Proof. Let T be the linear operator such that $f(\alpha, \beta) = (T\alpha|\beta)$ for all α and β in V. Then, since $f(\alpha, \beta) = \overline{f(\beta, \alpha)}$ and $(\overline{T\beta|\alpha}) = (\alpha|T\beta)$, it follows that

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

for all α and β; hence $T = T^*$. By Theorem 18 of Chapter 8, there is an orthonormal basis of V which consists of characteristic vectors for T. Suppose $\{\alpha_1, \ldots, \alpha_n\}$ is an orthonormal basis and that

$$T\alpha_j = c_j\alpha_j$$

for $1 \leq j \leq n$. Then

$$f(\alpha_k, \alpha_j) = (T\alpha_k|\alpha_j) = \delta_{kj}c_k$$

and by Theorem 15 of Chapter 8 each c_k is real. ∎

Corollary. *Under the above conditions*

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

Exercises

1. Which of the following functions f, defined on vectors $\alpha = (x_1, x_2)$ and $\beta = (y_1, y_2)$ in C^2, are (sesqui-linear) forms on C^2?

(a) $f(\alpha, \beta) = 1$.
(b) $f(\alpha, \beta) = (x_1 - \bar{y}_1)^2 + x_2\bar{y}_2$.
(c) $f(\alpha, \beta) = (x_1 + \bar{y}_1)^2 - (x_1 - \bar{y}_1)^2$.
(d) $f(\alpha, \beta) = x_1\bar{y}_2 - \bar{x}_2y_1$.

2. Let f be the form on R^2 defined by

$$f((x_1, y_1), (x_2, y_2)) = x_1y_1 + x_2y_2.$$

Find the matrix of f in each of the following bases:

$$\{(1, 0), (0, 1)\}, \{(1, -1), (1, 1)\}, \{(1, 2), (3, 4)\}.$$

3. Let

$$A = \begin{bmatrix} 1 & i \\ -i & 2 \end{bmatrix}$$

and let g be the form (on the space of 2×1 complex matrices) defined by $g(X, Y) = Y^*AX$. Is g an inner product?

4. Let V be a complex vector space and let f be a (sesqui-linear) form on V which is symmetric: $f(\alpha, \beta) = f(\beta, \alpha)$. What is f?

5. Let f be the form on R^2 given by

$$f((x_1, x_2), (y_1, y_2)) = x_1y_1 + 4x_2y_2 + 2x_1y_2 + 2x_2y_1.$$

Find an ordered basis in which f is represented by a diagonal matrix.

6. Call the form f (left) **non-degenerate** if 0 is the only vector α such that $f(\alpha, \beta) = 0$ for all β. Let f be a form on an inner product space V. Prove that f is non-degenerate if and only if the associated linear operator T_f (Theorem 1) is non-singular.

7. Let f be a form on a finite-dimensional vector space V. Look at the definition of left non-degeneracy given in Exercise 6. Define right non-degeneracy and prove that the form f is left non-degenerate if and only if f is right non-degenerate.

8. Let f be a non-degenerate form (Exercises 6 and 7) on a finite-dimensional space V. Let L be a linear functional on V. Show that there exists one and only one vector β in V such that $L(\alpha) = f(\alpha, \beta)$ for all α.

9. Let f be a non-degenerate form on a finite-dimensional space V. Show that each linear operator S has an 'adjoint relative to f,' i.e., an operator S' such that $f(S\alpha, \beta) = f(\alpha, S'\beta)$ for all α, β.

9.3. Positive Forms

In this section, we shall discuss non-negative (sesqui-linear) forms and their relation to a given inner product on the underlying vector space.

Definitions. *A form* f *on a real or complex vector space* V *is* **non-negative** *if it is Hermitian and* $f(\alpha, \alpha) \geq 0$ *for every* α *in* V. *The form* f *is* **positive** *if* f *is Hermitian and* $f(\alpha, \alpha) > 0$ *for all* $\alpha \neq 0$.

A positive form on V is simply an inner product on V. A non-negative form satisfies all of the properties of an inner product except that some non-zero vectors may be 'orthogonal' to themselves.

Let f be a form on the finite-dimensional space V. Let $\mathfrak{B} = \{\alpha_1, \ldots, \alpha_n\}$ be an ordered basis for V, and let A be the matrix of f in the basis $\mathfrak{B}$, that is, $A_{jk} = f(\alpha_k, \alpha_j)$. If $\alpha = x_1\alpha_1 + \cdots + x_n\alpha_n$, then

$$\begin{aligned} f(\alpha, \alpha) &= f(\sum_j x_j\alpha_j, \sum_k x_k\alpha_k) \\ &= \sum_j \sum_k x_j\bar{x}_k f(\alpha_j, \alpha_k) \\ &= \sum_j \sum_k A_{kj}x_j\bar{x}_k. \end{aligned}$$

So, we see that f is non-negative if and only if

$$A = A^*$$

and

$$\sum_j \sum_k A_{kj}x_j\bar{x}_k \geq 0 \quad \text{for all scalars} \quad x_1, \ldots, x_n. \tag{9-3}$$

In order that f should be positive, the inequality in (9-3) must be strict for all $(x_1, \ldots, x_n) \neq 0$. The conditions we have derived state that f is a positive form on V if and only if the function

$$g(X, Y) = Y^*AX$$

is a positive form on the space of $n \times 1$ column matrices over the scalar field.

Theorem 5. *Let* F *be the field of real numbers or the field of complex numbers. Let* A *be an* $n \times n$ *matrix over* F. *The function* g *defined by*

$$g(X, Y) = Y^*AX \tag{9-4}$$

is a positive form on the space $F^{n\times 1}$ *if and only if there exists an invertible* $n \times n$ *matrix* P *with entries in* F *such that* $A = P^*P$.

Proof. For any $n \times n$ matrix A, the function g in (9-4) is a form on the space of column matrices. We are trying to prove that g is positive if and only if $A = P^*P$. First, suppose $A = P^*P$. Then g is Hermitian and

$$\begin{aligned} g(X, X) &= X^*P^*PX \\ &= (PX)^*PX \\ &\geq 0. \end{aligned}$$

If P is invertible and $X \neq 0$, then $(PX)^*PX > 0$.

Now, suppose that g is a positive form on the space of column matrices. Then it is an inner product and hence there exist column matrices $Q_1, \ldots, Q_n$ such that

$$\begin{aligned} \delta_{jk} &= g(Q_j, Q_k) \\ &= Q_k^*AQ_j. \end{aligned}$$

But this just says that, if Q is the matrix with columns $Q_1, \ldots, Q_n$, then $Q^*AQ = I$. Since $\{Q_1, \ldots, Q_n\}$ is a basis, Q is invertible. Let $P = Q^{-1}$ and we have $A = P^*P$. ∎

In practice, it is not easy to verify that a given matrix A satisfies the criteria for positivity which we have given thus far. One consequence of the last theorem is that if g is positive then $\det A > 0$, because $\det A = \det (P^*P) = \det P^* \det P = |\det P|^2$. The fact that $\det A > 0$ is by no means sufficient to guarantee that g is positive; however, there are n determinants associated with A which have this property: If $A = A^*$ and if each of those determinants is positive, then g is a positive form.

Definition. *Let* A *be an* $n \times n$ *matrix over the field* F. *The* **principal minors** *of* A *are the scalars* $\Delta_k(A)$ *defined by*

$$\Delta_k(A) = \det \begin{bmatrix} A_{11} & \cdots & A_{1k} \\ \vdots & & \vdots \\ A_{k1} & \cdots & A_{kk} \end{bmatrix}, \qquad 1 \leq k \leq n.$$

Lemma. *Let* A *be an invertible* $n \times n$ *matrix with entries in a field* F. *The following two statements are equivalent.*

(a) *There is an upper-triangular matrix* P *with* $P_{kk} = 1$ $(1 \leq k \leq n)$ *such that the matrix* $B = AP$ *is lower-triangular.*

(b) *The principal minors of* A *are all different from* 0.

Proof. Let P be any $n \times n$ matrix and set $B = AP$. Then

$$B_{jk} = \sum_r A_{jr}P_{rk}.$$

If P is upper-triangular and $P_{kk} = 1$ for every k, then

$$\sum_{r=1}^{k-1} A_{jr}P_{rk} = B_{jk} - A_{kk}, \qquad k > 1.$$

Now B is lower-triangular provided $B_{jk} = 0$ for $j < k$. Thus B will be lower-triangular if and only if

$$\text{(9-5)} \qquad \sum_{r=1}^{k-1} A_{jr}P_{rk} = -A_{kk}, \qquad \begin{matrix} 1 \leq j \leq k-1 \\ 2 \leq k \leq n. \end{matrix}$$

So, we see that statement (a) in the lemma is equivalent to the statement that there exist scalars P_{rk}, $1 \leq r \leq k$, $1 \leq k \leq n$, which satisfy (9-5) and $P_{kk} = 1$, $1 \leq k \leq n$.

In (9-5), for each $k > 1$ we have a system of $k - 1$ linear equations for the unknowns $P_{1k}, P_{2k}, \ldots, P_{k-1,k}$. The coefficient matrix of that system is

$$\begin{bmatrix} A_{11} & \cdots & A_{1,k-1} \\ \vdots & & \vdots \\ A_{k-1} & \cdots & A_{k-1,k-1} \end{bmatrix}$$

and its determinant is the principal minor $\Delta_{k-1}(A)$. If each $\Delta_{k-1}(A) \neq 0$, the systems (9-5) have unique solutions. We have shown that statement (b) implies statement (a) and that the matrix P is unique.

Now suppose that (a) holds. Then, as we shall see,

$$\text{(9-6)} \qquad \begin{aligned} \Delta_k(A) &= \Delta_k(B) \\ &= B_{11}B_{22} \cdots B_{kk}, \qquad k = 1, \ldots, n. \end{aligned}$$

To verify (9-6), let $A_1, \ldots, A_n$ and $B_1, \ldots, B_n$ be the columns of A and B, respectively. Then

$$\text{(9-7)} \qquad \begin{aligned} B_1 &= A_1 \\ B_r &= \sum_{j=1}^{r-1} P_{jr}A_j + A_r, \qquad r > 1. \end{aligned}$$

Fix k, $1 \leq k \leq n$. From (9-7) we see that the rth column of the matrix

$$\begin{bmatrix} B_{11} & \cdots & B_{kk} \\ \vdots & & \vdots \\ B_{k1} & \cdots & B_{kk} \end{bmatrix}$$

is obtained by adding to the rth column of

$$\begin{bmatrix} A_{11} & \cdots & A_{1k} \\ \vdots & & \vdots \\ A_{k1} & \cdots & A_{kk} \end{bmatrix}$$

a linear combination of its other columns. Such operations do not change determinants. That proves (9-6), except for the trivial observation that because B is triangular $\Delta_k(B) = B_{11} \cdots B_{kk}$. Since A and P are invertible, B is invertible. Therefore,

$$\Delta(B) = B_{11} \cdots B_{nn} \neq 0$$

and so $\Delta_k(A) \neq 0$, $k = 1, \ldots, n$. ∎

Theorem 6. *Let* f *be a form on a finite-dimensional vector space* V *and let* A *be the matrix of* f *in an ordered basis* $\mathcal{B}$. *Then* f *is a positive form if and only if* A = A* *and the principal minors of* A *are all positive.*

Proof. Let's do the interesting half of the theorem first. Suppose that $A = A^*$ and $\Delta_k(A) > 0$, $1 \leq k \leq n$. By the lemma, there exists an (unique) upper-triangular matrix P with $P_{kk} = 1$ such that $B = AP$ is lower-triangular. The matrix P^* is lower-triangular, so that $P^*B = P^*AP$ is also lower-triangular. Since A is self-adjoint, the matrix $D = P^*AP$ is self-adjoint. A self-adjoint triangular matrix is necessarily a diagonal matrix. By the same reasoning which led to (9-6),

$$\begin{aligned} \Delta_k(D) &= \Delta_k(P^*B) \\ &= \Delta_k(B) \\ &= \Delta_k(A). \end{aligned}$$

Since D is diagonal, its principal minors are

$$\Delta_k(D) = D_{11} \cdots D_{kk}.$$

From $\Delta_k(D) > 0$, $1 \leq k \leq n$, we obtain $D_{kk} > 0$ for each k.

If A is the matrix of the form f in the ordered basis $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$, then $D = P^*AP$ is the matrix of f in the basis $\{\alpha_1', \ldots, \alpha_n'\}$ defined by

$$\alpha_j' = \sum_{i=1}^{n} P_{ij}\alpha_i.$$

See (9-2). Since D is diagonal with positive entries on its diagonal, it is obvious that

$$X^*DX > 0, \qquad X \neq 0$$

from which it follows that f is a positive form.

Now, suppose we start with a positive form f. We know that $A = A^*$. How do we show that $\Delta_k(A) > 0$, $1 \leq k \leq n$? Let V_k be the subspace spanned by $\alpha_1, \ldots, \alpha_k$ and let f_k be the restriction of f to $V_k \times V_k$. Evidently f_k is a positive form on V_k and, in the basis $\{\alpha_1, \ldots, \alpha_k\}$ it is represented by the matrix

$$\begin{bmatrix} A_{11} & \cdots & A_{1k} \\ \vdots & & \vdots \\ A_{k1} & \cdots & A_{kk} \end{bmatrix}.$$

As a consequence of Theorem 5, we noted that the positivity of a form implies that the determinant of any representing matrix is positive. ▌

There are some comments we should make, in order to complete our discussion of the relation between positive forms and matrices. What is it that characterizes the matrices which represent positive forms? If f is a form on a complex vector space and A is the matrix of f in some ordered basis, then f will be positive if and only if $A = A^*$ and

$$X^*AX > 0, \qquad \text{for all complex } X \neq 0. \tag{9-8}$$

It follows from Theorem 3 that the condition $A = A^*$ is redundant, i.e., that (9-8) implies $A = A^*$. On the other hand, if we are dealing with a real vector space the form f will be positive if and only if $A = A^t$ and

$$X^tAX > 0, \qquad \text{for all real } X \neq 0. \tag{9-9}$$

We want to emphasize that if a real matrix A satisfies (9-9), it does not follow that $A = A^t$. One thing which is true is that, *if* $A = A^t$ and (9-9) holds, then (9-8) holds as well. That is because

$$\begin{aligned}(X + iY)^*A(X + iY) &= (X^t - iY^t)A(X + iY) \\ &= X^tAX + Y^tAY + i[X^tAY - Y^tAX]\end{aligned}$$

and if $A = A^t$ then $Y^tAX = X^tAY$.

If A is an $n \times n$ matrix with complex entries and if A satisfies (9-9), we shall call A a **positive matrix.** The comments which we have just made may be summarized by saying this: In either the real or complex case, a form f is positive if and only if its matrix in some (in fact, every) ordered basis is a positive matrix.

Now suppose that V is a finite-dimensional inner product space. Let f be a non-negative form on V. There is a unique self-adjoint linear operator T on V such that

$$f(\alpha, \beta) = (T\alpha|\beta). \tag{9-10}$$

and T has the additional property that $(T\alpha|\alpha) \geq 0$.

Definition. *A linear operator* T *on a finite-dimensional inner product space* V *is* **non-negative** *if* $T = T^*$ *and* $(T\alpha|\alpha) \geq 0$ *for all* α *in* V. *A* **positive** *linear operator is one such that* $T = T^*$ *and* $(T\alpha|\alpha) > 0$ *for all* $\alpha \neq 0$.