Lecture no 11: BILINEAR, QUADRATIC, AND HERMITIAN FORMS[1]

This chapter generalizes the notions of linear mappings and linear functionals. Specifically, we introduce the notion of a *bilinear form*. These bilinear maps also give rise to *quadratic* and *Hermitian* forms. Although the field K is arbitrary, we will later specialize to the cases $K = \mathbf{R}$ and $K = \mathbf{C}$.

1. Bilinear Forms

Let V be a vector space of finite dimension over a field K.

Definition: A *bilinear form* on V is a mapping $f: V \times V \to K$ such that, for all $a, b \in K$ and all $u_i, v_i \in V$:

(i) $f(au_1 + bu_2, \; v) = af(u_1, v) + bf(u_2, v)$,

(ii) $f(u, \; av_1 + bv_2) = af(u, v_1) + bf(u, v_2)$

Note: We express condition (i) by saying f is linear in the first variable, and condition (ii) by saying f is linear in the second variable.

Example

(a) Let f be the dot product on $\mathbf{R}^n$; that is, for $u = (a_i)$ and $v = (b_i)$,

$$f(u, v) = u \cdot v = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n$$

Then f is a bilinear form on $\mathbf{R}^n$. (In fact, any inner product on a real vector space V is a bilinear form on V.)

(b) Let ϕ and σ be arbitrarily linear functionals on V. Let $f: V \times V \to K$ be defined by $f(u, v) = \phi(u)\sigma(v)$. Then f is a bilinear form, because ϕ and σ are each linear.

(c) Let $A = [a_{ij}]$ be any $n \times n$ matrix over a field K. Then A may be identified with the following bilinear form F on K^n, where $X = [x_i]$ and $Y = [y_i]$ are column vectors of variables:

$$f(X, Y) = X^T A Y = \sum_{i,j} a_{ij} x_i y_i = a_{11} x_1 y_1 + a_{12} x_1 y_2 + \cdots + a_{nn} x_n y_n$$

Space of Bilinear Forms

Let $B(V)$ denote the set of all bilinear forms on V. A *vector space* structure is placed on $B(V)$, where for any $f, g \in B(V)$ and any $k \in K$, we define $f + g$ and $k f$ as follows:

$$(f + g)(u,v) = f(u,v) + g(u,v) \text{ and } (k f)(u,v) = k f(u,v)$$

[1]

11	Л	2	Лінійний і білінійний функціонал. Білінійна і квадратична форма.
12	ПР	2	Матриця квадратичної форми. Перетворення матриці квадратичної форми при переході до нового базису.

The following theorem applies.

Theorem: Let V be a vector space of dimension n over K. Let $\{\phi_1, ..., \phi_n\}$ be any basis of the dual space V^*. Then $\{ f_{ij}: i, j = 1, ..., n\}$ is a basis of $B(V)$, where f_{ij} is defined by $f_{ij}(u, v) = \phi_i(u)\phi_j(v)$. Thus, in particular, $\dim B(V) = n^2$.

Bilinear Forms and Matrices

Let f be a bilinear form on V and let $S = \{u_1, ..., u_n\}$ be a basis of V. Suppose $u, v \in V$ and

$$u = a_1u_1 + \ldots + a_nu_n \text{ and } v = b_1u_1 + ... + b_nu_n$$

Then

$$f(u, v) = f(a_1u_1 + \ldots + a_nu_n, b_1u_1 + ... + b_nu_n) = \sum_{i,j} a_ib_jf\left(u_iu_j\right)$$

Thus, f is completely determined by the n^2 values $f(u_i, u_j)$.

The matrix $A = [a_{ij}]$ where $a_{ij} = f(u_i, u_j)$ is called the *matrix representation* of f relative to the basis S or, simply, the "matrix of f in S." It "represents" f in the sense that, for all $u, v \in V$,

$$f(u, v) = \sum_{i,j} a_ib_jf\left(u_iu_j\right) = [u]_S^TA[v]_S$$

Note: As usual, $[u]_S$ denotes the coordinate (column) vector of u in the basis S.

Change of Basis

We now ask, how does a matrix representing a bilinear form transform when a new basis is selected? The answer is given in the following theorem

Theorem: Let P be a *change-of-basis matrix* from one basis S to another basis S_0. If A is the matrix representing a *bilinear form* f in the original basis S, then $B = P^TA\,P$ is the *matrix representing* f in the *new basis* S_0.

The above theorem motivates the following definition.

Definition: A matrix B is *congruent* to a matrix A, written $B \cong A$, if there exists a nonsingular matrix P such that $B = P^TAP$.

Thus, by Theorem, matrices representing the same bilinear form are *congruent*. We remark that congruent matrices have the same rank, because P and P^T are nonsingular; hence, the following definition is well defined.

Definition: The rank of a bilinear form f on V, written rank(f), is the rank of any matrix representation of f. We say f is *degenerate* or *nondegenerate* according to whether rank(f) < dim V or rank(f) > dim V.

Alternating Bilinear Forms

Definition: Let f be a bilinear form on V. Then f is called

(i) *alternating* if $f(v, v) = 0$ for every $v \in V$;

(ii) *skew-symmetric* if $f(u, v) = -f(v, u)$ for every $u, v \in V$.

The main structure theorem of alternating bilinear forms is as follows.

Theorem: Let f be an alternating bilinear form on V. Then there exists a basis of V in which f is represented by a block diagonal matrix M of the form

$$M = \operatorname{diag}\left(\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \ldots, \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, [0], [0], \ldots [0]\right)$$

Moreover, the number of nonzero blocks is uniquely determined by f [because it is equal to ½ rank(f)].

2. Symmetric Bilinear Forms, Quadratic Forms

Definition: Let f be a bilinear form on V. Then f is said to be *symmetric* if, for every u, $v \in V$, $f(u, v) = f(v, u)$. One can easily show that f is *symmetric* if and only if any matrix representation A of f is a *symmetric matrix*.

The main result for symmetric bilinear forms is as follows.

Theorem: Let f be a symmetric bilinear form on V. Then V has a basis $\{v_1, \ldots, v_n\}$ in which f is represented by a diagonal matrix—that is, where $f(v_i, v_j) = 0$ for $i \neq j$.

Theorem: (Alternative Form) Let A be a symmetric matrix over K. Then A is *congruent* to a *diagonal matrix*, that is, there exists a nonsingular matrix P such that P^TAP is *diagonal*.

Diagonalization Algorithm

One way of obtaining the diagonal form $D = P^T AP$ is by a sequence of elementary row operations and the same sequence of elementary column operations. This same sequence of elementary row operations on the identity matrix I will yield P^T. This algorithm is formalized below.

***ALGORITHM* 1**: (Congruence Diagonalization of a Symmetric Matrix) The input is a symmetric matrix $A = [a_{ij}]$ of order n

Step 1. Form the $n \times 2n$ (block) matrix $M = [A_1, I]$, where $A_1 = A$ is the left half of M and the identity matrix I is the right half of M.

Step 2. Examine the entry a_{11}. There are three cases.

Case I: $a_{11} \neq 0$. (Use a_{11} as a pivot to put 0's below a_{11} in M and to the right of a_{11} in A_1.) For $i = 2, \ldots, n$:

(a) Apply the row operation "Replace R_i by $-a_{i1}R_1 + a_{11}R_i$."

(b) Apply the corresponding column operation "Replace C_i by $-a_{i1}C_1 + a_{11}C_i$."

These operations reduce the matrix M to the form

$$M \sim \begin{bmatrix} a_{11} & 0 & * & * \\ 0 & A_1 & * & * \end{bmatrix} \qquad (*)$$

Case II: $a_{11} = 0$ but $a_{kk} \neq 0$, for some $k > 1$.

(a) Apply the row operation "Interchange R_1 and R_k."

(b) Apply the corresponding column operation "Interchange C_1 and C_k."

(These operations bring a_{kk} into the first diagonal position, which reduces the matrix to Case I.)

Case III: All diagonal entries $a_{ii} = 0$ but some $a_{ij} \neq 0$.

(a) Apply the row operation "Replace R_i by $R_j + R_i$."

(b) Apply the corresponding column operation "Replace C_i by $C_j + C_i$."

(These operations bring $2a_{ij}$ into the ith diagonal position, which reduces the matrix to Case II.)

Thus, M is finally reduced to the form $(*)$, where A_2 is a symmetric matrix of order less than A.

Step 3. Repeat Step 2 with each new matrix A_k (by neglecting the first row and column of the preceding matrix) until A is diagonalized. Then M is transformed into the form $M' = [D, Q]$, where D is diagonal.

Step 4. Set $P = Q^T$. Then $D = P^T AP$.

Remark: We emphasize that in **Step 2**, the row operations will change both sides of M, but the column operations will only change the left half of M.

Example Let $A = \begin{bmatrix} 1 & 2 & -3 \\ 2 & 5 & -4 \\ -3 & -4 & 8 \end{bmatrix}$. Apply Algorithm to find a nonsingular matrix P such that $D = P^T AP$ is diagonal.

First form the block matrix $M = [A, I]$; that is, let

$$M = [A, I] = \left[\begin{array}{rrr|rrr} 1 & 2 & -3 & 1 & 0 & 0 \\ 2 & 5 & -4 & 0 & 1 & 0 \\ -3 & -4 & 8 & 0 & 0 & 1 \end{array}\right]$$

Apply the row operations "Replace R_2 by $-2R_1 + R_2$" and "Replace R_3 by $3R_1 + R_3$" to M, and then apply the corresponding column operations "Replace C_2 by $-2C_1 + C_2$" and "Replace $C3$ by $3C_1 + C_3$" to obtain

$$\left[\begin{array}{rrr|rrr} 1 & 2 & -3 & 1 & 0 & 0 \\ 0 & 1 & 2 & -2 & 1 & 0 \\ 0 & 2 & -1 & 3 & 0 & 1 \end{array}\right] \quad \text{and then} \quad \left[\begin{array}{rrr|rrr} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 2 & -2 & 1 & 0 \\ 0 & 2 & -1 & 3 & 0 & 1 \end{array}\right]$$

Next apply the row operation "Replace R_3 by $-2R_2 + R_3$" and then the corresponding column operation "Replace $C3$ by $-2C_2 + C_3$" to obtain

$$\left[\begin{array}{rrr|rrr} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 2 & -2 & 1 & 0 \\ 0 & 0 & -5 & 7 & -2 & 1 \end{array}\right] \quad \text{and then} \quad \left[\begin{array}{rrr|rrr} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & -2 & 1 & 0 \\ 0 & 0 & -5 & 7 & -2 & 1 \end{array}\right]$$

Now A has been diagonalized. Set

$$P = \begin{bmatrix} 1 & -2 & 7 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{and then} \quad D = P^{-1}AP = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -5 \end{bmatrix}$$

We emphasize that P is the transpose of the right half of the final matrix

Quadratic Forms

Expressions of the form

$$a_1x_1 + a_2x_2 + \dots + a_nx_n$$

occurred in our study of linear equations and linear systems. If $a_1, a_2, \dots, a_n$ are treated as fixed constants, then this expression is a real-valued function of the n variables $x_1, x_2, \dots, x_n$ and is called a *linear form* on $\mathbf{R}^n$. All variables in a linear form occur to the first power and there are no products of variables.

Here we will be concerned with *quadratic forms* on $\mathbf{R}^n$, which are functions of the form

$$a_1x_1^2 + a_2x_2^2 + \ldots + a_nx_n^2 + \text{(all possible terms } a_kx_ix_j \text{ in which } i \neq j)$$

The terms of the form $a_kx_ix_j$ are called *cross product terms*. It is common to combine the cross product terms involving x_ix_j with those involving x_jx_i to avoid duplication.

Thus, a general quadratic form on $\mathbf{R}^2$ would typically be expressed as

$$a_1x_1^2 + a_2x_2^2 + 2a_3x_1x_2 \tag{1}$$

and a general quadratic form on $\mathbf{R}^3$ as

$$a_1x_1^2 + a_2x_2^2 + a_3x_3^2 + 2a_4x_1x_2 + 2a_5x_1x_3 + 2a_6x_2x_3 \tag{2}$$

The definition of the quadratic forms in a general form:

Definition A: A mapping q: $V \rightarrow K$ is a *quadratic form* if $q(v) = f(v, v)$ for some symmetric bilinear form f on V;

or

Definition B: A quadratic form q in variables $x_1, x_2, \ldots, x_n$ is a *polynomial* such that every term has degree two, that is

$$q(x_1, x_2, \ldots, x_n) = \sum_i c_i x_i^2 + \sum_{i<j} d_{ij} x_i x_j$$

Remark: The quadratic form q in Definition B determines a symmetric matrix $A = [a_{ij}]$ where $a_{ii} = c_i$ and $a_{ij} = a_{ji} = 1/2d_{ij}$.

If we let $\boldsymbol{x}$ be the column vector of variables, then (1) and (2) can be expressed in matrix form as

$$\begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} a_1 & a_3 \\ a_3 & a_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \mathbf{x}^T A \mathbf{x}$$

$$\begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} a_1 & a_4 & a_5 \\ a_4 & a_2 & a_6 \\ a_5 & a_6 & a_3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \mathbf{x}^T A \mathbf{x}$$

Note that the matrix A in these formulas is *symmetric*, that its *diagonal entries* are the coefficients of the squared terms, and its *off-diagonal entries* are half the coefficients of the cross product terms. In general, if A is a symmetric $n \times n$ matrix and $\boldsymbol{x}$ is an $n \times 1$ column vector of variables, then we call the function

$$Q_A(\boldsymbol{x}) = \boldsymbol{x}^T A \boldsymbol{x} \tag{3}$$

the quadratic form associated with A. When convenient, (3) can be expressed in *dot product notation* as

$$Q_A(\boldsymbol{x}) = \boldsymbol{x} \cdot A\boldsymbol{x} = A\boldsymbol{x} \cdot \boldsymbol{x}$$

In the case where A is a *diagonal matrix*, the quadratic form $\boldsymbol{x}^T A \boldsymbol{x}$ has no cross product terms, for example, if A has diagonal entries $\lambda_1, \lambda_2, \ldots, \lambda_n$, then

$$\mathbf{x}^T A\mathbf{x} = [x_1 \quad x_2 \quad \cdots \quad x_n] \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \lambda_1 x_1^2 + \lambda_2 x_2^2 + \cdots + \lambda_n x_n^2$$

The Principal AxesTheorem

Many of the techniques for solving these problems are based on simplifying the quadratic form $\boldsymbol{x}^T A\boldsymbol{x}$ by making a substitution

$$\boldsymbol{x} = P\boldsymbol{y} \tag{4}$$

that is,

to expresses the variables $x_1, x_2, \ldots, x_n$ in terms of new variables $y_1, y_2, \ldots, y_n$.

> If P is *invertible*, then we call (4) a *change of variable*, and if P is *orthogonal*, then we call (4) an *orthogonal change of variable*.

If we make the change of variable $\boldsymbol{x} = P\boldsymbol{y}$ in the quadratic form $\boldsymbol{x}^T A\boldsymbol{x}$, then we obtain

$$\boldsymbol{x}^T A\boldsymbol{x} = (P\,\boldsymbol{y})^T A(P\,\boldsymbol{y}) = \boldsymbol{y}^T P^T AP\boldsymbol{y} = \boldsymbol{y}^T(P^T AP)\boldsymbol{y} \tag{5}$$

Since the matrix $B = P^T AP$ is symmetric (verify), the effect of the change of variable is to produce a *new quadratic form* $\boldsymbol{y}^T B\boldsymbol{y}$ in the variables $y_1, y_2, \ldots, y_n$. In particular, if we choose P to *orthogonally diagonalize* A, then the new quadratic form will be $\boldsymbol{y}^T D\boldsymbol{y}$, where D is a diagonal matrix with the eigenvalues of A on the main diagonal, that is,

$$\mathbf{x}^T A\mathbf{x} = \mathbf{y}^T D\mathbf{y} = [y_1 \quad y_2 \quad \cdots \quad y_n] \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$
$$= \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2$$

Thus, we have the following result, called *the principal axes theorem*.

Theorem: If $\boldsymbol{A}$ is a symmetric $n \times n$ matrix, then there is an *orthogonal change of variable* that transforms the quadratic form $\boldsymbol{x}^T A\boldsymbol{x}$ into a quadratic form $\boldsymbol{y}^T D\boldsymbol{y}$ with no cross product terms. Specifically, if P *orthogonally diagonalizes* A, then making the change of variable $\boldsymbol{x} = P\boldsymbol{y}$ in the quadratic form $\boldsymbol{x}^T A\boldsymbol{x}$ yields the quadratic form

$$\boldsymbol{x}^T A\boldsymbol{x} = \boldsymbol{y}^T D\boldsymbol{y} = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \ldots + \lambda_n y_n^2$$

in which $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the eigenvalues of A corresponding to the eigenvectors that form the successive columns of P.

Remark. If A is a symmetric $n \times n$ matrix, then the quadratic form $\boldsymbol{x}^T A\boldsymbol{x}$ is a real-valued function whose range is the set of all possible values for $\boldsymbol{x}^T A\boldsymbol{x}$ as x varies over $\mathbf{R}^n$. It can be shown that an orthogonal change of variable $\boldsymbol{x} = P\boldsymbol{y}$ does not alter the range of a quadratic form, that is, the set of all values for $\boldsymbol{x}^T A\boldsymbol{x}$ as $\boldsymbol{x}$ varies over $\mathbf{R}^n$ is the same as the set of all values for $\boldsymbol{y}^T(P^T AP)\boldsymbol{y}$ as $\boldsymbol{y}$ varies over $\mathbf{R}^n$.