4 Quadratic Forms II: Symmetric bi-linear functions

Prof.Dr.P.M.Bajracharya

School of Mathematical Sciences
T.U., Kirtipur

April 25, 2024

Summary

1. Symmetric bi-linear functions
2. Change of Variable in a Quadratic Form
3. Connections with conic sections

Symmetric bi-linear functions

- We know that any matrix can be viewed as a representation of a linear transformation with respect to a basis.

- We know that any matrix can be viewed as a representation of a linear transformation with respect to a basis.
- We can also associate to any symmetric matrix something that is known as a **symmetric bilinear transformation**.

Let's start with the definition.

Symmetric bilinear function

Let V be a vector space, e.g., $\mathbb{R}^2$. Then a **symmetric bilinear function** on V is defined as a mapping $B : V \times V \to \mathbb{R}$ such that for any $u, v, w \in V$

(a) $B(u, v) = B(v, u)$

(b) $\forall\ a, b \in \mathbb{R} \quad B(au + bv, w) = aB(u, w) + bB(v, w)$.

Remark.

1. The first condition says that B is symmetric in its two arguments.
2. By combining these two conditions we also have that B is linear in its second argument.

Suppose that A is an $n \times n$ matrix. For $v, w \in \mathbb{R}^n$ we will define the function

$$f(v, w) = v^T A w \in \mathbb{R}.$$

Suppose that A is an $n \times n$ matrix. For $v, w \in \mathbb{R}^n$ we will define the function

$$f(v, w) = v^T A w \in \mathbb{R}.$$

The next two theorems asserts that this is a bilinear form, which might give us a hint to a similarity between this bilinear form and the linear transformations.

Theorem (Symmetricity)

If A is a symmetric $n \times n$ matrix, and $B_A(v, w) = v^T A w$, then

$$B_A(v, w) = B_A(w, v).$$

Proof.

(a) We know that $v^T A w \in \mathbb{R}$. So,

$$v^T A w = (v^T A w)^T.$$

This implies that

$$\begin{aligned} B_A(v,w) &= (v^T A w)^T \\ &= (Aw)^T v \\ &= w^T A^T v \\ &= w^T A v \qquad \text{(because } A \text{ is symmetric)} \\ &= B_A(w,v). \end{aligned}$$

$\square$

Theorem (Linearity)

If A is a symmetric $n \times n$ matrix and $B_A(v, w) = v^T A w$, then $B_A(v, w)$ is linear in the first variable v.

Proof. For any real number a

$$\begin{aligned} B_A(av, w) &= (av)^T Aw \\ &= a(v)^T Aw \\ &= a(v^T Aw) \\ &= aB_A(v, w). \end{aligned}$$

Proof... Also,

$$\begin{aligned} B_A(u+v,w) &= (u+v)^T Aw \\ &= (u^T + v^T)Aw \\ &= u^T Aw + v^T Aw \\ &= B_A(u,w) + B_A(v,w). \end{aligned}$$

Therefore, $B_A(v,w) = v^T Aw$ is linear in the first variable v. □

The next theorem claims that a symmetric bilinear function can uniquely be expressed as a product of three matrices.

Theorem.

If B is a symmetric bilinear function on $\mathbb{R}^n$, then it is of the form $B = B_A(v, w) = v^T A w$, for some unique symmetric matrix A.

Proof.

Let $v = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}, w = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} \in \mathbb{R}^n$. Then we can write

$$v = \sum_{i=1}^{n} v_i e_i, \quad w = \sum_{j=1}^{n} w_j e_j,$$

where $\{e_1, ..., e_n\}$ is a basis for $\mathbb{R}^n$.

Proof... Using the properties of bilinear functions, we have

$$B(v,w) = B\left(\sum_{i=1}^{n} v_i e_i, \sum_{j=1}^{n} w_j e_j\right)$$
$$= \sum_{i=1}^{n} v_i \sum_{j=1}^{n} w_j B(e_i, e_j)$$

Proof...

$$
\begin{aligned}
B(v,w) &= \sum_{i=1}^{n} v_i \sum_{j=1}^{n} B(e_i,e_j) w_j \\
&= \sum_{i=1}^{n} v_i \begin{pmatrix} B(e_1,e_1) & \cdots & B(e_1,e_n) \\ \vdots & \vdots & \vdots \\ B(e_n,e_1) & \cdots & B(e_n,e_n) \end{pmatrix} \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}
\end{aligned}
$$

Proof...

$$
\begin{aligned}
B(v,w) &= \sum_{i=1}^{n} v_i \sum_{j=1}^{n} B(e_i,e_j) w_j \\
&= \sum_{i=1}^{n} v_i \begin{pmatrix} B(e_1,e_1) & \cdots & B(e_1,e_n) \\ \vdots & \vdots & \vdots \\ B(e_n,e_1) & \cdots & B(e_n,e_n) \end{pmatrix} \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} \\
&= \begin{pmatrix} v_1 & \dots & v_n \end{pmatrix} \begin{pmatrix} B(e_1,e_1) & \cdots & B(e_1,e_n) \\ \vdots & \vdots & \vdots \\ B(e_n,e_1) & \cdots & B(e_n,e_n) \end{pmatrix} \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}
\end{aligned}
$$

Proof...

$$\begin{aligned}
B(v,w) &= \sum_{i=1}^{n} v_i \sum_{j=1}^{n} B(e_i,e_j) w_j \\
&= \sum_{i=1}^{n} v_i \begin{pmatrix} B(e_1,e_1) & \cdots & B(e_1,e_n) \\ \vdots & \vdots & \vdots \\ B(e_n,e_1) & \cdots & B(e_n,e_n) \end{pmatrix} \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} \\
&= (v_1 \ \dots \ v_n) \begin{pmatrix} B(e_1,e_1) & \cdots & B(e_1,e_n) \\ \vdots & \vdots & \vdots \\ B(e_n,e_1) & \cdots & B(e_n,e_n) \end{pmatrix} \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}
\end{aligned}$$

Setting $A = (B(e_i, e_j))_{n\times n}$, we have

$$B = v^T A w.$$

Proof... To prove the uniqueness of the matrix A, we assume that there is another matrix C such that

$$B = v^T C w.$$

Proof... To prove the uniqueness of the matrix A, we assume that there is another matrix C such that

$$B = v^T C w.$$

Then

$$v^T A w - v^T C w = 0 \Rightarrow v^T (A - C) w = 0$$

Proof... To prove the uniqueness of the matrix A, we assume that there is another matrix C such that

$$B = v^T C w.$$

Then

$$v^T A w - v^T C w = 0 \Rightarrow v^T (A - C) w = 0$$

Put $D = A - C$.

Proof... To prove the uniqueness of the matrix A, we assume that there is another matrix C such that

$$B = v^T C w.$$

Then

$$v^T A w - v^T C w = 0 \Rightarrow v^T (A - C) w = 0$$

Put $D = A - C$. Thus, for any $v, w \in \mathbb{R}^n$

$$v^T D w = 0.$$

Proof... In particular, if we choose $v, w \in \mathbb{R}^n$ such that $v_i = w_i = 1$ for all $i \in \{1, ..., n\}$, then we still have

$$v^T D w = 0.$$

Proof... In particular, if we choose $v, w \in \mathbb{R}^n$ such that $v_i = w_i = 1$ for all $i \in \{1, ..., n\}$, then we still have

$$v^T D w = 0.$$

This implies that

$$\sum_{i=1}^{n} v_i \sum_{j=1}^{n} D_{ij} w_j = 0$$

Proof... In particular, if we choose $v, w \in \mathbb{R}^n$ such that $v_i = w_i = 1$ for all $i \in \{1, ..., n\}$, then we still have

$$v^T D w = 0.$$

This implies that

$$\sum_{i=1}^{n} v_i \sum_{j=1}^{n} D_{ij} w_j = 0$$

$$\Rightarrow \quad \sum_{i=1}^{n} \sum_{j=1}^{n} v_i D_{ij} w_j = 0.$$

Proof... We now see that $v_i D_{ij} w_j = D_{ij}$ for all $i \in \{1, ..., n\}$.

Proof... We now see that $v_i D_{ij} w_j = D_{ij}$ for all $i \in \{1, ..., n\}$. Therefore, we must have

$$D_{ij} = 0 \quad \text{for all } i \in \{1, ..., n\}.$$

Proof... We now see that $v_i D_{ij} w_j = D_{ij}$ for all $i \in \{1, ..., n\}$. Therefore, we must have

$$D_{ij} = 0 \quad \text{for all } i \in \{1, ..., n\}.$$

That means,

$$D = 0$$

Proof... We now see that $v_i D_{ij} w_j = D_{ij}$ for all $i \in \{1, ..., n\}$. Therefore, we must have

$$D_{ij} = 0 \quad \text{for all } i \in \{1, ..., n\}.$$

That means,

$$\begin{aligned} & D = 0 \\ \Rightarrow \qquad & A - C = 0 \end{aligned}$$

Proof... We now see that $v_i D_{ij} w_j = D_{ij}$ for all $i \in \{1, ..., n\}$. Therefore, we must have

$$D_{ij} = 0 \quad \text{for all } i \in \{1, ..., n\}.$$

That means,

$$\begin{aligned} & D = 0 \\ \Rightarrow \quad & A - C = 0 \\ \Rightarrow \quad & A = C. \end{aligned}$$

$\square$

Change of Variable in a Quadratic Form

- In some cases, quadratic forms are easier to use when they have no cross-product terms – that is, when the matrix of the quadratic form is a diagonal matrix.
- Fortunately, the cross-product term can be eliminated by making a suitable change of variable.

If $x \in R^n$, then a change of variable is an equation of the form

$$x = Py, \text{ or equivalently, } y = P^{-1}x,$$

where P is an invertible matrix and y is a new variable vector in R^n. Here y is the coordinate vector of x relative to the basis of $\mathbb{R}^n$ determined by the columns of P.

Example

Consider the quadratic form

$$Q(x) = x_1^2 - 8x_1x_2 - 5x_2^2. \tag{1}$$

1. Determine the matrix A of Quadratic Form (1).
2. Find the matrix V of the eigenvectors of A.
3. Make a change of variable in (1)

$$x = Py, \text{ where } x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \text{ and } y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}.$$

   Find (1) in terms of y.

Example ...

4. Compute $Q(x)$ for $x = (2, -2)$.
5. Compute $Q(x)$ for $x = (2, -2)$ using new variable y.

Solution. The matrix of the quadratic form is

$$A = \begin{pmatrix} 1 & -4 \\ -4 & -5 \end{pmatrix}$$

The first step is to orthogonally diagonalize A. Its eigenvalues turn out to be $\lambda = 3$ and $\lambda = 7$. Associated unit eigenvectors are

$$\lambda = 3 : \begin{pmatrix} 2/\sqrt{5} \\ -1/\sqrt{5} \end{pmatrix};\ \lambda = 7 : \begin{pmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{pmatrix}.$$

These vectors are orthogonal (because they correspond to distinct eigenvalues) and so provide an orthonormal basis for $\mathbb{R}^2$.

Let

$$P = \begin{pmatrix} 2/\sqrt{5} & -1/\sqrt{5} \\ 1/\sqrt{5} & 2/\sqrt{5} \end{pmatrix}, \quad D = \begin{pmatrix} 3 & 0 \\ 0 & -7 \end{pmatrix}.$$

Then $A = PDP^{-1}$ and $D = P^{-1}AP = P^TAP$, as pointed out earlier. A suitable change of variable is

$$x = Py, \text{ where } x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \text{ and } y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}.$$

Then

$$\begin{aligned} x_1^2 - 8x_1x_2 - 5x_2^2 &= x^TAx = (Py)^TA(Py) \\ &= y^TP^TAPy = y^TDy \\ &= 3y_1^2 - 7y_2^2. \end{aligned}$$

□

To illustrate the meaning of the equality of quadratic forms in Example 4, we can compute $Q(x)$ for $x = (2, -2)$ using the new quadratic form. First, since $x = Py$,

$$y = P^{-1}x = P^T x$$

so

$$y = \begin{pmatrix} 2/\sqrt{5} & -1/\sqrt{5} \\ 1/\sqrt{5} & 2/\sqrt{5} \end{pmatrix} \begin{pmatrix} 2 \\ -2 \end{pmatrix} = \begin{pmatrix} 6/\sqrt{5} \\ -2/\sqrt{5} \end{pmatrix}$$

Hence

$$\begin{aligned} 3y_1^2 - 7y_2^2 &= 3(6/\sqrt{5})_1^2 - 7(-2/\sqrt{5})^2 \\ &= 3(36/5)/ - 7(4/5) \\ &= 80/5 = 16. \end{aligned}$$

This is the value of $Q(x)$ in Example 3 when $x = (2, -2)$. See Figure

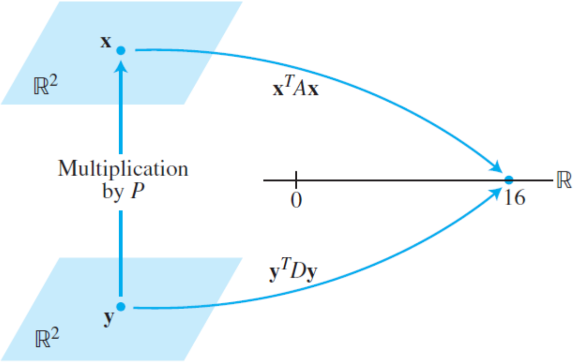

Figure: Change of variable in $x^T Ax$.

From the previous discussion, we can rewrite the above example as follows:

Example

- Make a change of variable that transforms the quadratic form

$$Q(x) = x_1^2 - 8x_1x_2 - 5x_2^2$$

  into a quadratic form with no cross-product term.
- Compute $Q(x)$ for $x = (2, -2)$.
- Compute $Q(x)$ for $x = (2, -2)$ using new variable y.

We can now state the theorem on the change of variable in a quadratic form.

The Principal Axes Theorem

Let A be an $n \times n$ symmetric matrix. Then there is an orthogonal matrix P such that the mapping defined by

$$x = Py$$

transforms the quadratic form $x^T Ax$ into a quadratic form $y^T Dy$ with no cross-product term.

The columns of P in the theorem are called the **principal axes** of the quadratic form $x^T Ax$. The vector y is the coordinate vector of x relative to the orthonormal basis of $\mathbb{R}^n$ given by these principal axes.

Proof. Let A be an $n \times n$ symmetric matrix. Then by the spectral theorem there is an orthogonal matrix P of eigenvectors of A. such that

$$P^T AP = \Lambda,$$

where Λ is the diagonal matrix of eigenvalues of A. Since P is orthogonal,

$$P^{-1} = P^T.$$

Proof... Put

$$x = Py.$$

Then

$$\begin{aligned} x^T A x &= (Py)^T A (Py) \\ &= y^T P^T A P y \\ &= y^T (P^T A P) y \\ &= y^T \Lambda y. \end{aligned}$$

Since Λ is a diagonal matrix, the quadratic form $y^T \Lambda y$ does not contain cross-product terms. $\square$