

$$\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$$

is an ordered basis for V, and that we are given a particular inner product on V; we shall show that the inner product is completely determined by the values

$$G_{jk} = (\alpha_k|\alpha_j) \tag{8-5}$$

it assumes on pairs of vectors in $\mathcal{B}$. If $\alpha = \sum_k x_k\alpha_k$ and $\beta = \sum_j y_j\alpha_j$, then

$$\begin{aligned}(\alpha|\beta) &= (\sum_k x_n\alpha_k|\beta)\\ &= \sum_k x_k(\alpha_k|\beta)\\ &= \sum_k x_k \sum_j \bar{y}_j(\alpha_k|\alpha_j)\\ &= \sum_{j,k} \bar{y}_j G_{jk} x_k\\ &= Y^*GX\end{aligned}$$

where X, Y are the coordinate matrices of α, β in the ordered basis $\mathcal{B}$, and G is the matrix with entries $G_{jk} = (\alpha_k|\alpha_j)$. We call G the **matrix of the inner product in the ordered basis** $\mathcal{B}$. It follows from (8-5) that G is hermitian, i.e., that $G = G^*$; however, G is a rather special kind of hermitian matrix. For G must satisfy the additional condition

$$X^*GX > 0, \qquad X \neq 0. \tag{8-6}$$

In particular, G must be invertible. For otherwise there exists an $X \neq 0$ such that $GX = 0$, and for any such X, (8-6) is impossible. More explicitly, (8-6) says that for any scalars $x_1, \ldots, x_n$ not all of which are 0

$$\sum_{j,k} x_j G_{jk} x_k > 0. \tag{8-7}$$

From this we see immediately that each diagonal entry of G must be positive; however, this condition on the diagonal entries is by no means sufficient to insure the validity of (8-6). Sufficient conditions for the validity of (8-6) will be given later.

The above process is reversible; that is, if G is any $n \times n$ matrix over F which satisfies (8-6) and the condition $G = G^*$, then G is the matrix in the ordered basis $\mathcal{B}$ of an inner product on V. This inner product is given by the equation

$$(\alpha|\beta) = Y^*GX$$

where X and Y are the coordinate matrices of α and β in the ordered basis $\mathcal{B}$.

Exercises

1. Let V be a vector space and $(\ |\ )$ an inner product on V.
   (a) Show that $(0|\beta) = 0$ for all β in V.
   (b) Show that if $(\alpha|\beta) = 0$ for all β in V, then $\alpha = 0$.
2. Let V be a vector space over F. Show that the sum of two inner products on V is an inner product on V. Is the difference of two inner products an inner product? Show that a positive multiple of an inner product is an inner product.
3. Describe explicitly all inner products on R^1 and on C^1.
4. Verify that the standard inner product on F^n is an inner product.
5. Let $(\ |\ )$ be the standard inner product on R^2.
   (a) Let $\alpha = (1, 2)$, $\beta = (-1, 1)$. If γ is a vector such that $(\alpha|\gamma) = -1$ and $(\beta|\gamma) = 3$, find γ.
   (b) Show that for any α in R^2 we have $\alpha = (\alpha|\epsilon_1)\epsilon_1 + (\alpha|\epsilon_2)\epsilon_2$.
6. Let $(\ |\ )$ be the standard inner product on R^2, and let T be the linear operator $T(x_1, x_2) = (-x_2, x_1)$. Now T is 'rotation through 90°' and has the property that $(\alpha|T\alpha) = 0$ for all α in R^2. Find all inner products $[\ |\ ]$ on R^2 such that $[\alpha|T\alpha] = 0$ for each α.
7. Let $(\ |\ )$ be the standard inner product on C^2. Prove that there is no non-zero linear operator on C^2 such that $(\alpha|T\alpha) = 0$ for every α in C^2. Generalize.

8. Let A be a 2×2 matrix with real entries. For X, Y in $R^{2\times 1}$ let

$$f_A(X, Y) = Y^t AX.$$

Show that f_A is an inner product on $R^{2\times 1}$ if and only if $A = A^t$, $A_{11} > 0$, $A_{22} > 0$, and $\det A > 0$.

9. Let V be a real or complex vector space with an inner product. Show that the quadratic form determined by the inner product satisfies the **parallelogram law**

$$\|\alpha + \beta\|^2 + \|\alpha - \beta\|^2 = 2\|\alpha\|^2 + 2\|\beta\|^2.$$

10. Let $(\ |\ )$ be the inner product on R^2 defined in Example 2, and let $\mathcal{B}$ be the standard ordered basis for R^2. Find the matrix of this inner product relative to $\mathcal{B}$.

11. Show that the formula

$$\left(\sum_j a_j x^j \Big| \sum_k b_k x^k\right) = \sum_{j,k} \frac{a_j b_k}{j + k + 1}$$

defines an inner product on the space $R[x]$ of polynomials over the field R. Let W be the subspace of polynomials of degree less than or equal to n. Restrict the above inner product to W, and find the matrix of this inner product on W, relative to the ordered basis $\{1, x, x^2, \ldots, x^n\}$. (*Hint:* To show that the formula defines an inner product, observe that

$$(f|g) = \int_0^1 f(t)g(t)\,dt$$

and work with the integral.)

12. Let V be a finite-dimensional vector space and let $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$ be a basis for V. Let $(\ |\ )$ be an inner product on V. If $c_1, \ldots, c_n$ are any n scalars, show that there is exactly one vector α in V such that $(\alpha|\alpha_j) = c_j$, $j = 1, \ldots, n$.

13. Let V be a complex vector space. A function J from V into V is called a **conjugation** if $J(\alpha + \beta) = J(\alpha) + J(\beta)$, $J(c\alpha) = \bar{c}J(\alpha)$, and $J(J(\alpha)) = \alpha$, for all scalars c and all α, β in V. If J is a conjugation show that:

(a) The set W of all α in V such that $J\alpha = \alpha$ is a vector space over R with respect to the operations defined in V.

(b) For each α in V there exist unique vectors β, γ in W such that $\alpha = \beta + i\gamma$.

14. Let V be a complex vector space and W a subset of V with the following properties:

(a) W is a real vector space with respect to the operations defined in V.

(b) For each α in V there exist unique vectors β, γ in W such that $\alpha = \beta + i\gamma$.

Show that the equation $J\alpha = \beta - i\gamma$ defines a conjugation on V such that $J\alpha = \alpha$ if and only if α belongs to W, and show also that J is the only conjugation on V with this property.

15. Find all conjugations on C^1 and C^2.

16. Let W be a finite-dimensional real subspace of a complex vector space V. Show that W satisfies condition (b) of Exercise 14 if and only if every basis of W is also a basis of V.

17. Let V be a complex vector space, J a conjugation on V, W the set of α in V such that $J\alpha = \alpha$, and f an inner product on W. Show that:

(a) There is a unique inner product g on V such that $g(\alpha, \beta) = f(\alpha, \beta)$ for all α, β in W,

(b) $g(J\alpha, J\beta) = g(\beta, \alpha)$ for all α, β in V.

What does part (a) say about the relation between the standard inner products on R^1 and C^1, or on R^n and C^n?

8.2. Inner Product Spaces

Now that we have some idea of what an inner product is, we shall turn our attention to what can be said about the combination of a vector space and some particular inner product on it. Specifically, we shall establish the basic properties of the concepts of 'length' and 'orthogonality' which are imposed on the space by the inner product.

Definition. *An* **inner product space** *is a real or complex vector space, together with a specified inner product on that space.*

A finite-dimensional real inner product space is often called a **Euclidean space.** A complex inner product space is often referred to as a **unitary space.**

Theorem 1. *If V is an inner product space, then for any vectors α, β in V and any scalar c*

(i) $\|c\alpha\| = |c|\,\|\alpha\|$;

(ii) $\|\alpha\| > 0$ *for* $\alpha \neq 0$;

(iii) $|(\alpha|\beta)| \leq \|\alpha\|\,\|\beta\|$;

(iv) $\|\alpha + \beta\| \leq \|\alpha\| + \|\beta\|$.

Proof. Statements (i) and (ii) follow almost immediately from the various definitions involved. The inequality in (iii) is clearly valid when $\alpha = 0$. If $\alpha \neq 0$, put

$$\gamma = \beta - \frac{(\beta|\alpha)}{\|\alpha\|^2}\alpha.$$

Then $(\gamma|\alpha) = 0$ and

$$\begin{aligned} 0 \leq \|\gamma\|^2 &= \left(\beta - \frac{(\beta|\alpha)}{\|\alpha\|^2}\alpha \,\Big|\, \beta - \frac{(\beta|\alpha)}{\|\alpha\|^2}\alpha\right) \\ &= (\beta|\beta) - \frac{(\beta|\alpha)(\alpha|\beta)}{\|\alpha\|^2} \\ &= \|\beta\|^2 - \frac{|(\alpha|\beta)|^2}{\|\alpha\|^2}. \end{aligned}$$

Hence $|(\alpha|\beta)|^2 \leq \|\alpha\|^2\|\beta\|^2$. Now using (c) we find that

$$\begin{aligned}\|\alpha+\beta\|^2 &= \|\alpha\|^2 + (\alpha|\beta) + (\beta|\alpha) + \|\beta\|^2\\ &= \|\alpha\|^2 + 2\,\mathrm{Re}\,(\alpha|\beta) + \|\beta\|^2\\ &\leq \|\alpha\|^2 + 2\,\|\alpha\|\,\|\beta\| + \|\beta\|^2\\ &= (\|\alpha\| + \|\beta\|)^2.\end{aligned}$$

Thus, $\|\alpha + \beta\| \leq \|\alpha\| + \|\beta\|$. ▌

The inequality in (iii) is called the **Cauchy-Schwarz inequality.** It has a wide variety of applications. The proof shows that if (for example) α is non-zero, then $|(\alpha|\beta)| < \|\alpha\|\,\|\beta\|$ unless

$$\beta = \frac{(\beta|\alpha)}{\|\alpha\|^2}\alpha.$$

Thus, equality occurs in (iii) if and only if α and β are linearly dependent.

Example 7. If we apply the Cauchy-Schwarz inequality to the inner products given in Examples 1, 2, 3, and 5, we obtain the following:

(a) $$|\Sigma\, x_k\bar{y}_k| \leq (\Sigma\, |x_k|^2)^{1/2}(\Sigma\, |y_k|^2)^{1/2}$$

(b) $$|x_1y_1 - x_2y_1 - x_1y_2 + 4x_2y_2| \leq ((x_1 - x_2)^2 + 3x_2^2)^{1/2}((y_1 - y_2)^2 + 3y_2^2)^{1/2}$$

(c) $$|\mathrm{tr}\,(AB^*)| \leq (\mathrm{tr}\,(AA^*))^{1/2}(\mathrm{tr}\,(BB^*))^{1/2}$$

(d) $$\left|\int_0^1 f(x)\overline{g(x)}\,dx\right| \leq \left(\int_0^1 |f(x)|^2\,dx\right)^{1/2}\left(\int_0^1 |g(x)|^2\,dx\right)^{1/2}.$$

Definitions. *Let α and β be vectors in an inner product space V. Then α is* **orthogonal** *to β if $(\alpha|\beta) = 0$; since this implies β is orthogonal to α, we often simply say that α and β are orthogonal. If S is a set of vectors in V, S is called an* **orthogonal set** *provided all pairs of distinct vectors in S are orthogonal. An* **orthonormal set** *is an orthogonal set S with the additional property that $\|\alpha\| = 1$ for every α in S.*

The zero vector is orthogonal to every vector in V and is the only vector with this property. It is appropriate to think of an orthonormal set as a set of mutually perpendicular vectors, each having length 1.

Example 8. The standard basis of either R^n or C^n is an orthonormal set with respect to the standard inner product.

Example 9. The vector (x, y) in R^2 is orthogonal to $(-y, x)$ with respect to the standard inner product, for

$$((x, y)|(-y, x)) = -xy + yx = 0.$$

However, if R^2 is equipped with the inner product of Example 2, then (x, y) and $(-y, x)$ are orthogonal if and only if

$$y = \tfrac{1}{2}(-3 \pm \sqrt{13})x.$$

Example 10. Let V be $C^{n\times n}$, the space of complex $n \times n$ matrices, and let E^{pq} be the matrix whose only non-zero entry is a 1 in row p and column q. Then the set of all such matrices E^{pq} is orthonormal with respect to the inner product given in Example 3. For

$$(E^{pq}|E^{rs}) = \mathrm{tr}\,(E^{pq}E^{sr}) = \delta_{qs}\,\mathrm{tr}\,(E^{pr}) = \delta_{qs}\delta_{pr}.$$

Example 11. Let V be the space of continuous complex-valued (or real-valued) functions on the interval $0 \leq x \leq 1$ with the inner product

$$(f|g) = \int_0^1 f(x)\overline{g(x)}\,dx.$$

Suppose $f_n(x) = \sqrt{2}\,\cos 2\pi nx$ and that $g_n(x) = \sqrt{2}\,\sin 2\pi nx$. Then $\{1, f_1, g_1, f_2, g_2, \ldots\}$ is an infinite orthonormal set. In the complex case, we may also form the linear combinations

$$\frac{1}{\sqrt{2}}(f_n + ig_n), \qquad n = 1, 2, \ldots.$$

In this way we get a new orthonormal set S which consists of all functions of the form

$$h_n(x) = e^{2\pi inx}, \qquad n = \pm 1, \pm 2, \ldots.$$

The set S' obtained from S by adjoining the constant function 1 is also orthonormal. We assume here that the reader is familiar with the calculation of the integrals in question.

The orthonormal sets given in the examples above are all linearly independent. We show now that this is necessarily the case.

Theorem 2. *An orthogonal set of non-zero vectors is linearly independent.*

Proof. Let S be a finite or infinite orthogonal set of non-zero vectors in a given inner product space. Suppose $\alpha_1, \alpha_2, \ldots, \alpha_m$ are distinct vectors in S and that

$$\beta = c_1\alpha_1 + c_2\alpha_2 + \cdots + c_m\alpha_m.$$

Then

$$\begin{aligned}(\beta|\alpha_k) &= (\sum_j c_j\alpha_j|\alpha_k)\\ &= \sum_j c_j(\alpha_j|\alpha_k)\\ &= c_k(\alpha_k|\alpha_k).\end{aligned}$$

Since $(\alpha_k|\alpha_k) \neq 0$, it follows that

$$c_k = \frac{(\beta|\alpha_k)}{||\alpha_k||^2}, \qquad 1 \le k \le m.$$

Thus when $\beta = 0$, each $c_k = 0$; so S is an independent set. ▮

Corollary. *If a vector β is a linear combination of an orthogonal sequence of non-zero vectors $\alpha_1, \ldots, \alpha_m$, then β is the particular linear combination*

$$\beta = \sum_{k=1}^{m} \frac{(\beta|\alpha_k)}{||\alpha_k||^2}\alpha_k. \tag{8-8}$$

This corollary follows from the proof of the theorem. There is another corollary which although obvious, should be mentioned. If $\{\alpha_1, \ldots, \alpha_m\}$ is an orthogonal set of non-zero vectors in a finite-dimensional inner product space V, then $m \le \dim V$. This says that the number of mutually orthogonal directions in V cannot exceed the algebraically defined dimension of V. The maximum number of mutually orthogonal directions in V is what one would intuitively regard as the geometric dimension of V, and we have just seen that this is not greater than the algebraic dimension. The fact that these two dimensions are equal is a particular corollary of the next result.

Theorem 3. *Let V be an inner product space and let $\beta_1, \ldots, \beta_n$ be any independent vectors in V. Then one may construct orthogonal vectors $\alpha_1, \ldots, \alpha_n$ in V such that for each k = 1, 2, . . . , n the set*

$$\{\alpha_1, \ldots, \alpha_k\}$$

is a basis for the subspace spanned by $\beta_1, \ldots, \beta_k$.

Proof. The vectors $\alpha_1, \ldots, \alpha_n$ will be obtained by means of a construction known as the **Gram-Schmidt orthogonalization process.** First let $\alpha_1 = \beta_1$. The other vectors are then given inductively as follows: Suppose $\alpha_1, \ldots, \alpha_m$ $(1 \le m < n)$ have been chosen so that for every k

$$\{\alpha_1, \ldots, \alpha_k\}, \qquad 1 \le k \le m$$

is an orthogonal basis for the subspace of V that is spanned by $\beta_1, \ldots, \beta_k$. To construct the next vector α_{m+1}, let

$$\alpha_{m+1} = \beta_{m+1} - \sum_{k=1}^{m} \frac{(\beta_{m+1}|\alpha_k)}{||\alpha_k||^2}\alpha_k. \tag{8-9}$$

Then $\alpha_{m+1} \neq 0$. For otherwise β_{m+1} is a linear combination of $\alpha_1, \ldots, \alpha_m$ and hence a linear combination of $\beta_1, \ldots, \beta_m$. Furthermore, if $1 \le j \le m$, then

$$\begin{aligned}(\alpha_{m+1}|\alpha_j) &= (\beta_{m+1}|\alpha_j) - \sum_{k=1}^{m} \frac{(\beta_{m+1}|\alpha_k)}{||\alpha_k||^2}(\alpha_k|\alpha_j)\\ &= (\beta_{m+1}|\alpha_j) - (\beta_{m+1}|\alpha_j)\\ &= 0.\end{aligned}$$

Therefore $\{\alpha_1, \ldots, \alpha_{m+1}\}$ is an orthogonal set consisting of $m + 1$ non-zero vectors in the subspace spanned by $\beta_1, \ldots, \beta_{m+1}$. By Theorem 2, it is a basis for this subspace. Thus the vectors $\alpha_1, \ldots, \alpha_n$ may be constructed one after the other in accordance with (8-9). In particular, when $n = 4$, we have

$$\begin{aligned}\alpha_1 &= \beta_1\\ \alpha_2 &= \beta_2 - \frac{(\beta_2|\alpha_1)}{||\alpha_1||^2}\alpha_1\\ \alpha_3 &= \beta_3 - \frac{(\beta_3|\alpha_1)}{||\alpha_1||^2}\alpha_1 - \frac{(\beta_3|\alpha_2)}{||\alpha_2||^2}\alpha_2\\ \alpha_4 &= \beta_4 - \frac{(\beta_4|\alpha_1)}{||\alpha_1||^2}\alpha_1 - \frac{(\beta_4|\alpha_2)}{||\alpha_2||^2}\alpha_2 - \frac{(\beta_4|\alpha_3)}{||\alpha_3||^2}\alpha_3.\end{aligned} \tag{8-10}$$ ▮

Corollary. *Every finite-dimensional inner product space has an orthonormal basis.*

Proof. Let V be a finite-dimensional inner product space and $\{\beta_1, \ldots, \beta_n\}$ a basis for V. Apply the Gram-Schmidt process to construct an orthogonal basis $\{\alpha_1, \ldots, \alpha_n\}$. Then to obtain an orthonormal basis, simply replace each vector α_k by $\alpha_k/||\alpha_k||$. ▮

One of the main advantages which orthonormal bases have over arbitrary bases is that computations involving coordinates are simpler. To indicate in general terms why this is true, suppose that V is a finite-dimensional inner product space. Then, as in the last section, we may use Equation (8-5) to associate a matrix G with every ordered basis $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$ of V. Using this matrix

$$G_{jk} = (\alpha_k|\alpha_j),$$

we may compute inner products in terms of coordinates. If $\mathcal{B}$ is an orthonormal basis, then G is the identity matrix, and for any scalars x_j and y_k

$$(\sum_j x_j\alpha_j|\sum_k y_k\alpha_k) = \sum_j x_j\bar{y}_j.$$

Thus in terms of an orthonormal basis, the inner product in V looks like the standard inner product in F^n.

Although it is of limited practical use for computations, it is interesting to note that the Gram-Schmidt process may also be used to test for linear dependence. For suppose $\beta_1, \ldots, \beta_n$ are linearly dependent vectors in an inner product space V. To exclude a trivial case, assume that $\beta_1 \neq 0$. Let m be the largest integer for which $\beta_1, \ldots, \beta_m$ are independent. Then $1 \le m < n$. Let $\alpha_1, \ldots, \alpha_m$ be the vectors obtained by applying the orthogonalization process to $\beta_1, \ldots, \beta_m$. Then the vector α_{m+1} given by (8-9) is necessarily 0. For α_{m+1} is in the subspace spanned by $\alpha_1, \ldots, \alpha_m$ and orthogonal to each of these vectors; hence it is 0 by (8-8). Conversely, if $\alpha_1, \ldots, \alpha_m$ are different from 0 and $\alpha_{m+1} = 0$, then $\beta_1, \ldots, \beta_{m+1}$ are linearly dependent.

EXAMPLE 12. Consider the vectors

$$\beta_1 = (3, 0, 4)$$
$$\beta_2 = (-1, 0, 7)$$
$$\beta_3 = (2, 9, 11)$$

in R^3 equipped with the standard inner product. Applying the Gram-Schmidt process to $\beta_1, \beta_2, \beta_3$, we obtain the following vectors.

$$\begin{aligned}
\alpha_1 &= (3, 0, 4)\\
\alpha_2 &= (-1, 0, 7) - \frac{((-1, 0, 7)|(3, 0, 4))}{25}(3, 0, 4)\\
&= (-1, 0, 7) - (3, 0, 4)\\
&= (-4, 0, 3)\\
\alpha_3 &= (2, 9, 11) - \frac{((2, 9, 11)|(3, 0, 4))}{25}(3, 0, 4)\\
&\quad - \frac{((2, 9, 11)|(-4, 0, 3))}{25}(-4, 0, 3)\\
&= (2, 9, 11) - 2(3, 0, 4) - -4, 0, 3)\\
&= (0, 9, 0).
\end{aligned}$$

These vectors are evidently non-zero and mutually orthogonal. Hence $\{\alpha_1, \alpha_2, \alpha_3\}$ is an orthogonal basis for R^3. To express an arbitrary vector (x_1, x_2, x_3) in R^3 as a linear combination of $\alpha_1, \alpha_2, \alpha_3$ it is *not* necessary to solve any linear equations. For it suffices to use (8-8). Thus

$$(x_1, x_2, x_3) = \frac{3x_1 + 4x_3}{25}\alpha_1 + \frac{-4x_1 + 3x_3}{25}\alpha_2 + \frac{x_2}{9}\alpha_3$$

as is readily verified. In particular,

$$(1, 2, 3) = \tfrac{3}{5}(3, 0, 4) + \tfrac{1}{5}(-4, 0, 3) + \tfrac{2}{9}(0, 9, 0).$$

To put this point in another way, what we have shown is the following: The basis $\{f_1, f_2, f_3\}$ of $(R^3)^*$ which is dual to the basis $\{\alpha_1, \alpha_2, \alpha_3\}$ is defined explicitly by the equations

$$f_1(x_1, x_2, x_3) = \frac{3x_1 + 4x_3}{25}$$
$$f_2(x_1, x_2, x_3) = \frac{-4x_1 + 3x_3}{25}$$
$$f_3(x_1, x_2, x_3) = \frac{x_2}{9}$$

and these equations may be written more generally in the form

$$f_j(x_1, x_2, x_3) = \frac{((x_1, x_2, x_3)|\alpha_j)}{||\alpha_j||^2}.$$

Finally, note that from $\alpha_1, \alpha_2, \alpha_3$ we get the orthonormal basis

$$\tfrac{1}{5}(3, 0, 4),\quad \tfrac{1}{5}(-4, 0, 3),\quad (0, 1, 0).$$

EXAMPLE 13. Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ where a, b, c, and d are complex numbers. Set $\beta_1 = (a, b)$, $\beta_2 = (c, d)$, and suppose that $\beta_1 \neq 0$. If we apply the orthogonalization process to β_1, β_2, using the standard inner product in C^2, we obtain the following vectors:

$$\begin{aligned}
\alpha_1 &= (a, b)\\
\alpha_2 &= (c, d) - \frac{((c, d)|(a, b))}{|a|^2 + |b|^2}(a, b)\\
&= (c, d) - \frac{(c\bar{a} + d\bar{b})}{|a|^2 + |b|^2}(a, b)\\
&= \left(\frac{cb\bar{b} - d\bar{b}a}{|a|^2 + |b|^2}, \frac{d\bar{a}a - c\bar{a}b}{|a|^2 + |b|^2}\right)\\
&= \frac{\det A}{|a|^2 + |b|^2}(-\bar{b}, \bar{a}).
\end{aligned}$$

Now the general theory tells us that $\alpha_2 \neq 0$ if and only if β_1, β_2 are linearly independent. On the other hand, the formula for α_2 shows that this is the case if and only if $\det A \neq 0$.

In essence, the Gram-Schmidt process consists of repeated applications of a basic geometric operation called orthogonal projection, and it is best understood from this point of view. The method of orthogonal projection also arises naturally in the solution of an important approximation problem.

Suppose W is a subspace of an inner product space V, and let β be an arbitrary vector in V. The problem is to find a best possible approximation to β by vectors in W. This means we want to find a vector α for which $||\beta - \alpha||$ is as small as possible subject to the restriction that α should belong to W. Let us make our language precise.

A **best approximation** to β by vectors in W is a vector α in W such that

$$||\beta - \alpha|| \leq ||\beta - \gamma||$$

for every vector γ in W.

By looking at this problem in R^2 or in R^3, one sees intuitively that a best approximation to β by vectors in W ought to be a vector α in W such that $\beta - \alpha$ is perpendicular (orthogonal) to W and that there ought to be exactly one such α. These intuitive ideas are correct for finite-dimensional subspaces and for some, but not all, infinite-dimensional subspaces. Since the precise situation is too complicated to treat here, we shall only prove the following result.

Theorem 4. *Let W be a subspace of an inner product space V and let β be a vector in V.*

(i) *The vector α in W is a best approximation to β by vectors in W if and only if $\beta - \alpha$ is orthogonal to every vector in W.*

(ii) *If a best approximation to β by vectors in W exists, it is unique.*

(iii) *If W is finite-dimensional and $\{\alpha_1, \ldots, \alpha_n\}$ is any orthonormal basis for W, then the vector*

$$\alpha = \sum_k \frac{(\beta|\alpha_k)}{\|\alpha_k\|^2}\alpha_k$$

is the (unique) best approximation to β by vectors in W.

Proof. First note that if γ is any vector in V, then $\beta - \gamma = (\beta - \alpha) + (\alpha - \gamma)$, and

$$\|\beta - \gamma\|^2 = \|\beta - \alpha\|^2 + 2\,\mathrm{Re}\,(\beta - \alpha|\alpha - \gamma) + \|\alpha - \gamma\|^2.$$

Now suppose $\beta - \alpha$ is orthogonal to every vector in W, that γ is in W and that $\gamma \neq \alpha$. Then, since $\alpha - \gamma$ is in W, it follows that

$$\begin{aligned}\|\beta - \gamma\|^2 &= \|\beta - \alpha\|^2 + \|\alpha - \gamma\|^2\\ &> \|\beta - \alpha\|^2.\end{aligned}$$

Conversely, suppose that $\|\beta - \gamma\| \geq \|\beta - \alpha\|$ for every γ in W. Then from the first equation above it follows that

$$2\,\mathrm{Re}\,(\beta - \alpha|\alpha - \gamma) + \|\alpha - \gamma\|^2 \geq 0$$

for all γ in W. Since every vector in W may be expressed in the form $\alpha - \gamma$ with γ in W, we see that

$$2\,\mathrm{Re}\,(\beta - \alpha|\tau) + \|\tau\|^2 \geq 0$$

for every τ in W. In particular, if γ is in W and $\gamma \neq \alpha$, we may take

$$\tau = -\frac{(\beta - \alpha|\alpha - \gamma)}{\|\alpha - \gamma\|^2}(\alpha - \gamma).$$

Then the inequality reduces to the statement

$$-2\frac{|(\beta - \alpha|\alpha - \gamma)|^2}{\|\alpha - \gamma\|^2} + \frac{|(\beta - \alpha|\alpha - \gamma)|^2}{\|\alpha - \gamma\|^2} \geq 0.$$

This holds if and only if $(\beta - \alpha|\alpha - \gamma) = 0$. Therefore, $\beta - \alpha$ is orthogonal to every vector in W. This completes the proof of the equivalence of the two conditions on α given in (i). The orthogonality condition is evidently satisfied by at most one vector in W, which proves (ii).

Now suppose that W is a finite-dimensional subspace of V. Then we know, as a corollary of Theorem 3, that W has an orthogonal basis. Let $\{\alpha_1, \ldots, \alpha_n\}$ be any orthogonal basis for W and define α by (8-11). Then, by the computation in the proof of Theorem 3, $\beta - \alpha$ is orthogonal to each of the vectors α_k ($\beta - \alpha$ is the vector obtained at the last stage when the orthogonalization process is applied to $\alpha_1, \ldots, \alpha_n, \beta$). Thus $\beta - \alpha$ is orthogonal to every linear combination of $\alpha_1, \ldots, \alpha_n$, i.e., to every vector in W. If γ is in W and $\gamma \neq \alpha$, it follows that $\|\beta - \gamma\| > \|\beta - \alpha\|$. Therefore, α is the best approximation to β that lies in W. ▌

Definition. *Let V be an inner product space and S any set of vectors in V. The* **orthogonal complement** *of S is the set $S^\perp$ of all vectors in V which are orthogonal to every vector in S.*

The orthogonal complement of V is the zero subspace, and conversely $\{0\}^\perp = V$. If S is any subset of V, its orthogonal complement $S^\perp$ (S perp) is always a subspace of V. For S is non-empty, since it contains 0; and whenever α and β are in $S^\perp$ and c is any scalar,

$$\begin{aligned}(c\alpha + \beta|\gamma) &= c(\alpha|\gamma) + (\beta|\gamma)\\ &= c0 + 0\\ &= 0\end{aligned}$$

for every γ in S, thus $c\alpha + \beta$ also lies in S. In Theorem 4 the characteristic property of the vector α is that it is the only vector in W such that $\beta - \alpha$ belongs to $W^\perp$.

Definition. *Whenever the vector α in Theorem 4 exists it is called the* **orthogonal projection of β on W.** *If every vector in V has an orthogonal projection on W, the mapping that assigns to each vector in V its orthogonal projection on W is called the* **orthogonal projection of V on W.**

By Theorem 4, the orthogonal projection of an inner product space on a finite-dimensional subspace always exists. But Theorem 4 also implies the following result.

Corollary. *Let V be an inner product space, W a finite-dimensional subspace, and E the orthogonal projection of V on W. Then the mapping*

$$\beta \to \beta - E\beta$$

is the orthogonal projection of V on $W^\perp$.

Proof. Let β be an arbitrary vector in V. Then $\beta - E\beta$ is in $W^\perp$, and for any γ in $W^\perp$, $\beta - \gamma = E\beta + (\beta - E\beta - \gamma)$. Since $E\beta$ is in W and $\beta - E\beta - \gamma$ is in $W^\perp$, it follows that

$$\|\beta - \gamma\|^2 = \|E\beta\|^2 + \|\beta - E\beta - \gamma\|^2$$
$$\geq \|\beta - (\beta - E\beta)\|^2$$

with strict inequality when $\gamma \neq \beta - E\beta$. Therefore, $\beta - E\beta$ is the best approximation to β by vectors in $W^\perp$. ∎

EXAMPLE 14. Give R^3 the standard inner product. Then the orthogonal projection of $(-10, 2, 8)$ on the subspace W that is spanned by $(3, 12, -1)$ is the vector

$$\alpha = \frac{((-10, 2, 8)|(3, 12, -1))}{9 + 144 + 1}(3, 12, -1)$$
$$= \frac{-14}{154}(3, 12, -1).$$

The orthogonal projection of R^3 on W is the linear transformation E defined by

$$(x_1, x_2, x_3) \to \left(\frac{3x_1 + 12x_2 - x_3}{154}\right)(3, 12, -1).$$

The rank of E is clearly 1; hence its nullity is 2. On the other hand,

$$E(x_1, x_2, x_3) = (0, 0, 0)$$

if and only if $3x_1 + 12x_2 - x_3 = 0$. This is the case if and only if (x_1, x_2, x_3) is in $W^\perp$. Therefore, $W^\perp$ is the null space of E, and $\dim (W^\perp) = 2$. Computing

$$(x_1, x_2, x_3) - \left(\frac{3x_1 + 12x_2 - x_3}{154}\right)(3, 12, -1)$$

we see that the orthogonal projection of R^3 on $W^\perp$ is the linear transformation $I - E$ that maps the vector (x_1, x_2, x_3) onto the vector

$$\frac{1}{154}(145x_1 - 36x_2 + 3x_3, -36x_1 + 10x_2 + 12x_3, 3x_1 + 12x_2 + 153x_3).$$

The observations made in Example 14 generalize in the following fashion.

Theorem 5. *Let W be a finite-dimensional subspace of an inner product space V and let E be the orthogonal projection of V on W. Then E is an idempotent linear transformation of V onto W, $W^\perp$ is the null space of E, and*

$$V = W \oplus W^\perp.$$

Proof. Let β be an arbitrary vector in V. Then $E\beta$ is the best approximation to β that lies in W. In particular, $E\beta = \beta$ when β is in W. Therefore, $E(E\beta) = E\beta$ for every β in V; that is, E is idempotent: $E^2 = E$. To prove that E is a linear transformation, let α and β be any vectors in

V and c an arbitrary scalar. Then, by Theorem 4, $\alpha - E\alpha$ and $\beta - E\beta$ are each orthogonal to every vector in W. Hence the vector

$$c(\alpha - E\alpha) + (\beta - E\beta) = (c\alpha + \beta) - (cE\alpha + E\beta)$$

also belongs to $W^\perp$. Since $cE\alpha + E\beta$ is a vector in W, it follows from Theorem 4 that

$$E(c\alpha + \beta) = cE\alpha + E\beta.$$

Of course, one may also prove the linearity of E by using (8-11). Again let β be any vector in V. Then $E\beta$ is the unique vector in W such that $\beta - E\beta$ is in $W^\perp$. Thus $E\beta = 0$ when β is in $W^\perp$. Conversely, β is in $W^\perp$ when $E\beta = 0$. Thus $W^\perp$ is the null space of E. The equation

$$\beta = E\beta + \beta - E\beta$$

shows that $V = W + W^\perp$; moreover, $W \cap W^\perp = \{0\}$. For if α is a vector in $W \cap W^\perp$, then $(\alpha|\alpha) = 0$. Therefore, $\alpha = 0$, and V is the direct sum of W and $W^\perp$. ∎

Corollary. *Under the conditions of the theorem, I − E is the orthogonal projection of V on $W^\perp$. It is an idempotent linear transformation of V onto $W^\perp$ with null space W.*

Proof. We have already seen that the mapping $\beta \to \beta - E\beta$ is the orthogonal projection of V on $W^\perp$. Since E is a linear transformation, this projection on $W^\perp$ is the linear transformation $I - E$. From its geometric properties one sees that $I - E$ is an idempotent transformation of V onto W. This also follows from the computation

$$(I - E)(I - E) = I - E - E + E^2$$
$$= I - E.$$

Moreover, $(I - E)\beta = 0$ if and only if $\beta = E\beta$, and this is the case if and only if β is in W. Therefore W is the null space of $I - E$. ∎

The Gram-Schmidt process may now be described geometrically in the following way. Given an inner product space V and vectors $\beta_1, \ldots, \beta_n$ in V, let P_k $(k > 1)$ be the orthogonal projection of V on the orthogonal complement of the subspace spanned by $\beta_1, \ldots, \beta_{k-1}$, and set $P_1 = I$. Then the vectors one obtains by applying the orthogonalization process to $\beta_1, \ldots, \beta_n$ are defined by the equations

$$\text{(8-12)} \qquad \alpha_k = P_k\beta_k, \qquad 1 \leq k \leq n.$$

Theorem 5 implies another result known as **Bessel's inequality**.

Corollary. *Let $\{\alpha_1, \ldots, \alpha_n\}$ be an orthogonal set of non-zero vectors in an inner product space V. If β is any vector in V, then*

$$\sum_k \frac{|(\beta|\alpha_k)|^2}{\|\alpha_k\|^2} \leq \|\beta\|^2$$

and equality holds if and only if

$$\beta = \sum_k \frac{(\beta|\alpha_k)}{||\alpha_k||^2}\alpha_k.$$

Proof. Let $\gamma = \sum_k [(\beta|\alpha_k)/||\alpha_k||^2]\,\alpha_k$. Then $\beta = \gamma + \delta$ where $(\gamma|\delta) = 0$. Hence

$$||\beta||^2 = ||\gamma||^2 + ||\delta||^2.$$

It now suffices to prove that

$$||\gamma||^2 = \sum_k \frac{|(\beta|\alpha_k)|^2}{||\alpha_k||^2}.$$

This is straightforward computation in which one uses the fact that $(\alpha_j|\alpha_k) = 0$ for $j \neq k$. ∎

In the special case in which $\{\alpha_1, \dots, \alpha_n\}$ is an orthonormal set, Bessel's inequality says that

$$\sum_k |(\beta|\alpha_k)|^2 \leq ||\beta||^2.$$

The corollary also tells us in this case that β is in the subspace spanned by $\alpha_1, \dots, \alpha_n$ if and only if

$$\beta = \sum_k (\beta|\alpha_k)\,\alpha_k$$

or if and only if Bessel's inequality is actually an equality. Of course, in the event that V is finite dimensional and $\{\alpha_1, \dots, \alpha_n\}$ is an orthogonal basis for V, the above formula holds for every vector β in V. In other words, if $\{\alpha_1, \dots, \alpha_n\}$ is an orthonormal basis for V, the kth coordinate of β in the ordered basis $\{\alpha_1, \dots, \alpha_n\}$ is $(\beta|\alpha_k)$.

EXAMPLE 15. We shall apply the last corollary to the orthogonal sets described in Example 11. We find that

(a) $$\sum_{k=-n}^{n} \left|\int_0^1 f(t)e^{-2\pi ikt}\,dt\right|^2 \leq \int_0^1 |f(t)|^2\,dt$$

(b) $$\int_0^1 \left|\sum_{k=-n}^{n} c_k e^{2\pi ikt}\right|^2 dt = \sum_{k=-n}^{n} |c_k|^2$$

(c) $$\int_0^1 (\sqrt{2}\cos 2\pi t + \sqrt{2}\sin 4\pi t)^2\,dt = 1 + 1 = 2.$$

Exercises

1. Consider R^4 with the standard inner product. Let W be the subspace of R^4 consisting of all vectors which are orthogonal to both $\alpha = (1, 0, -1, 1)$ and $\beta = (2, 3, -1, 2)$. Find a basis for W.

2. Apply the Gram-Schmidt process to the vectors $\beta_1 = (1, 0, 1)$, $\beta_2 = (1, 0, -1)$, $\beta_3 = (0, 3, 4)$, to obtain an orthonormal basis for R^3 with the standard inner product.

3. Consider C^3, with the standard inner product. Find an orthonormal basis for the subspace spanned by $\beta_1 = (1, 0, i)$ and $\beta_2 = (2, 1, 1 + i)$.

4. Let V be an inner product space. The **distance** between two vectors α and β in V is defined by

$$d(\alpha, \beta) = ||\alpha - \beta||.$$

Show that
  (a) $d(\alpha, \beta) \geq 0$;
  (b) $d(\alpha, \beta) = 0$ if and only if $\alpha = \beta$;
  (c) $d(\alpha, \beta) = d(\beta, \alpha)$;
  (d) $d(\alpha, \beta) \leq d(\alpha, \gamma) + d(\gamma, \beta)$.

5. Let V be an inner product space, and let α, β be vectors in V. Show that $\alpha = \beta$ if and only if $(\alpha|\gamma) = (\beta|\gamma)$ for every γ in V.

6. Let W be the subspace of R^2 spanned by the vector (3, 4). Using the standard inner product, let E be the orthogonal projection of R^2 onto W. Find
  (a) a formula for $E(x_1, x_2)$;
  (b) the matrix of E in the standard ordered basis;
  (c) $W^\perp$;
  (d) an orthonormal basis in which E is represented by the matrix

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

7. Let V be the inner product space consisting of R^2 and the inner product whose quadratic form is defined by

$$||(x_1, x_2)||^2 = (x_1 - x_2)^2 + 3x_2^2.$$

Let E be the orthogonal projection of V onto the subspace W spanned by the vector (3, 4). Now answer the four questions of Exercise 6.

8. Find an inner product on R^2 such that $(\epsilon_1, \epsilon_2) = 2$.

9. Let V be the subspace of $R[x]$ of polynomials of degree at most 3. Equip V with the inner product

$$(f|g) = \int_0^1 f(t)g(t)\,dt.$$

  (a) Find the orthogonal complement of the subspace of scalar polynomials.
  (b) Apply the Gram-Schmidt process to the basis $\{1, x, x^2, x^3\}$.

10. Let V be the vector space of all $n \times n$ matrices over C, with the inner product $(A|B) = \operatorname{tr}(AB^*)$. Find the orthogonal complement of the subspace of diagonal matrices.

11. Let V be a finite-dimensional inner product space, and let $\{\alpha_1, \dots, \alpha_n\}$ be an orthonormal basis for V. Show that for any vectors α, β in V

$$(\alpha|\beta) = \sum_{k=1}^{n} (\alpha|\alpha_k)\overline{(\beta|\alpha_k)}.$$

12. Let W be a finite-dimensional subspace of an inner product space V, and let E be the orthogonal projection of V on W. Prove that $(E\alpha|\beta) = (\alpha|E\beta)$ for all α, β in V.

13. Let S be a subset of an inner product space V. Show that $(S^\perp)^\perp$ contains the subspace spanned by S. When V is finite-dimensional, show that $(S^\perp)^\perp$ is the subspace spanned by S.

14. Let V be a finite-dimensional inner product space, and let $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$ be an orthonormal basis for V. Let T be a linear operator on V and A the matrix of T in the ordered basis $\mathcal{B}$. Prove that

$$A_{ij} = (T\alpha_j|\alpha_i).$$

15. Suppose $V = W_1 \oplus W_2$ and that f_1 and f_2 are inner products on W_1 and W_2, respectively. Show that there is a unique inner product f on V such that

(a) $W_2 = W_1^\perp$;

(b) $f(\alpha, \beta) = f_k(\alpha, \beta)$, when α, β are in W_k, $k = 1, 2$.

16. Let V be an inner product space and W a finite-dimensional subspace of V. There are (in general) many projections which have W as their range. One of these, the orthogonal projection on W, has the property that $||E\alpha|| \leq ||\alpha||$ for every α in V. Prove that if E is a projection with range W, such that $||E\alpha|| \leq ||\alpha||$ for all α in V, then E is the orthogonal projection on W.

17. Let V be the real inner product space consisting of the space of real-valued continuous functions on the interval, $-1 \leq t \leq 1$, with the inner product

$$(f|g) = \int_{-1}^{1} f(t)g(t)\,dt.$$

Let W be the subspace of odd functions, i.e., functions satisfying $f(-t) = -f(t)$. Find the orthogonal complement of W.

8.3. Linear Functionals and Adjoints

The first portion of this section treats linear functionals on an inner product space and their relation to the inner product. The basic result is that any linear functional f on a finite-dimensional inner product space is 'inner product with a fixed vector in the space,' i.e., that such an f has the form $f(\alpha) = (\alpha|\beta)$ for some fixed β in V. We use this result to prove the existence of the 'adjoint' of a linear operator T on V, this being a linear operator T^* such that $(T\alpha|\beta) = (\alpha|T^*\beta)$ for all α and β in V. Through the use of an orthonormal basis, this adjoint operation on linear operators (passing from T to T^*) is identified with the operation of forming the conjugate transpose of a matrix. We explore slightly the analogy between the adjoint operation and conjugation on complex numbers.

Let V be any inner product space, and let β be some fixed vector in V. We define a function f_β from V into the scalar field by

$$f_\beta(\alpha) = (\alpha|\beta).$$

This function f_β is a linear functional on V, because, by its very definition, $(\alpha|\beta)$ is linear as a function of α. If V is finite-dimensional, every linear functional on V arises in this way from some β.

Theorem 6. *Let V be a finite-dimensional inner product space, and f a linear functional on V. Then there exists a unique vector β in V such that $f(\alpha) = (\alpha|\beta)$ for all α in V.*

Proof. Let $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ be an orthonormal basis for V. Put

$$\beta = \sum_{j=1}^{n} \overline{f(\alpha_j)}\alpha_j \tag{8-13}$$

and let f_β be the linear functional defined by

$$f_\beta(\alpha) = (\alpha|\beta).$$

Then

$$f_\beta(\alpha_k) = (\alpha_k|\sum_j \overline{f(\alpha_j)}\alpha_j) = f(\alpha_k).$$

Since this is true for each α_k, it follows that $f = f_\beta$. Now suppose γ is a vector in V such that $(\alpha|\beta) = (\alpha|\gamma)$ for all α. Then $(\beta - \gamma|\beta - \gamma) = 0$ and $\beta = \gamma$. Thus there is exactly one vector β determining the linear functional f in the stated manner. ∎

The proof of this theorem can be reworded slightly, in terms of the representation of linear functionals in a basis. If we choose an orthonormal basis $\{\alpha_1, \ldots, \alpha_n\}$ for V, the inner product of $\alpha = x_1\alpha_1 + \cdots + x_n\alpha_n$ and $\beta = y_1\alpha_1 + \cdots + y_n\alpha_n$ will be

$$(\alpha|\beta) = x_1\bar{y}_1 + \cdots + x_n\bar{y}_n.$$

If f is any linear functional on V, then f has the form

$$f(\alpha) = c_1x_1 + \cdots + c_nx_n$$

for some fixed scalars $c_1, \ldots, c_n$ determined by the basis. Of course $c_j = f(\alpha_j)$. If we wish to find a vector β in V such that $(\alpha|\beta) = f(\alpha)$ for all α, then clearly the coordinates y_j of β must satisfy $\bar{y}_j = c_j$ or $y_j = \overline{f(\alpha_j)}$. Accordingly,

$$\beta = \overline{f(\alpha_1)}\alpha_1 + \cdots + \overline{f(\alpha_n)}\alpha_n$$

is the desired vector.

Some further comments are in order. The proof of Theorem 6 that we have given is admirably brief, but it fails to emphasize the essential geometric fact that β lies in the orthogonal complement of the null space of f. Let W be the null space of f. Then $V = W + W^\perp$, and f is completely determined by its values on $W^\perp$. In fact, if P is the orthogonal projection of V on $W^\perp$, then

$$f(\alpha) = f(P\alpha)$$

for all α in V. Suppose $f \neq 0$. Then f is of rank 1 and $\dim (W^\perp) = 1$. If γ is any non-zero vector in $W^\perp$, it follows that

$$P\alpha = \frac{(\alpha|\gamma)}{||\gamma||^2}\gamma$$

for all α in V. Thus

$$f(\alpha) = (\alpha|\gamma) \cdot \frac{f(\gamma)}{||\gamma||^2}$$

for all α, and $\beta = [\overline{f(\gamma)}/||\gamma||^2]\,\gamma$.

EXAMPLE 16. We should give one example showing that Theorem 6 is not true without the assumption that V is finite dimensional. Let V be the vector space of polynomials over the field of complex numbers, with the inner product

$$(f|g) = \int_0^1 f(t)\overline{g(t)}\,dt.$$

This inner product can also be defined algebraically. If $f = \Sigma\, a_k x^k$ and $g = \Sigma\, b_k x^k$, then

$$(f|g) = \sum_{j,k} \frac{1}{j+k+1} a_j \bar{b}_k.$$

Let z be a fixed complex number, and let L be the linear functional 'evaluation at z':

$$L(f) = f(z).$$

Is there a polynomial g such that $(f|g) = L(f)$ for every f? The answer is no; for suppose we have

$$f(z) = \int_0^1 f(t)\overline{g(t)}\,dt$$

for every f. Let $h = x - z$, so that for any f we have $(hf)(z) = 0$. Then

$$0 = \int_0^1 h(t)f(t)\overline{g(t)}\,dt$$

for all f. In particular this holds when $f = \bar{h}g$ so that

$$\int_0^1 |h(t)|^2|g(t)|^2\,dt = 0$$

and so $hg = 0$. Since $h \neq 0$, it must be that $g = 0$. But L is not the zero functional; hence, no such g exists.

One can generalize the example somewhat, to the case where L is a linear combination of point evaluations. Suppose we select fixed complex numbers $z_1, \ldots, z_n$ and scalars $c_1, \ldots, c_n$ and let

$$L(f) = c_1 f(z_1) + \cdots + c_n f(z_n).$$

Then L is a linear functional on V, but there is no g with $L(f) = (f|g)$, unless $c_1 = c_2 = \cdots = c_n = 0$. Just repeat the above argument with $h = (x - z_1) \cdots (x - z_n)$.

We turn now to the concept of the adjoint of a linear operator.

Theorem 7. *For any linear operator T on a finite-dimensional inner product space V, there exists a unique linear operator T* on V such that*

$$(8\text{-}14) \qquad (T\alpha|\beta) = (\alpha|T^*\beta)$$

for all α, β in V.

Proof. Let β be any vector in V. Then $\alpha \to (T\alpha|\beta)$ is a linear functional on V. By Theorem 6 there is a unique vector β' in V such that $(T\alpha|\beta) = (\alpha|\beta')$ for every α in V. Let T^* denote the mapping $\beta \to \beta'$:

$$\beta' = T^*\beta.$$

We have (8-14), but we must verify that T^* is a linear operator. Let β, γ be in V and let c be a scalar. Then for any α,

$$\begin{aligned}(\alpha|T^*(c\beta + \gamma)) &= (T\alpha|c\beta + \gamma)\\ &= (T\alpha|c\beta) + (T\alpha|\gamma)\\ &= \bar{c}(T\alpha|\beta) + (T\alpha|\gamma)\\ &= \bar{c}(\alpha|T^*\beta) + (\alpha|T^*\gamma)\\ &= (\alpha|cT^*\beta) + (\alpha|T^*\gamma)\\ &= (\alpha|cT^*\beta + T^*\gamma).\end{aligned}$$

Thus $T^*(c\beta + \gamma) = cT^*\beta + T^*\gamma$ and T^* is linear.

The uniqueness of T^* is clear. For any β in V, the vector $T^*\beta$ is uniquely determined as the vector β' such that $(T\alpha|\beta) = (\alpha|\beta')$ for every α. ∎

Theorem 8. *Let V be a finite-dimensional inner product space and let $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$ be an (ordered) orthonormal basis for V. Let T be a linear operator on V and let A be the matrix of T in the ordered basis $\mathcal{B}$. Then $A_{kj} = (T\alpha_j|\alpha_k)$.*

Proof. Since $\mathcal{B}$ is an orthonormal basis, we have

$$\alpha = \sum_{k=1}^{n} (\alpha|\alpha_k)\alpha_k.$$

The matrix A is defined by

$$T\alpha_j = \sum_{k=1}^{n} A_{kj}\alpha_k$$

and since

$$T\alpha_j = \sum_{k=1}^{n} (T\alpha_j|\alpha_k)\alpha_k$$

we have $A_{kj} = (T\alpha_j|\alpha_k)$. ∎

Corollary. *Let V be a finite-dimensional inner product space, and let T be a linear operator on V. In any orthonormal basis for V, the matrix of T* is the conjugate transpose of the matrix of T.*

Proof. Let $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$ be an orthonormal basis for V, let $A = [T]_{\mathcal{B}}$ and $B = [T^*]_{\mathcal{B}}$. According to Theorem 8,

$$A_{kj} = (T\alpha_j|\alpha_k)$$
$$B_{kj} = (T^*\alpha_j|\alpha_k).$$

By the definition of T^* we then have

$$\begin{aligned} B_{kj} &= (T^*\alpha_j|\alpha_k) \\ &= \overline{(\alpha_k|T^*\alpha_j)} \\ &= \overline{(T\alpha_k|\alpha_j)} \\ &= \overline{A_{jk}}. \quad \blacksquare \end{aligned}$$

Example 17. Let V be a finite-dimensional inner product space and E the orthogonal projection of V on a subspace W. Then for any vectors α and β in V.

$$\begin{aligned} (E\alpha|\beta) &= (E\alpha|E\beta + (1-E)\beta) \\ &= (E\alpha|E\beta) \\ &= (E\alpha + (1-E)\alpha|E\beta) \\ &= (\alpha|E\beta). \end{aligned}$$

From the uniqueness of the operator E^* it follows that $E^* = E$. Now consider the projection E described in Example 14. Then

$$A = \frac{1}{154}\begin{bmatrix} 9 & 36 & -3 \\ 36 & 144 & -12 \\ -3 & -12 & 1 \end{bmatrix}$$

is the matrix of E in the standard orthonormal basis. Since $E = E^*$, A is also the matrix of E^*, and because $A = A^*$, this does not contradict the preceding corollary. On the other hand, suppose

$$\begin{aligned} \alpha_1 &= (154, 0, 0) \\ \alpha_2 &= (145, -36, 3) \\ \alpha_3 &= (-36, 10, 12). \end{aligned}$$

Then $\{\alpha_1, \alpha_2, \alpha_3\}$ is a basis, and

$$\begin{aligned} E\alpha_1 &= (9, 36, -3) \\ E\alpha_2 &= (0, 0, 0) \\ E\alpha_3 &= (0, 0, 0). \end{aligned}$$

Since $(9, 36, -3) = -(154, 0, 0) - (145, -36, 3)$, the matrix B of E in the basis $\{\alpha_1, \alpha_2, \alpha_3\}$ is defined by the equation

$$B = \begin{bmatrix} -1 & 0 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

In this case $B \neq B^*$, and B^* is not the matrix of $E^* = E$ in the basis $\{\alpha_1, \alpha_2, \alpha_3\}$. Applying the corollary, we conclude that $\{\alpha_1, \alpha_2, \alpha_3\}$ is not an orthonormal basis. Of course this is quite obvious anyway.

Definition. *Let T be a linear operator on an inner product space V. Then we say that T* **has an adjoint on** *V if there exists a linear operator T* on V such that $(T\alpha|\beta) = (\alpha|T^*\beta)$ for all α and β in V.*

By Theorem 7 every linear operator on a finite-dimensional inner product space V has an adjoint on V. In the infinite-dimensional case this is not always true. But in any case there is at most one such operator T^*; when it exists, we call it the **adjoint** of T.

Two comments should be made about the finite-dimensional case.

1. The adjoint of T depends not only on T but on the inner product as well.
2. As shown by Example 17, in an arbitrary ordered basis $\mathcal{B}$, the relation between $[T]_{\mathcal{B}}$ and $[T^*]_{\mathcal{B}}$ is more complicated than that given in the corollary above.

Example 18. Let V be $C^{n\times 1}$, the space of complex $n \times 1$ matrices, with inner product $(X|Y) = Y^*X$. If A is an $n \times n$ matrix with complex entries, the adjoint of the linear operator $X \to AX$ is the operator $X \to A^*X$. For

$$(AX|Y) = Y^*AX = (A^*Y)^*X = (X|A^*Y).$$

The reader should convince himself that this is really a special case of the last corollary.

Example 19. This is similar to Example 18. Let V be $C^{n\times n}$ with the inner product $(A|B) = \operatorname{tr}(B^*A)$. Let M be a fixed $n \times n$ matrix over C. The adjoint of left multiplication by M is left multiplication by M^*. Of course, 'left multiplication by M' is the linear operator L_M defined by $L_M(A) = MA$.

$$\begin{aligned} (L_M(A)|B) &= \operatorname{tr}(B^*(MA)) \\ &= \operatorname{tr}(MAB^*) \\ &= \operatorname{tr}(AB^*M) \\ &= \operatorname{tr}(A(M^*B)^*) \\ &= (A|L_{M^*}(B)). \end{aligned}$$

Thus $(L_M)^* = L_{M^*}$. In the computation above, we twice used the characteristic property of the trace function: $\operatorname{tr}(AB) = \operatorname{tr}(BA)$.

Example 20. Let V be the space of polynomials over the field of complex numbers, with the inner product

$$(f|g) = \int_0^1 f(t)\overline{g(t)}\,dt.$$

If f is a polynomial, $f = \sum a_k x^k$, we let $\bar{f} = \sum \bar{a}_k x^k$. That is, $\bar{f}$ is the polynomial whose associated polynomial function is the complex conjugate of that for f:

$$\bar{f}(t) = \overline{f(t)}, \qquad t \text{ real}$$

Consider the operator 'multiplication by f,' that is, the linear operator M_f defined by $M_f(g) = fg$. Then this operator has an adjoint, namely, multiplication by $\bar{f}$. For

$$\begin{aligned}(M_f(g)|h) &= (fg|h)\\ &= \int_0^1 f(t)g(t)\overline{h(t)}\,dt\\ &= \int_0^1 g(t)\overline{[\overline{f(t)}h(t)]}\,dt\\ &= (g|\bar{f}h)\\ &= (g|M_{\bar{f}}(h))\end{aligned}$$

and so $(M_f)^* = M_{\bar{f}}$.

Example 21. In Example 20, we saw that some linear operators on an infinite-dimensional inner product space do have an adjoint. As we commented earlier, some do not. Let V be the inner product space of Example 21, and let D be the differentiation operator on $C[x]$. Integration by parts shows that

$$(Df|g) = f(1)g(1) - f(0)g(0) - (f|Dg).$$

Let us fix g and inquire when there is a polynomial D^*g such that $(Df|g) = (f|D^*g)$ for all f. If such a D^*g exists, we shall have

$$(f|D^*g) = f(1)g(1) - f(0)g(0) - (f|Dg)$$

or

$$(f|D^*g + Dg) = f(1)g(1) - f(0)g(0).$$

With g fixed, $L(f) = f(1)g(1) - f(0)g(0)$ is a linear functional of the type considered in Example 16 and cannot be of the form $L(f) = (f|h)$ unless $L = 0$. If D^*g exists, then with $h = D^*g + Dg$ we do have $L(f) = (f|h)$, and so $g(0) = g(1) = 0$. The existence of a suitable polynomial D^*g implies $g(0) = g(1) = 0$. Conversely, if $g(0) = g(1) = 0$, the polynomial $D^*g = -Dg$ satisfies $(Df|g) = (f|D^*g)$ for all f. If we choose any g for which $g(0) \neq 0$ or $g(1) \neq 0$, we cannot suitably define D^*g, and so we conclude that D has no adjoint.

We hope that these examples enhance the reader's understanding of the adjoint of a linear operator. We see that the adjoint operation, passing from T to T^*, behaves somewhat like conjugation on complex numbers. The following theorem strengthens the analogy.

Theorem 9. *Let V be a finite-dimensional inner product space. If T and U are linear operators on V and c is a scalar,*

(i) $(T + U)^* = T^* + U^*$;
(ii) $(cT)^* = \bar{c}T^*$;
(iii) $(TU)^* = U^*T^*$;
(iv) $(T^*)^* = T$.

Proof. To prove (i), let α and β be any vectors in V. Then

$$\begin{aligned}((T + U)\alpha|\beta) &= (T\alpha + U\alpha|\beta)\\ &= (T\alpha|\beta) + (U\alpha|\beta)\\ &= (\alpha|T^*\beta) + (\alpha|U^*\beta)\\ &= (\alpha|T^*\beta + U^*\beta)\\ &= (\alpha|(T^* + U^*)\beta)\end{aligned}$$

From the uniqueness of the adjoint we have $(T + U)^* = T^* + U^*$. We leave the proof of (ii) to the reader. We obtain (iii) and (iv) from the relations

$$(TU\alpha|\beta) = (U\alpha|T^*\beta) = (\alpha|U^*T^*\beta)$$

$$(T^*\alpha|\beta) = \overline{(\beta|T^*\alpha)} = \overline{(T\beta|\alpha)} = (\alpha|T\beta). \blacksquare$$

Theorem 9 is often phrased as follows: The mapping $T \to T^*$ is a conjugate-linear anti-isomorphism of period 2. The analogy with complex conjugation which we mentioned above is, of course, based upon the observation that complex conjugation has the properties $\overline{(z_1 + z_2)} = \bar{z}_1 + \bar{z}_2$, $\overline{(z_1 z_2)} = \bar{z}_1\bar{z}_2$, $\bar{\bar{z}} = z$. One must be careful to observe the reversal of order in a product, which the adjoint operation imposes: $(UT)^* = T^*U^*$. We shall mention extensions of this analogy as we continue our study of linear operators on an inner product space. We might mention something along these lines now. A complex number z is real if and only if $z = \bar{z}$. One might expect that the linear operators T such that $T = T^*$ behave in some way like the real numbers. This is in fact the case. For example, if T is a linear operator on a finite-dimensional *complex* inner product space, then

$$T = U_1 + iU_2 \tag{8-15}$$

where $U_1 = U_1^*$ and $U_2 = U_2^*$. Thus, in some sense, T has a 'real part' and an 'imaginary part.' The operators U_1 and U_2 satisfying $U_1 = U_1^*$, and $U_2 = U_2^*$, and (8-15) are unique, and are given by

$$U_1 = \frac{1}{2}(T + T^*)$$

$$U_2 = \frac{1}{2i}(T - T^*).$$

A linear operator T such that $T = T^*$ is called **self-adjoint** (or **Hermitian**). If $\mathfrak{B}$ is an orthonormal basis for V, then

$$[T^*]_{\mathfrak{B}} = [T]^*_{\mathfrak{B}}$$

and so T is self-adjoint if and only if its matrix in every orthonormal basis is a self-adjoint matrix. Self-adjoint operators are important, not simply because they provide us with some sort of real and imaginary part for the general linear operator, but for the following reasons: (1) Self-adjoint operators have many special properties. For example, for such an operator there is an orthonormal basis of characteristic vectors. (2) Many operators which arise in practice are self-adjoint. We shall consider the special properties of self-adjoint operators later.

Exercises

1. Let V be the space C^2, with the standard inner product. Let T be the linear operator defined by $T\epsilon_1 = (1, -2)$, $T\epsilon_2 = (i, -1)$. If $\alpha = (x_1, x_2)$, find $T^*\alpha$.

2. Let T be the linear operator on C^2 defined by $T\epsilon_1 = (1 + i, 2)$, $T\epsilon_2 = (i, i)$. Using the standard inner product, find the matrix of T^* in the standard ordered basis. Does T commute with T^*?

3. Let V be C^3 with the standard inner product. Let T be the linear operator on V whose matrix in the standard ordered basis is defined by

$$A_{jk} = i^{j+k}, \qquad (i^2 = -1).$$

Find a basis for the null space of T^*.

4. Let V be a finite-dimensional inner product space and T a linear operator on V. Show that the range of T^* is the orthogonal complement of the null space of T.

5. Let V be a finite-dimensional inner product space and T a linear operator on V. If T is invertible, show that T^* is invertible and $(T^*)^{-1} = (T^{-1})^*$.

6. Let V be an inner product space and β, γ fixed vectors in V. Show that $T\alpha = (\alpha|\beta)\gamma$ defines a linear operator on V. Show that T has an adjoint, and describe T^* explicitly.

Now suppose V is C^n with the standard inner product, $\beta = (y_1, \ldots, y_n)$, and $\gamma = (x_1, \ldots, x_n)$. What is the j, k entry of the matrix of T in the standard ordered basis? What is the rank of this matrix?

7. Show that the product of two self-adjoint operators is self-adjoint if and only if the two operators commute.

8. Let V be the vector space of the polynomials over R of degree less than or equal to 3, with the inner product

$$(f|g) = \int_0^1 f(t)g(t)\,dt.$$

If t is a real number, find the polynomial g_t in V such that $(f|g_t) = f(t)$ for all f in V.

9. Let V be the inner product space of Exercise 8, and let D be the differentiation operator on V. Find D^*.

10. Let V be the space of $n \times n$ matrices over the complex numbers, with the inner product $(A, B) = \operatorname{tr}(AB^*)$. Let P be a fixed invertible matrix in V, and let T_P be the linear operator on V defined by $T_P(A) = P^{-1}AP$. Find the adjoint of T_P.

11. Let V be a finite-dimensional inner product space, and let E be an idempotent linear operator on V, i.e., $E^2 = E$. Prove that E is self-adjoint if and only if $EE^* = E^*E$.

12. Let V be a finite-dimensional *complex* inner product space, and let T be a linear operator on V. Prove that T is self-adjoint if and only if $(T\alpha|\alpha)$ is real for every α in V.

8.4. Unitary Operators

In this section, we consider the concept of an isomorphism between two inner product spaces. If V and W are vector spaces, an isomorphism of V onto W is a one-one linear transformation from V onto W, i.e., a one-one correspondence between the elements of V and those of W, which 'preserves' the vector space operations. Now an inner product space consists of a vector space and a specified inner product on that space. Thus, when V and W are inner product spaces, we shall require an isomorphism from V onto W not only to preserve the linear operations, but also to preserve inner products. An isomorphism of an inner product space onto itself is called a 'unitary operator' on that space. We shall consider various examples of unitary operators and establish their basic properties.

Definition. *Let V and W be inner product spaces over the same field, and let T be a linear transformation from V into W. We say that T* **preserves inner products** *if $(T\alpha|T\beta) = (\alpha|\beta)$ for all α, β in V. An* **isomorphism** *of V onto W is a vector space isomorphism T of V onto W which also preserves inner products.*

If T preserves inner products, then $||T\alpha|| = ||\alpha||$ and so T is necessarily non-singular. Thus an isomorphism from V onto W can also be defined as a linear transformation from V onto W which preserves inner products. If T is an isomorphism of V onto W, then T^{-1} is an isomorphism of W onto V; hence, when such a T exists, we shall simply say V and W are **isomorphic.** Of course, isomorphism of inner product spaces is an equivalence relation.

Theorem 10. *Let V and W be finite-dimensional inner product spaces over the same field, having the same dimension. If T is a linear transformation from V into W, the following are equivalent.*

(i) T *preserves inner products.*

(ii) T *is an (inner product space) isomorphism.*

(iii) T *carries every orthonormal basis for* V *onto an orthonormal basis for* W.

(iv) T *carries some orthonormal basis for* V *onto an orthonormal basis for* W.

Proof. (i) $\rightarrow$ (ii) If T preserves inner products, then $||T\alpha|| = ||\alpha||$ for all α in V. Thus T is non-singular, and since $\dim V = \dim W$, we know that T is a vector space isomorphism.

(ii) $\rightarrow$ (iii) Suppose T is an isomorphism. Let $\{\alpha_1, \ldots, \alpha_n\}$ be an orthonormal basis for V. Since T is a vector space isomorphism and $\dim W = \dim V$, it follows that $\{T\alpha_1, \ldots, T\alpha_n\}$ is a basis for W. Since T also preserves inner products, $(T\alpha_j|T\alpha_k) = (\alpha_j|\alpha_k) = \delta_{jk}$.

(iii) $\rightarrow$ (iv) This requires no comment.

(iv) $\rightarrow$ (i) Let $\{\alpha_1, \ldots, \alpha_n\}$ be an orthonormal basis for V such that $\{T\alpha_1, \ldots, T\alpha_n\}$ is an orthonormal basis for W. Then

$$(T\alpha_j|T\alpha_k) = (\alpha_j|\alpha_k) = \delta_{jk}.$$

For any $\alpha = x_1\alpha_1 + \cdots + x_n\alpha_n$ and $\beta = y_1\alpha_1 + \cdots + y_n\alpha_n$ in V, we have

$$\begin{aligned}(\alpha|\beta) &= \sum_{j=1}^{n} x_j\bar{y}_j \\ (T\alpha|T\beta) &= (\sum_j x_jT\alpha_j|\sum_k y_kT\alpha_k) \\ &= \sum_j \sum_k x_j\bar{y}_k(T\alpha_j|T\alpha_k) \\ &= \sum_{j=1}^{n} x_j\bar{y}_j\end{aligned}$$

and so T preserves inner products. ∎

Corollary. *Let* V *and* W *be finite-dimensional inner product spaces over the same field. Then* V *and* W *are isomorphic if and only if they have the same dimension.*

Proof. If $\{\alpha_1, \ldots, \alpha_n\}$ is an orthonormal basis for V and $\{\beta_1, \ldots, \beta_n\}$ is an orthonormal basis for W, let T be the linear transformation from V into W defined by $T\alpha_j = \beta_j$. Then T is an isomorphism of V onto W. ∎

EXAMPLE 22. If V is an n-dimensional inner product space, then each ordered orthonormal basis $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$ determines an isomorphism of V onto F^n with the standard inner product. The isomorphism is simply

$$T(x_1\alpha_1 + \cdots + x_n\alpha_n) = (x_1, \ldots, x_n).$$

There is the superficially different isomorphism which $\mathcal{B}$ determines of V onto the space $F^{n\times 1}$ with $(X|Y) = Y^*X$ as inner product. The isomorphism is

$$\alpha \rightarrow [\alpha]_{\mathcal{B}}$$

i.e., the transformation sending α into its coordinate matrix in the ordered basis $\mathcal{B}$. For any ordered basis $\mathcal{B}$, this is a vector space isomorphism; however, it is an isomorphism of the two inner product spaces if and only if $\mathcal{B}$ is orthonormal.

EXAMPLE 23. Here is a slightly less superficial isomorphism. Let W be the space of all 3×3 matrices A over R which are skew-symmetric, i.e., $A^t = -A$. We equip W with the inner product $(A|B) = \frac{1}{2}\operatorname{tr}(AB^t)$, the $\frac{1}{2}$ being put in as a matter of convenience. Let V be the space R^3 with the standard inner product. Let T be the linear transformation from V into W defined by

$$T(x_1, x_2, x_3) = \begin{bmatrix} 0 & -x_3 & x_2 \\ x_3 & 0 & -x_1 \\ -x_2 & x_1 & 0 \end{bmatrix}.$$

Then T maps V onto W, and putting

$$A = \begin{bmatrix} 0 & -x_3 & x_2 \\ x_3 & 0 & -x_1 \\ -x_2 & x_1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & -y_3 & y_2 \\ y_3 & 0 & -y_1 \\ -y_2 & y_1 & 0 \end{bmatrix}$$

we have

$$\begin{aligned}\operatorname{tr}(AB^t) &= x_3y_3 + x_2y_2 + x_3y_3 + x_2y_2 + x_1y_1 \\ &= 2(x_1y_1 + x_2y_2 + x_3y_3).\end{aligned}$$

Thus $(\alpha|\beta) = (T\alpha|T\beta)$ and T is a vector space isomorphism. Note that T carries the standard basis $\{\epsilon_1, \epsilon_2, \epsilon_3\}$ onto the orthonormal basis consisting of the three matrices

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

EXAMPLE 24. It is not always particularly convenient to describe an isomorphism in terms of orthonormal bases. For example, suppose $G = P^*P$ where P is an invertible $n \times n$ matrix with complex entries. Let V be the space of complex $n \times 1$ matrices, with the inner product $[X|Y] = Y^*GX$.

Let W be the same vector space, with the standard inner product $(X|Y) = Y^*X$. We know that V and W are isomorphic inner product spaces. It would seem that the most convenient way to describe an isomorphism between V and W is the following: Let T be the linear transformation from V into W defined by $T(X) = PX$. Then

$$\begin{aligned}(TX|TY) &= (PX|PY)\\ &= (PY)^*(PX)\\ &= Y^*P^*PX\\ &= Y^*GX\\ &= [X|Y].\end{aligned}$$

Hence T is an isomorphism.

EXAMPLE 25. Let V be the space of all continuous real-valued functions on the unit interval, $0 \le t \le 1$, with the inner product

$$[f|g] = \int_0^1 f(t)g(t)t^2\,dt.$$

Let W be the same vector space with the inner product

$$(f|g) = \int_0^1 f(t)g(t)\,dt.$$

Let T be the linear transformation from V into W given by

$$(Tf)(t) = tf(t).$$

Then $(Tf|Tg) = [f|g]$, and so T preserves inner products; however, T is *not* an isomorphism of V onto W, because the range of T is not all of W. Of course, this happens because the underlying vector space is not finite-dimensional.

Theorem 11. *Let V and W be inner product spaces over the same field, and let T be a linear transformation from V into W. Then T preserves inner products if and only if $||T\alpha|| = ||\alpha||$ for every α in V.*

Proof. If T preserves inner products, T 'preserves norms.' Suppose $||T\alpha|| = ||\alpha||$ for every α in V. Then $||T\alpha||^2 = ||\alpha||^2$. Now using the appropriate polarization identity, (8-3) or (8-4), and the fact that T is linear, one easily obtains $(\alpha|\beta) = (T\alpha|T\beta)$ for all α, β in V. ▌

Definition. *A* **unitary operator** *on an inner product space is an isomorphism of the space onto itself.*

The product of two unitary operators is unitary. For, if U_1 and U_2 are unitary, then U_2U_1 is invertible and $||U_2U_1\alpha|| = ||U_1\alpha|| = ||\alpha||$ for each α. Also, the inverse of a unitary operator is unitary, since $||U\alpha|| = ||\alpha||$ says $||U^{-1}\beta|| = ||\beta||$, where $\beta = U\alpha$. Since the identity operator is clearly unitary, we see that the set of all unitary operators on an inner product space is a group, under the operation of composition.

If V is a finite-dimensional inner product space and U is a linear operator on V, Theorem 10 tells us that U is unitary if and only if $(U\alpha|U\beta) = (\alpha|\beta)$ for each α, β in V; or, if and only if for some (every) orthonormal basis $\{\alpha_1, \ldots, \alpha_n\}$ it is true that $\{U\alpha_1, \ldots, U\alpha_n\}$ is an orthonormal basis.

Theorem 12. *Let U be a linear operator on an inner product space V. Then U is unitary if and only if the adjoint U^* of U exists and $UU^* = U^*U = I$.*

Proof. Suppose U is unitary. Then U is invertible and

$$(U\alpha|\beta) = (U\alpha|UU^{-1}\beta) = (\alpha|U^{-1}\beta)$$

for all α, β. Hence U^{-1} is the adjoint of U.

Conversely, suppose U^* exists and $UU^* = U^*U = I$. Then U is invertible, with $U^{-1} = U^*$. So, we need only show that U preserves inner products. We have

$$\begin{aligned}(U\alpha|U\beta) &= (\alpha|U^*U\beta)\\ &= (\alpha|I\beta)\\ &= (\alpha|\beta)\end{aligned}$$

for all α, β. ▌

EXAMPLE 26. Consider $C^{n\times 1}$ with the inner product $(X|Y) = Y^*X$. Let A be an $n \times n$ matrix over C, and let U be the linear operator defined by $U(X) = AX$. Then

$$(UX|UY) = (AX|AY) = Y^*A^*AX$$

for all X, Y. Hence, U is unitary if and only if $A^*A = I$.

Definition. *A complex $n \times n$ matrix A is called* **unitary**, *if $A^*A = I$.*

Theorem 13. *Let V be a finite-dimensional inner product space and let U be a linear operator on V. Then U is unitary if and only if the matrix of U in some (or every) ordered orthonormal basis is a unitary matrix.*

Proof. At this point, this is not much of a theorem, and we state it largely for emphasis. If $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$ is an ordered orthonormal basis for V and A is the matrix of U relative to $\mathcal{B}$, then $A^*A = I$ if and only if $U^*U = I$. The result now follows from Theorem 12. ▌

Let A be an $n \times n$ matrix. The statement that A is unitary simply means

$$(A^*A)_{jk} = \delta_{jk}$$

or

$$\sum_{r=1}^{n} \overline{A_{rj}} A_{rk} = \delta_{jk}.$$

In other words, it means that the columns of A form an orthonormal set of column matrices, with respect to the standard inner product $(X|Y) = Y^*X$. Since $A^*A = I$ if and only if $AA^* = I$, we see that A is unitary exactly when the rows of A comprise an orthonormal set of n-tuples in C_n (with the standard inner product). So, using standard inner products, A is unitary if and only if the rows and columns of A are orthonormal sets. One sees here an example of the power of the theorem which states that a one-sided inverse for a matrix is a two-sided inverse. Applying this theorem as we did above, say to real matrices, we have the following: Suppose we have a square array of real numbers such that the sum of the squares of the entries in each row is 1 and distinct rows are orthogonal. Then the sum of the squares of the entries in each column is 1 and distinct columns are orthogonal. Write down the proof of this for a 3×3 array, without using any knowledge of matrices, and you should be reasonably impressed.

Definition. *A real or complex* $n \times n$ *matrix* A *is said to be* **orthogonal**, *if* $A^tA = I$.

A real orthogonal matrix is unitary; and, a unitary matrix is orthogonal if and only if each of its entries is real.

EXAMPLE 27. We give some examples of unitary and orthogonal matrices.

(a) A 1×1 matrix $[c]$ is orthogonal if and only if $c = \pm 1$, and unitary if and only if $\bar{c}c = 1$. The latter condition means (of course) that $|c| = 1$, or $c = e^{i\theta}$, where θ is real.

(b) Let

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

Then A is orthogonal if and only if

$$A^t = A^{-1} = \frac{1}{ad - bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

The determinant of any orthogonal matrix is easily seen to be ± 1. Thus A is orthogonal if and only if

$$A = \begin{bmatrix} a & b \\ -b & a \end{bmatrix}$$

or

$$A = \begin{bmatrix} a & b \\ b & -a \end{bmatrix}$$

where $a^2 + b^2 = 1$. The two cases are distinguished by the value of det A.

(c) The well-known relations between the trigonometric functions show that the matrix

$$A_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

is orthogonal. If θ is a real number, then A_θ is the matrix in the standard ordered basis for R^2 of the linear operator U_θ, rotation through the angle θ. The statement that A_θ is a real orthogonal matrix (hence unitary) simply means that U_θ is a unitary operator, i.e., preserves dot products.

(d) Let

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

Then A is unitary if and only if

$$\begin{bmatrix} \bar{a} & \bar{c} \\ \bar{b} & \bar{d} \end{bmatrix} = \frac{1}{ad - bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

The determinant of a unitary matrix has absolute value 1, and is thus a complex number of the form $e^{i\theta}$, θ real. Thus A is unitary if and only if

$$A = \begin{bmatrix} a & b \\ -e^{i\theta}\bar{b} & e^{i\theta}\bar{a} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & e^{i\theta} \end{bmatrix}\begin{bmatrix} a & b \\ -\bar{b} & \bar{a} \end{bmatrix}$$

where θ is a real number, and a, b are complex numbers such that $|a|^2 + |b|^2 = 1$.

As noted earlier, the unitary operators on an inner product space form a group. From this and Theorem 13 it follows that the set $U(n)$ of all $n \times n$ unitary matrices is also a group. Thus the inverse of a unitary matrix and the product of two unitary matrices are again unitary. Of course this is easy to see directly. An $n \times n$ matrix A with complex entries is unitary if and only if $A^{-1} = A^*$. Thus, if A is unitary, we have $(A^{-1})^{-1} = A = (A^*)^{-1} = (A^{-1})^*$. If A and B are $n \times n$ unitary matrices, then $(AB)^{-1} = B^{-1}A^{-1} = B^*A^* = (AB)^*$.

The Gram-Schmidt process in C^n has an interesting corollary for matrices that involves the group $U(n)$.

Theorem 14. *For every invertible complex* $n \times n$ *matrix* B *there exists a unique lower-triangular matrix* M *with positive entries on the main diagonal such that* MB *is unitary.*

Proof. The rows $\beta_1, \ldots, \beta_n$ of B form a basis for C^n. Let $\alpha_1, \ldots, \alpha_n$ be the vectors obtained from $\beta_1, \ldots, \beta_n$ by the Gram-Schmidt process. Then, for $1 \leq k \leq n$, $\{\alpha_1, \ldots, \alpha_k\}$ is an orthogonal basis for the subspace spanned by $\{\beta_1, \ldots, \beta_k\}$, and

$$\alpha_k = \beta_k - \sum_{j<k} \frac{(\beta_k|\alpha_j)}{\|\alpha_j\|^2} \alpha_j.$$

Hence, for each k there exist unique scalars C_{kj} such that

$$\alpha_k = \beta_k - \sum_{j<k} C_{kj}\beta_j.$$

Let U be the unitary matrix with rows

$$\frac{\alpha_1}{\|\alpha_1\|}, \ldots, \frac{\alpha_n}{\|\alpha_n\|}$$

and M the matrix defined by

$$M_{kj} = \begin{cases} -\dfrac{1}{\|\alpha_k\|} \cdot C_{kj}, & \text{if } j < k \\ \dfrac{1}{\|\alpha_k\|}, & \text{if } j = k \\ 0, & \text{if } j > k. \end{cases}$$

Then M is lower-triangular, in the sense that its entries above the main diagonal are 0. The entries M_{kk} of M on the main diagonal are all > 0, and

$$\frac{\alpha_k}{\|\alpha_k\|} = \sum_{j=1}^{n} M_{kj}\beta_j, \qquad 1 \leq k \leq n.$$

Now these equations simply say that

$$U = MB.$$

To prove the uniqueness of M, let $T^+(n)$ denote the set of all complex $n \times n$ lower-triangular matrices with positive entries on the main diagonal. Suppose M_1 and M_2 are elements of $T^+(n)$ such that M_iB is in $U(n)$ for $i = 1, 2$. Then because $U(n)$ is a group

$$(M_1B)(M_2B)^{-1} = M_1M_2^{-1}$$

lies in $U(n)$. On the other hand, although it is not entirely obvious, $T^+(n)$ is also a group under matrix multiplication. One way to see this is to consider the geometric properties of the linear transformations

$$X \to MX, \qquad (M \text{ in } T^+(n))$$

on the space of column matrices. Thus M_2^{-1}, $M_1M_2^{-1}$, and $(M_1M_2^{-1})^{-1}$ are all in $T^+(n)$. But, since $M_1M_2^{-1}$ is in $U(n)$, $(M_1M_2^{-1})^{-1} = (M_1M_2^{-1})^*$. The transpose or conjugate transpose of any lower-triangular matrix is an upper-triangular matrix. Therefore, $M_1M_2^{-1}$ is simultaneously upper- and lower-triangular, i.e., diagonal. A diagonal matrix is unitary if and only if each of its entries on the main diagonal has absolute value 1; if the diagonal entries are all positive, they must equal 1. Hence $M_1M_2^{-1} = I$ and $M_1 = M_2$. ∎

Let $GL(n)$ denote the set of all invertible complex $n \times n$ matrices. Then $GL(n)$ is also a group under matrix multiplication. This group is called the **general linear group.** Theorem 14 is equivalent to the following result.

Corollary. *For each* B *in* GL(n) *there exist unique matrices* N *and* U *such that* N *is in* T⁺(n), U *is in* U(n), *and*

$$\mathrm{B} = \mathrm{N} \cdot \mathrm{U}.$$

Proof. By the theorem there is a unique matrix M in $T^+(n)$ such that MB is in $U(n)$. Let $MB = U$ and $N = M^{-1}$. Then N is in $T^+(n)$ and $B = N \cdot U$. On the other hand, if we are given any elements N and U such that N is in $T^+(n)$, U is in $U(n)$, and $B = N \cdot U$, then $N^{-1}B$ is in $U(n)$ and N^{-1} is the unique matrix M which is characterized by the theorem; furthermore U is necessarily $N^{-1}B$. ∎

EXAMPLE 28. Let x_1 and x_2 be real numbers such that $x_1^2 + x_2^2 = 1$ and $x_1 \neq 0$. Let

$$B = \begin{bmatrix} x_1 & x_2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Applying the Gram-Schmidt process to the rows of B, we obtain the vectors

$$\begin{aligned} \alpha_1 &= (x_1, x_2, 0) \\ \alpha_2 &= (0, 1, 0) - x_2(x_1, x_2, 0) \\ &= x_1(-x_2, x_1, 0) \\ \alpha_3 &= (0, 0, 1). \end{aligned}$$

Let U be the matrix with rows α_1, (α_2/x_1), α_3. Then U is unitary, and

$$U = \begin{bmatrix} x_1 & x_2 & 0 \\ -x_2 & x_1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -\dfrac{x_2}{x_1} & \dfrac{1}{x_1} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 & x_2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Now multiplying by the inverse of

$$M = \begin{bmatrix} 1 & 0 & 0 \\ -\dfrac{x_2}{x_1} & \dfrac{1}{x_1} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

we find that

$$\begin{bmatrix} x_1 & x_2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ x_2 & x_1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 & x_2 & 0 \\ -x_2 & x_1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Let us now consider briefly change of coordinates in an inner product space. Suppose V is a finite-dimensional inner product space and that $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$ and $\mathcal{B}' = \{\alpha_1', \ldots, \alpha_n'\}$ are two ordered *orthonormal* bases for V. There is a unique (necessarily invertible) $n \times n$ matrix P such that

$$[\alpha]_{\mathcal{B}'} = P^{-1}[\alpha]_{\mathcal{B}}$$

for every α in V. If U is the unique linear operator on V defined by $U\alpha_j = \alpha_j'$, then P is the matrix of U in the ordered basis $\mathcal{B}$:

$$\alpha_k' = \sum_{j=1}^{n} P_{jk}\alpha_j.$$

Since $\mathcal{B}$ and $\mathcal{B}'$ are orthonormal bases, U is a unitary operator and P is a unitary matrix. If T is any linear operator on V, then

$$[T]_{\mathcal{B}'} = P^{-1}[T]_{\mathcal{B}}P = P^*[T]_{\mathcal{B}}P.$$

Definition. *Let* A *and* B *be complex* n $\times$ n *matrices. We say that* B **is unitarily equivalent to** A *if there is an* n $\times$ n *unitary matrix* P *such that* B = P^{-1}AP. *We say that* B *is* **orthogonally equivalent** *to* A *if there is an* n $\times$ n *orthogonal matrix* P *such that* B = P^{-1}AP.

With this definition, what we observed above may be stated as follows: If $\mathcal{B}$ and $\mathcal{B}'$ are two ordered orthonormal bases for V, then, for each linear operator T on V, the matrix $[T]_{\mathcal{B}'}$ is unitarily equivalent to the matrix $[T]_{\mathcal{B}}$. In case V is a real inner product space, these matrices are orthogonally equivalent, via a real orthogonal matrix.

Exercises

1. Find a unitary matrix which is not orthogonal, and find an orthogonal matrix which is not unitary.

2. Let V be the space of complex $n \times n$ matrices with inner product $(A|B) = \operatorname{tr}(AB^*)$. For each M in V, let T_M be the linear operator defined by $T_M(A) = MA$. Show that T_M is unitary if and only if M is a unitary matrix.

3. Let V be the set of complex numbers, regarded as a *real* vector space.

(a) Show that $(\alpha|\beta) = \operatorname{Re}(\alpha\bar{\beta})$ defines an inner product on V.

(b) Exhibit an (inner product space) isomorphism of V onto R^2 with the standard inner product.

(c) For each γ in V, let M_γ be the linear operator on V defined by $M_\gamma(\alpha) = \gamma\alpha$. Show that $(M_\gamma)^* = M_{\bar{\gamma}}$.

(d) For which complex numbers γ is M_γ self-adjoint?

(e) For which γ is M_γ unitary?

(f) For which γ is M_γ positive?

(g) What is $\det(M_\gamma)$?

(h) Find the matrix of M_γ in the basis $\{1, i\}$.

(i) If T is a linear operator on V, find necessary and sufficient conditions for T to be an M_γ.

(j) Find a unitary operator on V which is not an M_γ.

4. Let V be R^2, with the standard inner product. If U is a unitary operator on V, show that the matrix of U in the standard ordered basis is either

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix}$$

for some real θ, $0 \leq \theta < 2\pi$. Let U_θ be the linear operator corresponding to the first matrix, i.e., U_θ is rotation through the angle θ. Now convince yourself that every unitary operator on V is either a rotation, or reflection about the ϵ_1-axis followed by a rotation.

(a) What is $U_\theta U_\phi$?

(b) Show that $U_\theta^* = U_{-\theta}$.

(c) Let ϕ be a fixed real number, and let $\mathcal{B} = \{\alpha_1, \alpha_2\}$ be the orthonormal basis obtained by rotating $\{\epsilon_1, \epsilon_2\}$ through the angle ϕ, i.e., $\alpha_j = U_\phi\epsilon_j$. If θ is another real number, what is the matrix of U_θ in the ordered basis $\mathcal{B}$?

5. Let V be R^3, with the standard inner product. Let W be the plane spanned by $\alpha = (1, 1, 1)$ and $\beta = (1, 1, -2)$. Let U be the linear operator defined, geometrically, as follows: U is rotation through the angle θ, about the straight line through the origin which is orthogonal to W. There are actually two such rotations —choose one. Find the matrix of U in the standard ordered basis. (Here is one way you might proceed. Find α_1 and α_2 which form an orthonormal basis for W. Let α_3 be a vector of norm 1 which is orthogonal to W. Find the matrix of U in the basis $\{\alpha_1, \alpha_2, \alpha_3\}$. Perform a change of basis.)

6. Let V be a finite-dimensional inner product space, and let W be a subspace of V. Then $V = W \oplus W^\perp$, that is, each α in V is uniquely expressible in the form $\alpha = \beta + \gamma$, with β in W and γ in $W^\perp$. Define a linear operator U by $U\alpha = \beta - \gamma$.

(a) Prove that U is both self-adjoint and unitary.

(b) If V is R^3 with the standard inner product and W is the subspace spanned by $(1, 0, 1)$, find the matrix of U in the standard ordered basis.

7. Let V be a *complex* inner product space and T a *self-adjoint* linear operator on V. Show that

(a) $\|\alpha + iT\alpha\| = \|\alpha - iT\alpha\|$ for every α in V.

(b) $\alpha + iT\alpha = \beta + iT\beta$ if and only if $\alpha = \beta$.

(c) $I + iT$ is non-singular.

(d) $I - iT$ is non-singular.

(e) Now suppose V is finite-dimensional, and prove that

$$U = (I - iT)(I + iT)^{-1}$$

is a unitary operator; U is called the **Cayley transform** of T. In a certain sense, $U = f(T)$, where $f(x) = (1 - ix)/(1 + ix)$.

8. If θ is a real number, prove that the following matrices are unitarily equivalent

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}, \quad \begin{bmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{bmatrix}.$$

9. Let V be a finite-dimensional inner product space and T a positive linear operator on V. Let p_T be the inner product on V defined by $p_T(\alpha, \beta) = (T\alpha|\beta)$. Let U be a linear operator on V and U^* its adjoint with respect to $(\ |\ )$. Prove that U is unitary with respect to the inner product p_T if and only if $T = U^*TU$.

10. Let V be a finite-dimensional inner product space. For each α, β in V, let $T_{\alpha,\beta}$ be the linear operator on V defined by $T_{\alpha,\beta}(\gamma) = (\gamma|\beta)\alpha$. Show that

(a) $T^*_{\alpha,\beta} = T_{\beta,\alpha}$.
(b) trace $(T_{\alpha,\beta}) = (\alpha|\beta)$.
(c) $T_{\alpha,\beta}T_{\gamma,\delta} = T_{\alpha,(\beta|\gamma)\delta}$.
(d) Under what conditions is $T_{\alpha,\beta}$ self-adjoint?

11. Let V be an n-dimensional inner product space over the field F, and let $L(V, V)$ be the space of linear operators on V. Show that there is a unique inner product on $L(V, V)$ with the property that $||T_{\alpha,\beta}||^2 = ||\alpha||^2||\beta||^2$ for all α, β in V. ($T_{\alpha,\beta}$ is the operator defined in Exercise 10.) Find an isomorphism between $L(V, V)$ with this inner product and the space of $n \times n$ matrices over F, with the inner product $(A|B) = \operatorname{tr}(AB^*)$.

12. Let V be a finite-dimensional inner product space. In Exercise 6, we showed how to construct some linear operators on V which are both self-adjoint and unitary. Now prove that there are no others, i.e., that every self-adjoint unitary operator arises from some subspace W as we described in Exercise 6.

13. Let V and W be finite-dimensional inner product spaces having the same dimension. Let U be an isomorphism of V onto W. Show that:

(a) The mapping $T \to UTU^{-1}$ is an isomorphism of the vector space $L(V, V)$ onto the vector space $L(W, W)$.
(b) trace (UTU^{-1}) = trace (T) for each T in $L(V, V)$.
(c) $UT_{\alpha,\beta}U^{-1} = T_{U\alpha,U\beta}$ ($T_{\alpha,\beta}$ defined in Exercise 10).
(d) $(UTU^{-1})^* = UT^*U^{-1}$.
(e) If we equip $L(V, V)$ with inner product $(T_1|T_2) = \operatorname{trace}(T_1T_2^*)$, and similarly for $L(W, W)$, then $T \to UTU^{-1}$ is an inner product space isomorphism.

14. If V is an inner product space, a **rigid motion** is any function T from V into V (not necessarily linear) such that $||T\alpha - T\beta|| = ||\alpha - \beta||$ for all α, β in V. One example of a rigid motion is a linear unitary operator. Another example is translation by a fixed vector γ:

$$T_\gamma(\alpha) = \alpha + \gamma$$

(a) Let V be R^2 with the standard inner product. Suppose T is a rigid motion of V and that $T(0) = 0$. Prove that T is linear and a unitary operator.
(b) Use the result of part (a) to prove that every rigid motion of R^2 is composed of a translation, followed by a unitary operator.
(c) Now show that a rigid motion of R^2 is either a translation followed by a rotation, or a translation followed by a reflection followed by a rotation.

15. A unitary operator on R^4 (with the standard inner product) is simply a linear operator which preserves the quadratic form

$$||(x, y, z, t)||^2 = x^2 + y^2 + z^2 + t^2$$

that is, a linear operator U such that $||U\alpha||^2 = ||\alpha||^2$ for all α in R^4. In a certain part of the theory of relativity, it is of interest to find the linear operators T which preserve the form

$$||(x, y, z, t)||^2_L = t^2 - x^2 - y^2 - z^2.$$

Now $||\ \ ||^2_L$ does not come from an inner product, but from something called the 'Lorentz metric' (which we shall not go into). For that reason, a linear operator T on R^4 such that $||T\alpha||^2_L = ||\alpha||^2_L$, for every α in R^4, is called a **Lorentz transformation.**

(a) Show that the function U defined by

$$U(x, y, z, t) = \begin{bmatrix} t+x & y+iz \\ y-iz & t-x \end{bmatrix}$$

is an isomorphism of R^4 onto the real vector space H of all self-adjoint 2×2 complex matrices.
(b) Show that $||\alpha||^2_L = \det(U\alpha)$.
(c) Suppose T is a (real) linear operator on the space H of 2×2 self-adjoint matrices. Show that $L = U^{-1}TU$ is a linear operator on R^4.
(d) Let M be any 2×2 complex matrix. Show that $T_M(A) = M^*AM$ defines a linear operator T_M on H. (Be sure you check that T_M maps H into H.)
(e) If M is a 2×2 matrix such that $|\det M| = 1$, show that $L_M = U^{-1}T_MU$ is a Lorentz transformation on R^4.
(f) Find a Lorentz transformation which is not an L_M.

8.5. Normal Operators

The principal objective in this section is the solution of the following problem. If T is a linear operator on a finite-dimensional inner product space V, under what conditions does V have an orthonormal basis consisting of characteristic vectors for T? In other words, when is there an *orthonormal* basis $\mathfrak{B}$ for V, such that the matrix of T in the basis $\mathfrak{B}$ is diagonal?

We shall begin by deriving some necessary conditions on T, which we shall subsequently show are sufficient. Suppose $\mathfrak{B} = \{\alpha_1, \ldots, \alpha_n\}$ is an orthonormal basis for V with the property

$$T\alpha_j = c_j\alpha_j, \quad j = 1, \ldots, n. \tag{8-16}$$

This simply says that the matrix of T in the ordered basis $\mathfrak{B}$ is the diagonal matrix with diagonal entries $c_1, \ldots, c_n$. The adjoint operator T^* is represented in this same ordered basis by the conjugate transpose matrix, i.e., the diagonal matrix with diagonal entries $\bar{c}_1, \ldots, \bar{c}_n$. If V is a real inner product space, the scalars $c_1, \ldots, c_n$ are (of course) real, and so it must be that $T = T^*$ .In other words, if V is a finite-dimensional *real* inner probut space and T is a linear operator for which there is an orthonormal basis of characteristic vectos, then T must be self-adjoint. If V is a complex inner product space, the scalars $c_1, \ldots, c_n$ need not be real, i.e., T need not be self-adjoint. But notice that T must satisfy

$$TT^* = T^*T. \tag{8-17}$$

For, any two diagonal matrices commute, and since T and T^* are both represented by diagonal matrices in the ordered basis $\mathcal{B}$, we have (8-17). It is a rather remarkable fact that in the complex case this condition is also sufficient to imply the existence of an orthonormal basis of characteristic vectors.

Definition. *Let* V *be a finite-dimensional inner product space and* T *a linear operator on* V*. We say that* T *is* **normal** *if it commutes with its adjoint i.e.,* TT* = T*T.

Any self-adjoint operator is normal, as is any unitary operator. Any scalar multiple of a normal operator is normal; however, sums and products of normal operators are not generally normal. Although it is by no means necessary, we shall begin our study of normal operators by considering self-adjoint operators.

Theorem 15. *Let* V *be an inner product space and* T *a self-adjoint linear operator on* V*. Then each characteristic value of* T *is real, and characteristic vectors of* T *associated with distinct characteristic values are orthogonal.*

Proof. Suppose c is a characteristic value of T, i.e., that $T\alpha = c\alpha$ for some non-zero vector α. Then

$$\begin{aligned} c(\alpha|\alpha) &= (c\alpha|\alpha) \\ &= (T\alpha|\alpha) \\ &= (\alpha|T\alpha) \\ &= (\alpha|c\alpha) \\ &= \bar{c}(\alpha|\alpha). \end{aligned}$$

Since $(\alpha|\alpha) \neq 0$, we must have $c = \bar{c}$. Suppose we also have $T\beta = d\beta$ with $\beta \neq 0$. Then

$$\begin{aligned} c(\alpha|\beta) &= (T\alpha|\beta) \\ &= (\alpha|T\beta) \\ &= (\alpha|d\beta) \\ &= \bar{d}(\alpha|\beta) \\ &= d(\alpha|\beta). \end{aligned}$$

If $c \neq d$, then $(\alpha|\beta) = 0$. ∎

It should be pointed out that Theorem 15 says nothing about the existence of characteristic values or characteristic vectors.

Theorem 16. *On a finite-dimensional inner product space of positive dimension, every self-adjoint operator has a (non-zero) characteristic vector.*

Proof. Let V be an inner product space of dimension n, where $n > 0$, and let T be a self-adjoint operator on V. Choose an orthonormal basis $\mathcal{B}$ for V and let $A = [T]_{\mathcal{B}}$. Since $T = T^*$, we have $A = A^*$. Now let W be the space of $n \times 1$ matrices over C, with inner product $(X|Y) = Y^*X$. Then $U(X) = AX$ defines a self-adjoint linear operator U on W. The characteristic polynomial, det $(xI - A)$, is a polynomial of degree n over the complex numbers; every polynomial over C of positive degree has a root. Thus, there is a complex number c such that det $(cI - A) = 0$. This means that $A - cI$ is singular, or that there exists a non-zero X such that $AX = cX$. Since the operator U (multiplication by A) is self-adjoint, it follows from Theorem 15 that c is real. If V is a real vector space, we may choose X to have real entries. For then A and $A - cI$ have real entries, and since $A - cI$ is singular, the system $(A - cI)X = 0$ has a non-zero real solution X. It follows that there is a non-zero vector α in V such that $T\alpha = c\alpha$. ∎

There are several comments we should make about the proof.

(1) The proof of the existence of a non-zero X such that $AX = cX$ had nothing to do with the fact that A was Hermitian (self-adjoint). It shows that any linear operator on a finite-dimensional complex vector space has a characteristic vector. In the case of a real inner product space, the self-adjointness of A is used very heavily, to tell us that each characteristic value of A is real and hence that we can find a suitable X with real entries.

(2) The argument shows that the characteristic polynomial of a self-adjoint matrix has real coefficients, in spite of the fact that the matrix may not have real entries.

(3) The assumption that V is finite-dimensional is necessary for the theorem; a self-adjoint operator on an infinite-dimensional inner product space need not have a characteristic value.

EXAMPLE 29. Let V be the vector space of continuous complex-valued (or real-valued) continuous functions on the unit interval, $0 \leq t \leq 1$, with the inner product

$$(f|g) = \int_0^1 f(t)\overline{g(t)}\,dt.$$

The operator 'multiplication by t,' $(Tf)(t) = tf(t)$, is self-adjoint. Let us suppose that $Tf = cf$. Then

$$(t - c)f(t) = 0, \qquad 0 \leq t \leq 1$$

and so $f(t) = 0$ for $t \neq c$. Since f is continuous, $f = 0$. Hence T has no characteristic values (vectors).

Theorem 17. *Let V be a finite-dimensional inner product space, and let T be any linear operator on V. Suppose W is a subspace of V which is invariant under T. Then the orthogonal complement of W is invariant under T*.*

Proof. We recall that the fact that W is invariant under T does not mean that each vector in W is left fixed by T; it means that if α is in W then $T\alpha$ is in W. Let β be in $W^\perp$. We must show that $T^*\beta$ is in $W^\perp$, that is, that $(\alpha|T^*\beta) = 0$ for every α in W. If α is in W, then $T\alpha$ is in W, so $(T\alpha|\beta) = 0$. But $(T\alpha|\beta) = (\alpha|T^*\beta)$. ∎

Theorem 18. *Let V be a finite-dimensional inner product space, and let T be a self-adjoint linear operator on V. Then there is an orthonormal basis for V, each vector of which is a characteristic vector for T.*

Proof. We are assuming $\dim V > 0$. By Theorem 16, T has a characteristic vector α. Let $\alpha_1 = \alpha/||\alpha||$ so that α_1 is also a characteristic vector for T and $||\alpha_1|| = 1$. If $\dim V = 1$, we are done. Now we proceed by induction on the dimension of V. Suppose the theorem is true for inner product spaces of dimension less than $\dim V$. Let W be the one-dimensional subspace spanned by the vector α_1. The statement that α_1 is a characteristic vector for T simply means that W is invariant under T. By Theorem 17, the orthogonal complement $W^\perp$ is invariant under $T^* = T$. Now $W^\perp$, with the inner product from V, is an inner product space of dimension one less than the dimension of V. Let U be the linear operator induced on $W^\perp$ by T, that is, the restriction of T to $W^\perp$. Then U is self-adjoint, and by the induction hypothesis, $W^\perp$ has an orthonormal basis $\{\alpha_2, \ldots, \alpha_n\}$ consisting of characteristic vectors for U. Now each of these vectors is also a characteristic vector for T, and since $V = W \oplus W^\perp$, we conclude that $\{\alpha_1, \ldots, \alpha_n\}$ is the desired basis for V. ∎

Corollary. *Let A be an n × n Hermitian (self-adjoint) matrix. Then there is a unitary matrix P such that $P^{-1}AP$ is diagonal (A is unitarily equivalent to a diagonal matrix). If A is a real symmetric matrix, there is a real orthogonal matrix P such that $P^{-1}AP$ is diagonal.*

Proof. Let V be $C^{n\times 1}$, with the standard inner product, and let T be the linear operator on V which is represented by A in the standard ordered basis. Since $A = A^*$, we have $T = T^*$. Let $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$ be an ordered orthonormal basis for V, such that $T\alpha_j = c_j\alpha_j$, $j = 1, \ldots, n$. If $D = [T]_{\mathcal{B}}$, then D is the diagonal matrix with diagonal entries $c_1, \ldots, c_n$. Let P be the matrix with column vectors $\alpha_1, \ldots, \alpha_n$. Then $D = P^{-1}AP$.

In case each entry of A is real, we can take V to be R^n, with the standard inner product, and repeat the argument. In this case, P will be a unitary matrix with real entries, i.e., a real orthogonal matrix. ∎

Combining Theorem 18 with our comments at the beginning of this section, we have the following: If V is a finite-dimensional *real* inner product space and T is a linear operator on V, then V has an orthonormal basis of characteristic vectors for T if and only if T is self-adjoint. Equivalently, if A is an $n \times n$ matrix with *real* entries, there is a real orthogonal matrix P such that P^tAP is diagonal if and only if $A = A^t$. There is no such result for complex symmetric matrices. In other words, for complex matrices there is a significant difference between the conditions $A = A^t$ and $A = A^*$.

Having disposed of the self-adjoint case, we now return to the study of normal operators in general. We shall prove the analogue of Theorem 18 for normal operators, in the *complex* case. There is a reason for this restriction. A normal operator on a real inner product space may not have any non-zero characteristic vectors. This is true, for example, of all but two rotations in R^2.

Theorem 19. *Let V be a finite-dimensional inner product space and T a normal operator on V. Suppose α is a vector in V. Then α is a characteristic vector for T with characteristic value c if and only if α is a characteristic vector for T* with characteristic value $\bar{c}$.*

Proof. Suppose U is any normal operator on V. Then $||U\alpha|| = ||U^*\alpha||$. For using the condition $UU^* = U^*U$ one sees that

$$\begin{aligned} ||U\alpha||^2 = (U\alpha|U\alpha) &= (\alpha|U^*U\alpha) \\ &= (\alpha|UU^*\alpha) = (U^*\alpha|U^*\alpha) = ||U^*\alpha||^2. \end{aligned}$$

If c is any scalar, the operator $U = T - cI$ is normal. For $(T - cI)^* = T^* - \bar{c}I$, and it is easy to check that $UU^* = U^*U$. Thus

$$||(T - cI)\alpha|| = ||(T^* - \bar{c}I)\alpha||$$

so that $(T - cI)\alpha = 0$ if and only if $(T^* - \bar{c}I)\alpha = 0$. ∎

Definition. *A complex n × n matrix A is called* **normal** *if $AA^* = A^*A$.*

It is not so easy to understand what normality of matrices or operators really means; however, in trying to develop some feeling for the concept, the reader might find it helpful to know that a triangular matrix is normal if and only if it is diagonal.

Theorem 20. *Let V be a finite-dimensional inner product space, T a linear operator on V, and $\mathcal{B}$ an orthonormal basis for V. Suppose that the* *matrix A of T in the basis* $\mathfrak{B}$ *is upper triangular. Then T is normal if and only if A is a diagonal matrix.*

Proof. Since $\mathfrak{B}$ is an orthonormal basis, A^* is the matrix of T^* in $\mathfrak{B}$. If A is diagonal, then $AA^* = A^*A$, and this implies $TT^* = T^*T$. Conversely, suppose T is normal, and let $\mathfrak{B} = \{\alpha_1, \ldots, \alpha_n\}$. Then, since A is upper-triangular, $T\alpha_1 = A_{11}\alpha_1$. By Theorem 19 this implies, $T^*\alpha_1 = \bar{A}_{11}\alpha_1$. On the other hand,

$$\begin{aligned} T^*\alpha_1 &= \sum_j (A^*)_{j1}\alpha_j \\ &= \sum_j \bar{A}_{1j}\alpha_j. \end{aligned}$$

Therefore, $A_{1j} = 0$ for every $j > 1$. In particular, $A_{12} = 0$, and since A is upper-triangular, it follows that

$$T\alpha_2 = A_{22}\alpha_2.$$

Thus $T^*\alpha_2 = \bar{A}_{22}\alpha_2$ and $A_{2j} = 0$ for all $j \neq 2$. Continuing in this fashion, we find that A is diagonal. ▮

Theorem 21. *Let V be a finite-dimensional complex inner product space and let T be any linear operator on V. Then there is an orthonormal basis for V in which the matrix of T is upper triangular.*

Proof. Let n be the dimension of V. The theorem is true when $n = 1$, and we proceed by induction on n, assuming the result is true for linear operators on complex inner product spaces of dimension $n - 1$. Since V is a finite-dimensional complex inner product space, there is a unit vector α in V and a scalar c such that

$$T^*\alpha = c\alpha.$$

Let W be the orthogonal complement of the subspace spanned by α and let S be the restriction of T to W. By Theorem 17, W is invariant under T. Thus S is a linear operator on W. Since W has dimension $n - 1$, our inductive assumption implies the existence of an orthonormal basis $\{\alpha_1, \ldots, \alpha_{n-1}\}$ for W in which the matrix of S is upper-triangular; let $\alpha_n = \alpha$. Then $\{\alpha_1, \ldots, \alpha_n\}$ is an orthonormal basis for V in which the matrix of T is upper-triangular. ▮

This theorem implies the following result for matrices.

Corollary. *For every complex* n × n *matrix A there is a unitary matrix U such that* $U^{-1}AU$ *is upper-triangular.*

Now combining Theorem 21 and Theorem 20, we immediately obtain the following analogue of Theorem 18 for normal operators.

Theorem 22. *Let V be a finite-dimensional complex inner product space and T a normal operator on V. Then V has an orthonormal basis consisting of characteristic vectors for T.*

Again there is a matrix interpretation.

Corollary. *For every normal matrix A there is a unitary matrix P such that* $P^{-1}AP$ *is a diagonal matrix.*

Exercises

1. For each of the following real symmetric matrices A, find a real orthogonal matrix P such that P^tAP is diagonal.

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}, \quad \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix}$$

2. Is a complex symmetric matrix self-adjoint? Is it normal?

3. For

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{bmatrix}$$

there is a real orthogonal matrix P such that $P^tAP = D$ is diagonal. Find such a diagonal matrix D.

4. Let V be C^2, with the standard inner product. Let T be the linear operator on V which is represented in the standard ordered basis by the matrix

$$A = \begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix}.$$

Show that T is normal, and find an orthonormal basis for V, consisting of characteristic vectors for T.

5. Give an example of a 2×2 matrix A such that A^2 is normal, but A is not normal.

6. Let T be a normal operator on a finite-dimensional complex inner product space. Prove that T is self-adjoint, positive, or unitary according as every characteristic value of T is real, positive, or of absolute value 1. (Use Theorem 22 to reduce to a similar question about diagonal matrices.)

7. Let T be a linear operator on the finite-dimensional inner product space V, and suppose T is both positive and unitary. Prove $T = I$.

8. Prove T is normal if and only if $T = T_1 + iT_2$, where T_1 and T_2 are self-adjoint operators which commute.

9. Prove that a real symmetric matrix has a real symmetric cube root; i.e., if A is real symmetric, there is a real symmetric B such that $B^3 = A$.

10. Prove that every positive matrix is the square of a positive matrix.

11. Prove that a normal and nilpotent operator is the zero operator.

12. If T is a normal operator, prove that characteristic vectors for T which are associated with distinct characteristic values are orthogonal.

13. Let T be a normal operator on a finite-dimensional complex inner product space. Prove that there is a polynomial f, with complex coefficients, such that $T^* = f(T)$. (Represent T by a diagonal matrix, and see what f must be.)

14. If two normal operators commute, prove that their product is normal.

9. Operators on Inner Product Spaces

9.1. Introduction

We regard most of the topics treated in Chapter 8 as fundamental, the material that everyone should know. The present chapter is for the more advanced student or for the reader who is eager to expand his knowledge concerning operators on inner product spaces. With the exception of the Principal Axis theorem, which is essentially just another formulation of Theorem 18 on the orthogonal diagonalization of self adjoint operators, and the other results on forms in Section 9.2, the material presented here is more sophisticated and generally more involved technically. We also make more demands of the reader, just as we did in the later parts of Chapters 5 and 7. The arguments and proofs are written in a more condensed style, and there are almost no examples to smooth the way; however, we have seen to it that the reader is well supplied with generous sets of exercises.

The first three sections are devoted to results concerning forms on inner product spaces and the relation between forms and linear operators. The next section deals with spectral theory, i.e., with the implications of Theorems 18 and 22 of Chapter 8 concerning the diagonalization of self-adjoint and normal operators. In the final section, we pursue the study of normal operators treating, in particular, the real case, and in so doing we examine what the primary decomposition theorem of Chapter 6 says about normal operators.