5

EIGENVALUES OF OPERATORS ACTING ON EUCLIDEAN SPACES

5.1 Eigenvalues and inner products

This chapter describes some properties of eigenvalues and eigenvectors of linear transformations that operate on Euclidean spaces, that is, on linear spaces having an inner product. We recall the fundamental properties of inner products.

In a *real* Euclidean space an inner product (x, y) of two elements x andy is a real number satisfying the following properties :

(1) $(x, y) = (y, x)$ (symmetry)
(2) $(x + z, y) = (x, y) + (z, y)$ (linearity)
(3) $(cx, y) = c(x, y)$ (homogeneity)
(4) $(x, x) > 0$ if $x \neq O$ (positivity).

In a ***complex*** Euclidean space the inner product is a complex number satisfying the same properties, with the exception that symmetry is replaced by ***Hermitian symmetry,***

$$(1') \qquad (x, y) = \overline{(y, x)},$$

where the bar denotes the complex conjugate. In (3) the scalar c is complex. From (1') and (3) we obtain

$$(3') \qquad (x, cy) = \bar{c}(x, y),$$

which tells us that scalars are conjugated when taken out of the second factor. Taking $x = y$ in (1') we see that (x, x) is real so property (4) is meaningful if the space is complex.

When we use the term Euclidean space without further designation it is to be understood that the space can be real or complex. Although most of our applications will be to finite-dimensional spaces, we do not require this restriction at the outset.

The first theorem shows that eigenvalues (if they exist) can be expressed in terms of the inner product.

THEOREM 5.1. ***Let E be a Euclidean space, let V be a subspace of E, and let $T: V \to E$ be a linear transformation having an eigenvalue λ with a corresponding eigenvector x. Then we have***

$$(5.1) \qquad \lambda = \frac{(T(x), x)}{(x, x)}.$$

Proof. Since $T(x) = \lambda x$ we have

$$(T(x), x) = (\lambda x, x) = \lambda(x, x).$$

Since $x \neq 0$ we can divide by (x, x) to get (5.1).

Several properties of eigenvalues are easily deduced from Equation (5.1). For example, from the Hermitian symmetry of the inner product we have the companion formula

$$\bar{\lambda} = \frac{(x, T(x))}{(x, x)} \tag{5.2}$$

for the complex conjugate of λ. From (5.1) and (5.2) we see that λ is real $(\lambda = \bar{\lambda})$ if and only if $(T(x), x)$ is real, that is, if and only if

$$(T(x), x) = (x, T(x)) \qquad \text{for the eigenvector } x .$$

(This condition is trivially satisfied in a real Euclidean space.) Also, λ is pure imaginary $(\lambda = -\bar{\lambda})$ if and only if $(T(x), x)$ is pure imaginary, that is, if and only if

$$(T(x), x) = -(x, T(x)) \qquad \text{for the eigenvector } x.$$

5.2 Hermitian and skew-Hermitian transformations

In this section we introduce two important types of linear operators which act on Euclidean spaces. These operators have two categories of names, depending on whether the underlying Euclidean space has a real or complex inner product. In the real case the transformations are called *symmetric* and *skew-symmetric*. In the complex case they are called *Hermitian* and *skew-Hermitian*. These transformations occur in many different applications. For example, Hermitian operators on infinite-dimensional spaces play an important role in quantum mechanics. We shall discuss primarily the complex case since it presents no added difficulties.

DEFINITION. *Let E be a Euclidean space and let V be a subspace of E. A linear transformation $T: V \to E$ is called Hermitian on V if*

$$(T(x), y) = (x, T(y)) \qquad \textit{for all } x \textit{ and } y \textit{ in } V.$$

Operator T is called skew-Hermitian on V if

$$(T(x), y) = -(x, T(y)) \qquad \textit{for all } x \textit{ and } y \textit{ in } V.$$

In other words, a Hermitian operator T can be shifted from one factor of an inner product to the other without changing the value of the product. Shifting a skew-Hermitian operator changes the sign of the product.

Note: As already mentioned, if E is a *real* Euclidean space, Hermitian transformations are also called *symmetric;* skew-Hermitian transformations are called *skew-symmetric.*

EXAMPLE 1. ***Symmetry and skew-symmetry in the space C(a, b).*** Let $C(a, b)$ denote the space of all real functions continuous on a closed interval $[a, b]$, with the real inner product

$$(f, g) = \int_a^b f(t)g(t)\, dt.$$

Let V be a subspace of $C(a, b)$. If $T: V \to C(a, b)$ is a linear transformation then $(f, T(g)) = \int_a^b f(t)Tg(t)\, dt$, where we have written $Tg(t)$ for $T(g)(t)$. Therefore the conditions for symmetry and skew-symmetry become

$$\int_a^b \{f(t)Tg(t) - g(t)Tf(t)\}\, dt = 0 \quad \text{if } T \text{ is symmetric,} \tag{5.3}$$

and

$$\int_a^b \{f(t)Tg(t) + g(t)Tf(t)\}\, dt = 0 \quad \text{if } T \text{ is skew-symmetric.} \tag{5.4}$$

EXAMPLE 2. ***Multiplication by a fixed function.*** In the space $C(a, b)$ of Example 1, choose a fixed function p and define $T(f) = pf$, the product of p and f. For this T, Equation (5.3) is satisfied for all f and g in $C(a, b)$ since the integrand is zero. Therefore, multiplication by a fixed function is a symmetric operator.

EXAMPLE 3. ***The differentiation operator.*** In the space $C(a, b)$ of Example 1, let V be the subspace consisting of all functions f which have a continuous derivative in the open interval (a, b) and which also satisfy the *boundary condition* $f(a) = f(b)$. Let $D: V \to C(a, b)$ be the differentiation operator given by $D(f) = f'$. It is easy to prove that D is skew-symmetric. In this case the integrand in (5.4) is the derivative of the product fg, so the integral is equal to

$$\int_a^b (fg)'(t)\, dt = f(b)g(b) - f(a)g(a).$$

Since both f and g satisfy the boundary condition, we have $f(b)g(b) - f(a)g(a) = 0$. Thus, the boundary condition implies skew-symmetry of D. The only eigenfunctions in the subspace V are the constant functions. They belong to the eigenvalue 0.

EXAMPLE 4. ***Sturm-Liouville operators.*** This example is important in the theory of linear second-order differential equations. We use the space $C(a, b)$ of Example 1 once more and let V be the subspace consisting of all f which have a continuous second derivative in $[a, b]$ and which also satisfy the two boundary conditions

$$p(a)f(a) = 0, \qquad p(b)f(b) = 0, \tag{5.5}$$

where p is a fixed function in $C(a, b)$ with a continuous derivative on $[a, b]$. Let q be another fixed function in $C(a, b)$ and let $T: V \to C(a, b)$ be the operator defined by the equation

$$T(f) = (pf')' + qf.$$

This is called a *Sturm-Liouville operator*. To test for symmetry we note that $fT(g) - gT(f) = f(pg')' - g(pf')'$. Using this in (5.3) and integrating both $\int_a^b f \cdot (pg')'\, dt$ and $\int_a^b g \cdot (pf')'\, dt$ by parts, we find

$$\int_a^b \{fT(g) - gT(f)\}\, dt = fpg' \Big|_a^b - \int_a^b pg'f'\, dt - gpf' \Big|_a^b + \int_a^b pf'g'\, dt = 0,$$

since both f and g satisfy the boundary conditions (5.5). Hence T is symmetric on V. The eigenfunctions of T are those nonzero f which satisfy, for some real λ, a differential equation of the form

$$(pf')' + qf = \lambda f$$

on $[a, b]$, and also satisfy the boundary conditions (5.5).

5.3 Eigenvalues and eigenvectors of Hermitian and skew-Hermitian operators

Regarding eigenvalues we have the following theorem;

THEOREM 5.2. *Assume T has an eigenvalue λ. Then we have:*
(a) *If T is Hermitian, λ is real: $\lambda = \bar{\lambda}$.*
(b) *If T is skew-Hermitian, λ is pure imaginary: $\lambda = -\bar{\lambda}$.*

Proof. Let x be an eigenvector corresponding to λ. Then we have

$$\lambda = \frac{(T(x), x)}{(x, x)} \quad \text{and} \quad \bar{\lambda} = \frac{(x, T(x))}{(x, x)}.$$

If T is Hermitian we have $(T(x), x) = (x, T(x))$ so $\lambda = \bar{\lambda}$. If T is skew-Hermitian we have $(T(x), x) = -(x, T(x))$ so $\lambda = -\bar{\lambda}$.

Note: If T is *symmetric*, Theorem 5.2 tells us nothing new about the eigenvalues of T since all the eigenvalues must be real if the inner product is real.. If T is *skew-symmetric*, the eigenvalues of T must be both real and pure imaginary. Hence all the eigenvalues of a skew-symmetric operator must be zero (if any exist).

5.4 Orthogonality of eigenvectors corresponding to distinct eigenvalues

Distinct eigenvalues of any linear transformation correspond to independent eigenvectors (by Theorem 4.2). For Hermitian and skew-Hermitian transformations more is true.

THEOREM 5.3. *Let T be a Hermitian or skew-Hermitian transformation, and let λ and μ be distinct eigenvalues of T with corresponding eigenvectors x and y. Then x and y are orthogonal; that is, $(x, y) = 0$.*

Proof. We write $T(x) = \lambda x$, $T(y) = \mu y$ and compare the two inner products $(T(x), y)$ and $(x, T(y))$. We have

$$(T(x), y) = (\lambda x, y) = \lambda(x, y) \quad \text{and} \quad (x, T(y)) = (x, \mu y) = \bar{\mu}(x, y).$$

If T is Hermitian this gives us $\lambda(x, y) = \bar{\mu}(x, y) = \mu(x, y)$ since $\mu = \bar{\mu}$. Therefore $(x, y) = 0$ since $\lambda \neq \mu$. If T is skew-Hermitian we obtain $\lambda(x, y) = -\bar{\mu}(x, y) = \mu(x, y)$ which again implies $(x, y) = 0$.

EXAMPLE. We apply Theorem 5.3 to those nonzero functions which satisfy a differential equation of the form

$$(pf')' + qf = \lambda f \tag{5.6}$$

on an interval $[a, b]$, and which also satisfy the boundary conditions $p(a)f(a) = p(b)f(b) = 0$. The conclusion is that any two solutions f and g corresponding to two distinct values of λ are orthogonal. For example, consider the differential equation of simple harmonic motion,

$$f'' + k^2 f = 0$$

on the interval $[0, \pi]$, where $k \neq 0$. This has the form (5.6) with $p = 1$, $q = 0$, and $\lambda = -k^2$. All solutions are given by $f(t) = c_1 \cos kt + c_2 \sin kt$. The boundary condition $f(0) = 0$ implies $c_1 = 0$. The second boundary condition, $f(\pi) = 0$, implies $c_2 \sin k\pi = 0$. Since $c_2 \neq 0$ for a nonzero solution, we must have $\sin k\pi = 0$, which means that k is an integer. In other words, nonzero solutions which satisfy the boundary conditions exist if and only if k is an integer. These solutions are $f(t) = \sin nt$, $n = \pm 1, \pm 2, \ldots$. The orthogonality condition implied by Theorem 5.3 now becomes the familiar relation

$$\int_0^{\pi} \sin nt \sin mt \, dt = 0$$

if m^2 and n^2 are distinct integers.

5.5 Exercises

1. Let E be a Euclidean space, let V be a subspace, and let $T: V \to E$ be a given linear transformation. Let λ be a scalar and x a nonzero element of V. Prove that λ is an eigenvalue of T with x as an eigenvector if and only if

$$(T(x), y) = \lambda(x, y) \qquad \text{for every } y \text{ in } E.$$

2. Let $T(x) = cx$ for every x in a linear space V, where c is a fixed scalar. Prove that T is symmetric if V is a real Euclidean space.
3. Assume $T: V \to V$ is a Hermitian transformation.
   (a) Prove that T^n is Hermitian for every positive integer n, and that T^{-1} is Hermitian if T is invertible.
   (b) What can you conclude about T^n and T^{-1} if T is skew-Hermitian?
4. Let $T_1: V \to E$ and $T_2: V \to E$ be two Hermitian transformations.
   (a) Prove that $aT_1 + bT_2$ is Hermitian for all real scalars a and b.
   (b) Prove that the product (composition) T_1T_2 is Hermitian if T_1 and T_2 commute, that is, if $T_1T_2 = T_2T_1$.
5. Let $V = V_3(\mathbf{R})$ with the usual dot product as inner product. Let T be a reflection in the xy-plane; that is, let $T(i) = i$, $T(j) = j$, and $T(k) = -k$. Prove that T is symmetric.

6. Let $C(0, 1)$ be the real linear space of all real functions continuous on $[0, 1]$ with inner product $(f, g) = \int_0^1 f(t)g(t)\,dt$. Let V be the subspace of all f such that $\int_0^1 f(t)\,dt = 0$. Let $T: V \to C(0, 1)$ be the integration operator defined by $Tf(x) = \int_0^x f(t)\,dt$. Prove that T is skew-symmetric.
7. Let V be the real Euclidean space of all real polynomials with the inner product $(f, g) = \int_{-1}^1 f(t)g(t)\,dt$. Determine which of the following transformations $T: V \to V$ is symmetric or skew-symmetric:
   (a) $Tf(x) = f(-x)$.
   (b) $Tf(x) = f(x)f(-x)$.
   (c) $Tf(x) = f(x) + f(-x)$.
   (d) $Tf(x) = f(x) - f(-x)$.
8. Refer to Example 4 of Section 5.2. Modify the inner product as follows:

$$(f, g) = \int_a^b f(t)g(t)w(t)\,dt,$$

where w is a fixed positive function in $C(a, b)$. Modify the Sturm-Liouville operator T by writing

$$T(f) = \frac{(pf')' + qf}{w}.$$

Prove that the modified operator is symmetric on the subspace V.
9. Let V be a subspace of a complex Euclidean space E. Let $T: V \to E$ be a linear transformation and define a scalar-valued function Q on V as follows:

$$Q(x) = (T(x), x) \qquad \text{for all } x \text{ in } V.$$

(a) If T is Hermitian on V, prove that $Q(x)$ is real for all x.
(b) If T is skew-Hermitian, prove that $Q(x)$ is pure imaginary for all x.
(c) Prove that $Q(tx) = t\bar{t}Q(x)$ for every scalar t.
(d) Prove that $Q(x + y) = Q(x) + Q(y) + (T(x), y) + (T(y), x)$, and find a corresponding formula for $Q(x + ty)$.
(e) If $Q(x) = 0$ for all x prove that $T(x) = 0$ for all x.
(f) If $Q(x)$ is real for all x prove that T is Hermitian. [*Hint:* Use the fact that $Q(x + ty)$ equals its conjugate for every scalar t.]
10. This exercise shows that the Legendre polynomials (introduced in Section 1.14) are eigenfunctions of a Sturm-Liouville operator. The Legendre polynomials are defined by the equation

$$P_n(t) = \frac{1}{2^n n!} f_n^{(n)}(t), \qquad \text{where } f_n(t) = (t^2 - 1)^n.$$

(a) Verify that $(t^2 - 1) f_n'(t) = 2nt f_n(t)$.
(b) Differentiate the equation in (a) $n + 1$ times, using Leibniz's formula (see p. 222 of Volume I) to obtain

$$(t^2 - 1)f_n^{(n+2)}(t) + 2t(n + 1)f_n^{(n+1)}(t) + n(n + 1)f_n^{(n)}(t) = 2ntf_n^{(n+1)}(t) + 2n(n + 1)f_n^{(n)}(t).$$

(c) Show that the equation in (b) can be rewritten in the form

$$[(t^2 - 1)P_n'(t)]' = n(n + 1)P_n(t).$$

This shows that $P_n(t)$ is an eigenfunction of the Sturm-Liouville operator T given on the interval [-1 , 1] by $T(f) = (pf')'$, where $p(t) = t^2 - 1$. The eigenfunction $P_n(t)$ belongs to the eigenvalue $\lambda = n(n + 1)$. In this example the boundary conditions for symmetry are automatically satisfied since $p(1) = p(-1) = 0$.

5.6 Existence of an orthonormal set of eigenvectors for Hermitian and skew-Hermitian operators acting on finite-dimensional spaces

Both Theorems 5.2 and 5.3 are based on the assumption that T has an eigenvalue. As we know, eigenvalues need not exist. However, if T acts on a *finite-dimensional complex* space, then eigenvalues always exist since they are the roots of the characteristic polynomial. If T is Hermitian, all the eigenvalues are real. If T is skew-Hermitian, all the eigenvalues are pure imaginary.

We also know that two distinct eigenvalues belong to orthogonal eigenvectors if T is Hermitian or skew-Hermitian. Using this property we can prove that T has an orthonormal set of eigenvectors which spans the whole space. (We recall that an orthogonal set is called orthonormal if each of its (elements has norm 1.)

THEOREM 5.4. *Assume* dim $V = n$ *and let* $T: V \to V$ *be Hermitian or skew-Hermitian. Then there exist n eigenvectors* $u_1, \ldots, u_n$ *of T which form an orthonormal basis for V. Hence the matrix of T relative to this basis is the diagonal matrix* $\Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$, *where* λ_k *is the eigenvalue belonging to* u_k.

Proof. We use inductian on the dimension n. If $n = 1$, then T has exactly one eigenvalue. Any eigenvector u_1 of norm 1 is an orthonormal basis for V.

Now assume the theorem is true for every Euclidean space of dimension $n - 1$. To prove it is also true for V we choose an eigenvalue λ_1 for T and a corresponding eigenvector u_1 of norm 1. Then $T(u_1) = \lambda_1 u_1$ and $\|u_1\| = 1$. Let S be the subspace spanned by u_1. We shall apply the induction hypothesis to the subspace $S^\perp$ consisting of all elements in V which are orthogonal to u_1,

$$S^\perp = \{x \mid x \in V, (x, u_1) = 0\}.$$

To do this we need to know that dim $S^\perp = n - 1$ and that T maps $S^\perp$ into itself.

From Theorem 1.7(a) we know that u_1 is part of a basis for V, say the basis $(u_1, v_2, \ldots, v_n)$. We can assume, without loss in generality, that this is an orthonormal basis. (If not, we apply the Gram-Schmidt process to convert it into an orthonormal basis, keeping u_1 as the first basis element.) Now take any x in $S^\perp$ and write

$$x = x_1 u_1 + x_2 v_2 + \cdots + x_n v_n .$$

Then $x_1 = (x, u_1) = 0$ since the basis is orthonormal, so x is in the space spanned by $v_2, \ldots, v_n$. Hence dim $S^\perp = n - 1$.

Next we show that T maps $S^\perp$ into itself. Assume T is Hermitian. If $x \in S^\perp$ we have

$$(T(x), u_1) = (x, T(u_1)) = (x, \lambda_1 u_1) = \lambda_1 (x, u_1) = 0,$$

so $T(x) \in S^\perp$. Since T is Hermitian on $S^\perp$ we can apply the induction hypothesis to find that T has $n - 1$ eigenvectors $u_2, \ldots, u_n$ which form an orthonormal basis for $S^\perp$. Therefore the orthogonal set $u_1, \ldots, u_n$ is an orthonormal basis for V. This proves the theorem if T is Hermitian. A similar argument works if T is skew-Hermitian.

5.7 Matrix representations for Hermitian and skew-Hermitian operators

In this section we assume that V is a finite-dimensional Euclidean space. A Hermitian or skew-Hermitian transformation can be characterized in terms of its action on the elements of any basis.

THEOREM 5.5. *Let $(e_1, \ldots, e_n)$ be a basis for V and let $T: V \to V$ be a linear transformation. Then we have:*

(a) *T is Hermitian if and only if $(T(e_j), e_i) = (e_j, T(e_i))$ for all i and j.*
(b) *T is skew-Hermitian if and only if $(T(e_j), e_i) = -(e_j, T(e_i))$ for all i and j.*

Proof. Take any two elements x and y in V and express each in terms of the basis elements, say $x = \sum x_j e_j$ and $y = \sum y_i e_i$. Then we have

$$(T(x), y) = \left(\sum_{j=1}^{n} x_j T(e_j), y\right) = \sum_{j=1}^{n} x_j \left(T(e_j), \sum_{i=1}^{n} y_i e_i\right) = \sum_{j=1}^{n} \sum_{i=1}^{n} x_j \bar{y}_i (T(e_j), e_i).$$

Similarly we find

$$(x, T(y)) = \sum_{j=1}^{n} \sum_{i=1}^{n} x_j \bar{y}_i (e_j, T(e_i)).$$

Statements (a) and (b) following immediately from these equations.

Now we characterize these concepts in terms of a matrix representation of T.

THEOREM 5.6. *Let $(e_1, \ldots, e_n)$ be an orthonormal basis for V, and let $A = (a_{ij})$ be the matrix representation of a linear transformation $T: V \to V$ relative to this basis. Then we have:*

(a) *T is Hermitian if and only if $a_{ij} = \bar{a}_{ji}$ for all i and j.*
(b) *T is skew-Hermitian if and only if $a_{ij} = -\bar{a}_{ji}$ for all i and j.*

Proof. Since A is the matrix of T we have $T(e_j) = \sum_{k=1}^{n} a_{kj} e_k$. Taking the inner product of $T(e_j)$ with e_i and using the linearity of the inner product we obtain

$$(T(e_j), e_i) = \left(\sum_{k=1}^{n} a_{kj} e_k, e_i\right) = \sum_{k=1}^{n} a_{kj} (e_k, e_i).$$

But $(e_k, e_i) = 0$ unless $k = i$, so the last sum simplifies to $a_{ij}(e_i, e_i) = a_{ij}$ since $(e_i, e_i) = 1$. Hence we have

$$a_{ij} = (T(e_j), e_i) \qquad \text{for all } i, j.$$

Interchanging i and j, taking conjugates, and using the Hermitian symmetry of the inner product, we find

$$\bar{a}_{ji} = (e_j, T(e_i)) \qquad \text{for all } i, j.$$

Now we apply Theorem 5.5 to complete the proof.

5.8 Hermitian and skew-Hermitian matrices. The adjoint of a matrix

The following definition is suggested by Theorem 5.6.

DEFINITION. ***A square matrix $A = (a_{ij})$ is called Hermitian if $a_{ij} = \bar{a}_{ji}$ for all i and j. Matrix A is called skew-Hermitian if $a_{ij} = -\bar{a}_{ji}$ for all i and j.***

Theorem 5.6 states that a transformation T on a finite-dimensional space V is Hermitian or skew-Hermitian according as its matrix relative to an orthonormal basis is Hermitian or skew-Hermitian.

These matrices can be described in another way. Let $\bar{A}$ denote the matrix obtained by replacing each entry of A by its complex conjugate. Matrix $\bar{A}$ is called the ***conjugate*** of A. Matrix A is Hermitian if and only if it is equal to the transpose of its conjugate, $A = \bar{A}^t$. It is skew-Hermitian if $A = -\bar{A}^t$.

The transpose of the conjugate is given a special name.

DEFINITION OF THE ADJOINT OF A MATRIX. ***For any matrix A, the transpose of the conjugate, $\bar{A}^t$, is also called the adjoint of A and is denoted by A^*.***

Thus, a square matrix A is Hermitian if $A = A^*$, and skew-Hermitian if $A = -A^*$. A Hermitian matrix is also called ***self-adjoint.***

> ***Note:*** Much of the older matrix literature uses the term *adjoint* for the transpose of the cofactor matrix, an entirely different object. The definition given here conforms to the current nomenclature in the theory of linear operators.

5.9 Diagonalization of a Hermitian or skew-Hermitian matrix

THEOREM ***5.7. Every $n \times n$ Hermitian or skew-Hermitian matrix A is similar to the diagonal matrix $\Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ of its eigenvalues. Moreover, we have***

$$\Lambda = C^{-1}AC,$$

where C is a nonsingular matrix whose inverse is its adjoint, $C^{-1} = C^*$.

Proof. Let V be the space of n-tuples of complex numbers, and let $(e_1, \ldots, e_n)$ be the orthonormal basis of unit coordinate vectors. If $x = \sum x_i e_i$ and $y = \sum y_i e_i$ let the inner product be given by $(x, y) = \sum x_i \bar{y}_i$. For the given matrix A, let T be the transformation represented by A relative to the chosen basis. Then Theorem 5.4 tells us that V has an orthonormal basis of eigenvectors $(u_1, \ldots, u_n)$, relative to which T has the diagonal matrix representation $\Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$, where λ_k is the eigenvalue belonging to u_k. Since both A and Λ represent T they are similar, so we have $\Lambda = C^{-1}AC$, where $C = (c_{ij})$ is the nonsingular matrix relating the two bases:

$$[u_1, \ldots, u_n] = [e_1, \ldots, e_n]C.$$

This equation shows that the jth column of C consists of the components of u_j relative to $(e_1, \ldots, e_n)$. Therefore c_{ij} is the ith component of u_j. The inner product of u_j and u_i is given by

$$(u_j, u_i) = \sum_{k=1}^{n} c_{kj}\bar{c}_{ki}.$$

Since $\{u_1, \ldots, u_n\}$ is an orthonormal set, this shows that $CC^* = I$, so $C^{-1} = C^*$.

Note: The proof of Theorem 5.7 also tells us how to determine the diagonalizing matrix C. We find an orthonormal set of eigenvectors $u_1, \ldots, u_n$ and then use the components of u_j (relative to the basis of unit coordinate vectors) as the entries of the jth column of C.

EXAMPLE 1. The real Hermitian matrix $A = \begin{bmatrix} 2 & 2 \\ 2 & 5 \end{bmatrix}$ has eigenvalues $\lambda_1 = 1$ and $\lambda_2 = 6$. The eigenvectors belonging to 1 are $t(2, -1)$, $t \neq 0$. Those belonging to 6 are $t(1, 2)$, $t \neq 0$. The two eigenvectors $u_1 = t(2, -1)$ and $u_2 = t(1, 2)$ with $t = 1/\sqrt{5}$ form an orthonormal set. Therefore the matrix

$$C = \frac{1}{\sqrt{5}}\begin{bmatrix} 2 & 1 \\ -1 & 2 \end{bmatrix}$$

is a diagonalizing matrix for A. In this case $C^* = C^t$ since C is real. It is easily verified that $C^tAC = \begin{bmatrix} 1 & 0 \\ 0 & 6 \end{bmatrix}$.

EXAMPLE 2. If A is already a diagonal matrix, then the diagonalizing matrix C of Theorem 5.7 either leaves A unchanged or merely rearranges the diagonal entries.

5.10 Unitary matrices. Orthogonal matrices

DEFINITION. *A square matrix A is called unitary if $AA^* = I$. It is called orthogonal if $AA^t = I$.*

Note: Every real unitary matrix is orthogonal since $A^* = A^t$.

Theorem 5.7 tells us that a Hermitian or skew-Hermitian matrix can always be diagonalized by a unitary matrix. A real Hermitian matrix has real eigenvalues and the corresponding eigenvectors can be taken real. Therefore a real Hermitian matrix can be diagonalized by a real orthogonal matrix. This is *not* true for real skew-Hermitian matrices. (See Exercise 11 in Section 5.11.)

We also have the following related concepts.

DEFINITION. *A square matrix A with real or complex entries is called symmetric if $A = A^t$; it is called skew-symmetric if $A = -A^t$.*

EXAMPLE 3. If A is real, its adjoint is equal to its transpose, $A^* = A^t$. Thus, every *real* Hermitian matrix is symmetric, but a symmetric matrix need not be Hermitian.

EXAMPLE: 4. If $A = \begin{bmatrix} 1+i & 2 \\ 3-i & 4i \end{bmatrix}$, then $\bar{A} = \begin{bmatrix} 1-i & 2 \\ 3+i & -4i \end{bmatrix}$, $A^t = \begin{bmatrix} 1+i & 3-i \\ 2 & 4i \end{bmatrix}$ and $A^* = \begin{bmatrix} 1-i & 3+i \\ 2 & -4i \end{bmatrix}$.

EXAMPLE: 5. Both matrices $\begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}$ and $\begin{bmatrix} 1 & 2+i \\ 2-i & 3 \end{bmatrix}$ are Hermitian. The first is symmetric, the second is not.

EXAMPLE 6. Both matrices $\begin{bmatrix} 0 & -2 \\ 2 & 0 \end{bmatrix}$ and $\begin{bmatrix} i & -2 \\ 2 & 3i \end{bmatrix}$ are skew-Hermitian. The first is skew-symmetric, the second is not.

EXAMPLE 7. All the diagonal elements of a Hermitian matrix are real. All the diagonal elements of a skew-Hermitian matrix are pure imaginary. All the diagonal elements of a skew-symmetric matrix are zero.

EXAMPLE 8. For any square matrix A, the matrix $B = \frac{1}{2}(A + A^*)$ is Hermitian, and the matrix $C = \frac{1}{2}(A - A^*)$ is skew-Hermitian. Their sum is A. Thus, every square matrix A can be expressed as a sum $A = B + C$, where B is Hermitian and C is skew-Hermitian. It is an easy exercise to verify that this decomposition is unique. Also every square matrix A can be expressed uniquely as the sum of a symmetric matrix, $\frac{1}{2}(A + A^t)$, and a skew-symmetric matrix, $\frac{1}{2}(A - A^t)$.

EXAMPLE 9. If A is orthogonal we have $1 = \det (AA^t) = (\det A)(\det A^t) = (\det A)^2$, so $\det A =$ f 1.

5.11 Exercises

1. Determine which of the following matrices are symmetric, skew-symmetric, Hermitian, skew-Hermitian.

(a) $\begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 3 \\ 2 & 3 & 4 \end{bmatrix}$, (b) $\begin{bmatrix} 0 & i & 2 \\ i & 0 & 3 \\ -2 & -3 & 4i \end{bmatrix}$, (c) $\begin{bmatrix} 0 & i & 2 \\ -i & 0 & 3 \\ -2 & -3 & 0 \end{bmatrix}$, (d) $\begin{bmatrix} 0 & 1 & 2 \\ -1 & 0 & 3 \\ -2 & -3 & 0 \end{bmatrix}$.

2. (a) Verify that the 2×2 matrix $A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ is an orthogonal matrix.

   (b) Let T be the linear transformation with the above matrix A relative to the usual basis $\{i, j\}$. Prove that T maps each point in the plane with polar coordinates (r, α) onto the point with polar coordinates $(r, \alpha + \theta)$. Thus, T is a rotation of the plane about the origin, θ being the angle of rotation.

3. Let V be real 3-space with the usual basis vectors i, j, k. Prove that each of the following matrices is orthogonal and represents the transformation indicated.

   (a) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$ (reflection in the xy-plane).

   (b) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$ (reflection through the x-axis).

   (c) $\begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$ (reflection through the origin).

   (d) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix}$ (rotation about the x-axis).

   (e) $\begin{bmatrix} -1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix}$ (rotation about x-axis followed by reflection in the yz-plane).

4. A real orthogonal matrix A is called ***proper*** if $\det A = 1$, and ***improper*** if $\det A = -1$.

   (a) If A is a proper 2×2 matrix, prove that $A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ for some θ. This represents a rotation through an angle θ.

   (b) Prove that $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ and $\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$ are improper matrices. The first matrix represents a reflection of the xy-plane through the x-axis; the second represents a reflection through the y-axis. Find all improper 2×2 matrices.

In each of Exercises 5 through 8, find (a) an orthogonal set of eigenvectors for A, and (b) a unitary matrix C such that $C^{-1}AC$ is a diagonal matrix.

5. $A = \begin{bmatrix} 9 & 12 \\ 12 & 16 \end{bmatrix}$.

6. $A = \begin{bmatrix} 0 & -2 \\ 2 & 0 \end{bmatrix}$.

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

8. $A = \begin{bmatrix} 1 & 3 & 4 \\ 3 & 1 & 0 \\ 4 & 0 & 1 \end{bmatrix}$.

9. Determine which of the following matrices are unitary, and which are orthogonal (a, b, θ real).

$$\text{(a)} \begin{bmatrix} e^{ia} & 0 \\ 0 & e^{ib} \end{bmatrix}, \qquad \text{(b)} \begin{bmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{bmatrix}, \qquad \text{(c)} \begin{bmatrix} \frac{1}{2}\sqrt{2} & -\frac{1}{3}\sqrt{3} & \frac{1}{6}\sqrt{6} \\ 0 & \frac{1}{3}\sqrt{3} & \frac{1}{3}\sqrt{6} \\ \frac{1}{2}\sqrt{2} & \frac{1}{3}\sqrt{3} & -\frac{1}{6}\sqrt{6} \end{bmatrix}.$$

10. The special theory of relativity makes use of the equations

$$x' = a(x - vt), \qquad y' = y, \qquad z' = z, \qquad t' = a(t - vx/c^2).$$

Here v is the velocity of a moving object, c the speed of light, and $a = c/\sqrt{c^2 - v^2}$. The linear transformation which maps (x, y, z, t) onto (x', y', z', t') is called a *Lorentz transformation.*
(a) Let $(x_1, x_2, x_3, x_4) = (x, y, z, ict)$ and $(x_1', x_2', x_3', x_4') = (x', y', z', ict')$. Show that the four equations can be written as one matrix equation,

$$[x_1', x_2', x_3', x_4'] = [x_1, x_2, x_3, x_4] \begin{bmatrix} a & 0 & 0 & -iav/c \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ iav/c & 0 & 0 & a \end{bmatrix}.$$

(b) Prove that the 4×4 matrix in (a) is orthogonal but not unitary.

11. Let a be a nonzero real number and let A be the skew-symmetric matrix $A = \begin{bmatrix} 0 & a \\ -a & 0 \end{bmatrix}$.
(a) Find an orthonormal set of eigenvectors for A.
(b) Find a unitary matrix C such that $C^{-1}AC$ is a diagonal matrix.
(c) Prove that there is no real orthogonal matrix C such that $C^{-1}AC$ is a diagonal matrix.

12. If the eigenvalues of a Hermitian or skew-Hermitian matrix A are all equal to c, prove that $A = cI$.

13. If A is a real skew-symmetric matrix, prove that both $I - A$ and $I + A$ are nonsingular and that $(I - A)(I + A)^{-1}$ is orthogonal.

14. For each of the following statements about $n \times n$ matrices, give a proof or exhibit a counter example.
(a) If A and B are unitary, then $A + B$ is unitary.
(b) If A and B are unitary, then AB is unitary.
(c) If A and AB are unitary, then B is unitary.
(d) If A and B are unitary, then $A + B$ is not unitary.

5.12 Quadratic forms

Let V be a *real* Euclidean space and let $T: V \to V$ be a symmetric operator. This means that T can be shifted from one factor of an inner product to the other,

$$(T(x), y) = (x, T(y)) \qquad \text{for all } x \text{ and } y \text{ in } V.$$

Given T, we define a real-valued function Q on V by the equation

$$Q(x) = (T(x), x).$$

The function Q is called the *quadratic form* associated with T. The term "quadratic" is suggested by the following theorem which shows that in the finite-dimensional case $Q(x)$ is a quadratic polynomial in the components of x.

THEOREM 5.8. *Let $(e_1, \ldots, e_n)$ be an orthonormal basis for a real Euclidean space V. Let $T: V \to V$ be a symmetric transformation, and let $A = (a_{ij})$ be the matrix of T relative to this basis. Then the quadratic form $Q(x) = (T(x), x)$ is related to A as follows:*

$$Q(x) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j \qquad \text{if} \quad x = \sum_{i=1}^{n} x_i e_i . \tag{5.7}$$

Proof. By linearity we have $T(x) = \sum x_i T(e_i)$. Therefore

$$Q(x) = \left(\sum_{i=1}^{n} x_i T(e_i), \sum_{j=1}^{n} x_j e_j \right) = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i x_j (T(e_i), e_j) .$$

This proves (5.7) since $a_{ij} = a_{ji} = (T(e_i), e_j)$.

The sum appearing in (5.7) is meaningful even if the matrix A is not symmetric.

DEFINITION. *Let V be any real Euclidean space with an orthonormal basis $(e_1, \ldots, e_n)$, and let $A = (a_{ij})$ by any $n \times n$ matrix of scalars. The scalar-valued function Q defined at each element $x = \sum x_i e_i$ of V by the double sum*

$$Q(x) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j \tag{5.8}$$

is called the quadratic form associated with A.

If A is a diagonal matrix, then $a_{ij} = 0$ if $i \neq j$ so the sum in (5.8) contains only squared terms and can be written more simply as

$$Q(x) = \sum_{i=1}^{n} a_{ii} x_i^2 .$$

In this case the quadratic form is called a *diagonalform.*

The double sum appearing in (5.8) can also be expressed as a product of three matrices.

THEOREM 5.9. *Let $X = [x_1, \ldots, x_n]$ be a $1 \times n$ row matrix, and let $A = (a_{ij})$ be an $n \times n$ matrix. Then XAX^t is a 1×1 matrix with entry*

$$\sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j . \tag{5.9}$$

Proof. Th.e product XA is a $1 \times n$ matrix, $XA = [y_1, \ldots, y_n]$, where entry y_j is the dot product of X with thejth column of A,

$$y_j = \sum_{i=1}^{n} x_i a_{ij} .$$