Quadratic Forms

Definition of a Quadratic Form

Expressions of the form

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n$$

occurred in our study of linear equations and linear systems. If $a_1, a_2, \ldots, a_n$ are treated as constants, then this expression is a real-valued function of the ***variables*** $x_1, x_2, \ldots, x_n$ and is called a ***linear form*** on 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 R^n, which are functions of the form

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

The terms of the form $a_kx_ix_j$ in which i is $\neq j$ 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 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 R^3 as

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

If, as usual, we do not distinguish between the number a and the 1×1 matrix $[a]$, and if we let $\mathbf{x}$ be the column vector of variables, then (1) and (2) can be expressed in matrix form as

$$[x_1 \quad x_2]\begin{bmatrix} a_1 & a_3 \\ a_3 & a_2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \mathbf{x}^TA\mathbf{x}$$

$$[x_1 \quad x_2 \quad x_3]\begin{bmatrix} a_1 & a_4 & a_5 \\ a_4 & a_2 & a_6 \\ a_5 & a_6 & a_3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \mathbf{x}^TA\mathbf{x}$$

(verify). 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 $\mathbf{x}$ is an $n \times 1$ column vector of variables, then we call the function

$$Q_A(\mathbf{x}) = \mathbf{x}^TA\mathbf{x} \tag{3}$$

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

$$\boxed{\mathbf{x}^TA\mathbf{x} = \mathbf{x} \cdot A\mathbf{x} = A\mathbf{x} \cdot \mathbf{x}} \tag{4}$$

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

$$\mathbf{x}^TA\mathbf{x} = [x_1 \quad x_2 \quad \cdots \quad x_n]\begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \lambda_1x_1^2 + \lambda_2x_2^2 + \cdots + \lambda_nx_n^2$$

EXAMPLE 1 | Expressing Quadratic Forms in Matrix Notation

In each part, express the quadratic form in the matrix notation $\mathbf{x}^T A\mathbf{x}$, where A is symmetric.

(a) $2x^2 + 6xy - 5y^2$ (b) $x_1^2 + 7x_2^2 - 3x_3^2 + 4x_1x_2 - 2x_1x_3 + 8x_2x_3$

Solution The diagonal entries of A are the coefficients of the squared terms, and the off-diagonal entries are half the coefficients of the cross product terms, so

$$2x^2 + 6xy - 5y^2 = [x \quad y]\begin{bmatrix} 2 & 3 \\ 3 & -5 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$$

$$x_1^2 + 7x_2^2 - 3x_3^2 + 4x_1x_2 - 2x_1x_3 + 8x_2x_3 = [x_1 \quad x_2 \quad x_3]\begin{bmatrix} 1 & 2 & -1 \\ 2 & 7 & 4 \\ -1 & 4 & -3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

Change of Variable in a Quadratic Form

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

$$\mathbf{x} = P\mathbf{y} \tag{5}$$

that 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 (5) a ***change of variable***, and if P is orthogonal, then we call (5) an ***orthogonal change of variable***.

The following result, called the ***Principal Axes Theorem***, shows that by making an appropriate orthogonal change of variable in a quadratic form it is possible to eliminate its cross product terms, thereby producing a simpler quadratic form that is generally easier to work with.

Theorem 7.3.1

The Principal Axes Theorem

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

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

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.

EXAMPLE 2 | An Illustration of the Principal Axes Theorem

Find an orthogonal change of variable that eliminates the cross product terms in the quadratic form $Q = x_1^2 - x_3^2 - 4x_1x_2 + 4x_2x_3$, and express Q in terms of the new variables.

Solution The quadratic form can be expressed in matrix notation as

$$Q = \mathbf{x}^T A\mathbf{x} = [x_1 \quad x_2 \quad x_3]\begin{bmatrix} 1 & -2 & 0 \\ -2 & 0 & 2 \\ 0 & 2 & -1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

The characteristic equation of the matrix A is

$$\begin{vmatrix} \lambda - 1 & 2 & 0 \\ 2 & \lambda & -2 \\ 0 & -2 & \lambda + 1 \end{vmatrix} = \lambda^3 - 9\lambda = \lambda(\lambda+3)(\lambda-3) = 0$$

so the eigenvalues are $\lambda = 0, -3, 3$. We leave it for you to show that orthonormal bases for the three eigenspaces are

$$\lambda = 0: \begin{bmatrix} \frac{2}{3} \\ \frac{1}{3} \\ \frac{2}{3} \end{bmatrix}, \quad \lambda = -3: \begin{bmatrix} -\frac{1}{3} \\ -\frac{2}{3} \\ \frac{2}{3} \end{bmatrix}, \quad \lambda = 3: \begin{bmatrix} -\frac{2}{3} \\ \frac{2}{3} \\ \frac{1}{3} \end{bmatrix}$$

Thus, a substitution $\mathbf{x} = P\mathbf{y}$ that eliminates the cross product terms is

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} \frac{2}{3} & -\frac{1}{3} & -\frac{2}{3} \\ \frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{2}{3} & \frac{1}{3} \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$$

This produces the new quadratic form

$$Q = \mathbf{y}^T(P^T AP)\mathbf{y} = [y_1 \quad y_2 \quad y_3]\begin{bmatrix} 0 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 3 \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = -3y_2^2 + 3y_3^2$$

in which there are no cross product terms.

Question:

In Exercises 1–2, express the quadratic form in the matrix notation x TAx, where A is a symmetric matrix.

c. $9x_1^2 - x_2^2 + 4x_3^2 + 6x_1x_2 - 8x_1x_3 + x_2x_3$

Solution:

(c) $$9x_1^2 - x_2^2 + 4x_3^2 + 6x_1x_2 - 8x_1x_3 + x_2x_3 = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} 9 & 3 & -4 \\ 3 & -1 & \frac{1}{2} \\ -4 & \frac{1}{2} & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

Question:

In Exercises 3–4, find a formula for the quadratic form that does not use matrices.

4. $$\begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} -2 & \frac{7}{2} & 1 \\ \frac{7}{2} & 0 & 6 \\ 1 & 6 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

Solution:

$$\begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} -2 & \frac{7}{2} & 1 \\ \frac{7}{2} & 0 & 6 \\ 1 & 6 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = -2x_1^2 + 3x_3^2 + 7x_1x_2 + 2x_1x_3 + 12x_2x_3$$

Question:

In Exercises 5–8, find an orthogonal change of variables that eliminates the cross product terms in the quadratic form Q, and express Q in terms of the new variables.

8. $Q = 2x_1^2 + 5x_2^2 + 5x_3^2 + 4x_1x_2 - 4x_1x_3 - 8x_2x_3$

Solution:

$Q = \mathbf{x}^T A\mathbf{x} = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} 2 & 2 & -2 \\ 2 & 5 & -4 \\ -2 & -4 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$; the characteristic polynomial of the matrix A is

$\det(\lambda I - A) = \begin{vmatrix} \lambda-2 & -2 & 2 \\ -2 & \lambda-5 & 4 \\ 2 & 4 & \lambda-5 \end{vmatrix} = (\lambda-1)^2(\lambda-10)$ so the eigenvalues of A are 1 and 10.

The reduced row echelon form of $1I - A$ is $\begin{bmatrix} 1 & 2 & -2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ so that the eigenspace corresponding to

$\lambda = 1$ contains vectors $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ where $x_1 = -2s + 2t$, $x_2 = s$, $x_3 = t$. Vectors $\mathbf{p}_1 = \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}$ and $\mathbf{p}_2 = \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}$ form a basis for this eigenspace. We apply the Gram-Schmidt process to find an orthogonal basis for this eigenspace: $\mathbf{v}_1 = \mathbf{p}_1 = \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}$ and $\mathbf{v}_2 = \mathbf{p}_2 - \frac{\langle \mathbf{p}_2, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 = \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} - \frac{-4}{5}\begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{2}{5} \\ \frac{4}{5} \\ 1 \end{bmatrix}$, then proceed to normalize the two vectors to yield an orthonormal basis: $\mathbf{q}_1 = \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|} = \begin{bmatrix} -\frac{2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \\ 0 \end{bmatrix}$ and $\mathbf{q}_2 = \frac{\mathbf{v}_2}{\|\mathbf{v}_2\|} = \begin{bmatrix} \frac{2}{3\sqrt{5}} \\ \frac{4}{3\sqrt{5}} \\ \frac{5}{3\sqrt{5}} \end{bmatrix}$.

The reduced row echelon form of $10I - A$ is $\begin{bmatrix} 1 & 0 & \frac{1}{2} \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$ so that the eigenspace corresponding to

$\lambda = 10$ contains vectors $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ where $x_1 = -\frac{1}{2}t$, $x_2 = -t$, $x_3 = t$. A vector $\mathbf{p}_3 = \begin{bmatrix} -1 \\ -2 \\ 2 \end{bmatrix}$ forms a basis for this eigenspace.

Applying the Gram-Schmidt process to $\{\mathbf{p}_3\}$ amounts to simply normalizing this vector.

Therefore an orthogonal change of variables $\mathbf{x} = P\mathbf{y}$ that eliminates the cross product terms in Q is

$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -\frac{2}{\sqrt{5}} & \frac{2}{3\sqrt{5}} & -\frac{1}{3} \\ \frac{1}{\sqrt{5}} & \frac{4}{3\sqrt{5}} & -\frac{2}{3} \\ 0 & \frac{5}{3\sqrt{5}} & \frac{2}{3} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$. In terms of the new variables, we have

$$Q = \mathbf{x}^T A\mathbf{x} = \mathbf{y}^T \left(P^T AP\right)\mathbf{y} = \begin{bmatrix} y_1 & y_2 & y_3 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 10 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = y_1^2 + y_2^2 + 10y_3^2 .$$