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.

9.2. Forms on Inner Product Spaces

If T is a linear operator on a finite-dimensional inner product space V the function f defined on $V \times V$ by

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

may be regarded as a kind of substitute for T. Many questions about T are equivalent to questions concerning f. In fact, it is easy to see that f determines T. For if $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$ is an orthonormal basis for V, then the entries of the matrix of T in $\mathcal{B}$ are given by

$$A_{jk} = f(\alpha_k, \alpha_j).$$

It is important to understand why f determines T from a more abstract point of view. The crucial properties of f are described in the following definition.

Definition. *A* **(sesqui-linear) form** *on a real or complex vector space* V *is a function* f *on* V $\times$ V *with values in the field of scalars such that*

$$\text{(a)} \qquad f(c\alpha + \beta, \gamma) = cf(\alpha, \gamma) + f(\beta, \gamma)$$
$$\text{(b)} \qquad f(\alpha, c\beta + \gamma) = \bar{c}f(\alpha, \beta) + f(\alpha, \gamma)$$

for all α, β, γ *in* V *and all scalars* c.

Thus, a sesqui-linear form is a function on $V \times V$ such that $f(\alpha, \beta)$ is a linear function of α for fixed β and a conjugate-linear function of β for fixed α. In the real case, $f(\alpha, \beta)$ is linear as a function of each argument; in other words, f is a **bilinear form.** In the complex case, the sesqui-linear form f is not bilinear unless $f = 0$. In the remainder of this chapter, we shall omit the adjective 'sesqui-linear' unless it seems important to include it.

If f and g are forms on V and c is a scalar, it is easy to check that $cf + g$ is also a form. From this it follows that any linear combination of forms on V is again a form. Thus the set of all forms on V is a subspace of the vector space of all scalar-valued functions on $V \times V$.

Theorem 1. *Let* V *be a finite-dimensional inner product space and* f *a form on* V. *Then there is a unique linear operator* T *on* V *such that*

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

for all α, β *in* V, *and the map* f $\rightarrow$ T *is an isomorphism of the space of forms onto* L(V, V).

Proof. Fix a vector β in V. Then $\alpha \rightarrow f(\alpha, \beta)$ is a linear function on V. By Theorem 6 there is a unique vector β' in V such that $f(\alpha, \beta) = (\alpha|\beta')$ for every α. We define a function U from V into V by setting $U\beta = \beta'$. Then

$$\begin{aligned} f(\alpha|c\beta + \gamma) &= (\alpha|U(c\beta + \gamma)) \\ &= \bar{c}f(\alpha, \beta) + f(\alpha, \gamma) \\ &= \bar{c}(\alpha|U\beta) + (\alpha|U\gamma) \\ &= (\alpha|cU\beta + U\gamma) \end{aligned}$$

for all α, β, γ in V and all scalars c. Thus U is a linear operator on V and $T = U^*$ is an operator such that $f(\alpha, \beta) = (T\alpha|\beta)$ for all α and β. If we also have $f(\alpha, \beta) = (T'\alpha|\beta)$, then

$$(T\alpha - T'\alpha|\beta) = 0$$

for all α and β; so $T\alpha = T'\alpha$ for all α. Thus for each form f there is a unique linear operator T_f such that

$$f(\alpha, \beta) = (T_f\alpha|\beta)$$

for all α, β in V. If f and g are forms and c a scalar, then

$$\begin{aligned} (cf + g)(\alpha, \beta) &= (T_{cf+g}\alpha|\beta) \\ &= cf(\alpha, \beta) + g(\alpha, \beta) \\ &= c(T_f\alpha|\beta) + (T_g\alpha|\beta) \\ &= ((cT_f + T_g)\alpha|\beta) \end{aligned}$$

for all α and β in V. Therefore,

$$T_{cf+g} = cT_f + T_g$$

so $f \to T_f$ is a linear map. For each T in $L(V, V)$ the equation

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

defines a form such that $T_f = T$, and $T_f = 0$ if and only if $f = 0$. Thus $f \to T_f$ is an isomorphism. ∎

Corollary. *The equation*

$$(f|g) = tr\,(T_fT_g^*)$$

defines an inner product on the space of forms with the property that

$$(f|g) = \sum_{j,k} f(\alpha_k, \alpha_j)\overline{g(\alpha_k, \alpha_j)}$$

for every orthonormal basis $\{\alpha_1, \ldots, \alpha_n\}$ *of* V.

Proof. It follows easily from Example 3 of Chapter 8 that $(T, U) \to \text{tr}\,(TU^*)$ is an inner product on $L(V, V)$. Since $f \to T_f$ is an isomorphism, Example 6 of Chapter 8 shows that

$$(f|g) = \text{tr}\,(T_fT_g^*)$$

is an inner product. Now suppose that A and B are the matrices of T_f and T_g in the orthonormal basis $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$. Then

$$A_{jk} = (T_f\alpha_k|\alpha_j) = f(\alpha_k, \alpha_j)$$

and $B_{jk} = (T_g\alpha_k|\alpha_j) = g(\alpha_k, \alpha_j)$. Since AB^* is the matrix of $T_fT_g^*$ in the basis $\mathcal{B}$, it follows that

$$(f|g) = \operatorname{tr}(AB^*) = \sum_{j,k} A_{jk}\overline{B}_{jk}. \quad \blacksquare$$

Definition. *If* f *is a form and* $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$ *an arbitrary ordered basis of* V, *the matrix* A *with entries*

$$A_{jk} = f(\alpha_k, \alpha_j)$$

is called the **matrix of** f **in the ordered basis** $\mathcal{B}$.

When $\mathcal{B}$ is an orthonormal basis, the matrix of f in $\mathcal{B}$ is also the matrix of the linear transformation T_f, but in general this is not the case.

If A is the matrix of f in the ordered basis $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$, it follows that

$$f\left(\sum_s x_s\alpha_s, \sum_r y_r\alpha_r\right) = \sum_{r,s} \overline{y}_r A_{rs} x_s \tag{9-1}$$

for all scalars x_s and y_r $(1 \le r, s \le n)$. In other words, the matrix A has the property that

$$f(\alpha, \beta) = Y^*AX$$

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

The matrix of f in another basis

$$\alpha_j' = \sum_{i=1}^{n} P_{ij}\alpha_i, \qquad (1 \le j \le n)$$

is given by the equation

$$A' = P^*AP. \tag{9-2}$$

For

$$\begin{aligned} A'_{jk} &= f(\alpha_k', \alpha_j') \\ &= f\left(\sum_s P_{sk}\alpha_s, \sum_r P_{rj}\alpha_r\right) \\ &= \sum_{r,s} \overline{P_{rj}} A_{rs} P_{sk} \\ &= (P^*AP)_{jk}. \end{aligned}$$

Since $P^* = P^{-1}$ for unitary matrices, it follows from (9-2) that results concerning unitary equivalence may be applied to the study of forms.

Theorem 2. *Let* f *be a form on a finite-dimensional complex inner product space* V. *Then there is an orthonormal basis for* V *in which the matrix of* f *is upper-triangular.*

Proof. Let T be the linear operator on V such that $f(\alpha, \beta) = (T\alpha|\beta)$ for all α and β. By Theorem 21, there is an orthonormal basis

$\{\alpha_1, \ldots, \alpha_n\}$ in which the matrix of T is upper-triangular. Hence,

$$f(\alpha_k, \alpha_j) = (T\alpha_k|\alpha_j) = 0$$

when $j > k$. ∎

Definition. *A form* f *on a real or complex vector space* V *is called* **Hermitian** *if*

$$\mathrm{f}(\alpha, \beta) = \overline{\mathrm{f}(\beta, \alpha)}$$

for all α and β in V.

If T is a linear operator on a finite-dimensional inner product space V and f is the form

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

then $\overline{f(\beta, \alpha)} = (\alpha|T\beta) = (T^*\alpha|\beta)$; so f is Hermitian if and only if T is self-adjoint.

When f is Hermitian $f(\alpha, \alpha)$ is real for every α, and on complex spaces this property characterizes Hermitian forms.

Theorem 3. *Let* V *be a complex vector space and* f *a form on* V *such that* f(α, α) *is real for every α. Then* f *is Hermitian.*

Proof. Let α and β be vectors in V. We must show that $f(\alpha, \beta) = \overline{f(\beta, \alpha)}$. Now

$$f(\alpha + \beta, \alpha + \beta) = f(\alpha, \beta) + f(\alpha, \beta) + f(\beta, \alpha) + f(\beta, \beta).$$

Since $f(\alpha + \beta, \alpha + \beta)$, $f(\alpha, \alpha)$, and $f(\beta, \beta)$ are real, the number $f(\alpha, \beta) + f(\beta, \alpha)$ is real. Looking at the same argument with $\alpha + i\beta$ instead of $\alpha + \beta$, we see that $-if(\alpha, \beta) + if(\beta, \alpha)$ is real. Having concluded that two numbers are real, we set them equal to their complex conjugates and obtain

$$f(\alpha, \beta) + f(\beta, \alpha) = \overline{f(\alpha, \beta)} + \overline{f(\beta, \alpha)}$$
$$-if(\alpha, \beta) + if(\beta, \alpha) = i\overline{f(\alpha, \beta)} - i\overline{f(\beta, \alpha)}$$

If we multiply the second equation by i and add the result to the first equation, we obtain

$$2f(\alpha, \beta) = 2\overline{f(\beta, \alpha)}.$$ ∎

Corollary. *Let* T *be a linear operator on a complex finite-dimensional inner product space* V. *Then* T *is self-adjoint if and only if* $(\mathrm{T}\alpha|\alpha)$ *is real for every α in* V.

Theorem 4 (Principal Axis Theorem). *For every Hermitian form* f *on a finite-dimensional inner product space* V, *there is an orthonormal basis of* V *in which* f *is represented by a diagonal matrix with real entries.*