Natalia Lazzati

Mathematics for Economics

Note 4: Quadratic Forms

Note 4 is based on Searle and Willett (2001) and Simon and Blume (1994, Ch. 16).

Quadratic Forms in Economics

Quadratic forms are important in testing the second order conditions that distinguish maxima from minima in economic optimization problems, in checking the concavity of functions that are twice continuously differentiable and in the theory of variance in statistics. Like linear functions, quadratic forms have matrix representations. So studying the properties of quadratic forms reduces to studying properties of a symmetric matrix.

Let us discuss two examples before formalizing the concept. The first example sheds light on the relevance of quadratic forms; the second one describes an economic application.

Example 1. The simplest functions with a unique global extremum are the pure **quadratics** $y = x^2$ and $y = -x^2$. The former has a global minimum at $x = 0$; and the latter has a global maximum at $x = 0$. ▲

Example 2. One simple production function in economics takes the form

$$y = \theta + \lambda_1 x_1 + \lambda_2 x_2 + a_{11} x_1^2 + a_{22} x_2^2 + a_{12} x_1 x_2$$

where y is an output and x_1 and x_2 are two inputs. This can be written as

$$\begin{aligned} y &= \theta + \begin{pmatrix} \lambda_1 & \lambda_2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} a_{11} & \frac{1}{2} a_{12} \\ \frac{1}{2} a_{12} & a_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\ &= \theta + \boldsymbol{\lambda}^T \mathbf{x} + \mathbf{x}^T A \mathbf{x}. \end{aligned}$$

Here $\mathbf{x}^T A \mathbf{x}$ is called a quadratic form. ▲

Definition of Quadratic Forms

The definition below formalizes de concept of quadratic form.

Definition 1. A quadratic form on $\mathbb{R}^n$ is a real-valued function $Q : \mathbb{R}^n \to \mathbb{R}$ of the form

$$Q(x_1, ..., x_n) = \textstyle\sum_{i,j=1}^{n} a_{ij} x_i x_j = \mathbf{x}^T A \mathbf{x} \tag{1}$$

in which each term is a monomial of degree two, where $\mathbf{x}^T = (x_1, ..., x_n)$ and $A = \{a_{ij}\}$ for $i = 1, 2, ..., n$ and $j = 1, 2, ..., n$.

By definition, the matrix A in (1) is always a square matrix. As we argue next, this matrix can also be taken as symmetric. If $\mathbf{x}^T B \mathbf{x}$ is a quadratic form, then the next equalities apply

$$\begin{aligned}
\mathbf{x}^T B \mathbf{x} &= \left(\mathbf{x}^T B \mathbf{x}\right)^T \\
&= \mathbf{x}^T B^T \mathbf{x} \\
&= \frac{1}{2}\left(\mathbf{x}^T B^T \mathbf{x} + \mathbf{x}^T B^T \mathbf{x}\right) \\
&= \frac{1}{2}\left(\mathbf{x}^T B \mathbf{x} + \mathbf{x}^T B^T \mathbf{x}\right) \\
&= \mathbf{x}^T \left[\frac{1}{2}\left(B + B^T\right)\right] \mathbf{x} \\
&= \mathbf{x}^T A \mathbf{x}.
\end{aligned}$$

Here $A = \frac{1}{2}\left(B + B^T\right)$ which is obviously symmetric. (**Justify each of the previous equalities.**) This means that every quadratic form $\mathbf{x}^T B \mathbf{x}$ can be written as $\mathbf{x}^T A \mathbf{x}$ with A symmetric.

Definiteness of Quadratic Forms

A quadratic form always takes on the value zero at the point $\mathbf{x} = \mathbf{0}$. Its distinguishing characteristic is the set of values it takes when $\mathbf{x} \neq \mathbf{0}$.

The general quadratic form of one variable is $y = ax^2$. If $a > 0$, then ax^2 is always ≥ 0 and equals 0 only when $x = 0$. Such a form is called positive definite, and $x = 0$ is its global minimizer. If $a < 0$, then ax^2 is always ≤ 0 and equals 0 only when $x = 0$. Such a form is called negative definite, and $x = 0$ is its global maximizer.

In two dimensions, the quadratic form $Q_1(x_1, x_2) = x_1^2 + x_2^2$ is always greater than zero at $(x_1, x_2) \neq (0, 0)$. Then, we call Q_1 positive definite. Quadratic forms as $Q_2(x_1, x_2) = -x_1^2 - x_2^2$, which are strictly negative except at the origin, are called negative definite. Quadratic forms like $Q_3(x_1, x_2) = x_1^2 - x_2^2$, which take on both positive and negative values are called indefinite.

There are two intermediate cases. Quadratic forms which are always greater than or equal to zero, but may equal zero at some nonzero $\mathbf{x}'s$ are called positive semi-definite. For instance,

$$Q_4(x_1, x_2) = (x_1 + x_2)^2 = x_1^2 + 2x_1x_2 + x_2^2$$

which is never negative, but which equals zero at nonzero points such as $(x_1, x_2) = (1, -1)$. A quadratic form like $Q_5(x_1, x_2) = -(x_1 + x_2)^2$, which is never positive but can be zero at points other than the origin, is called negative semi-definite.

As we showed, quadratic forms are always associated to symmetric matrices. It is then non surprising that a symmetric matrix A is called positive definite, positive semi-definite, negative definite, negative semi-definite or indefinite according to the definiteness of the corresponding quadratic form $Q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$.

The next definition extends these concepts to higher dimensions.

Definition 2. (Definiteness) Let A be an $n \times n$ symmetric matrix and $\mathbf{x}$ an n-dimensional vector, then A is

(a) positive definite if $\mathbf{x}^T A\mathbf{x} > 0$ for all $\mathbf{x} \neq \mathbf{0}$;

(b) positive semi-definite if $\mathbf{x}^T A\mathbf{x} \geq 0$ for all $\mathbf{x}$;

(c) negative definite if $\mathbf{x}^T A\mathbf{x} < 0$ for all $\mathbf{x} \neq \mathbf{0}$; and

(d) negative semi-definite if $\mathbf{x}^T A\mathbf{x} \leq 0$ for all $\mathbf{x}$.

Example 3. Let $A = \begin{pmatrix} 4 & 2 \\ 2 & 1 \end{pmatrix}$. Then,

$$\mathbf{x}^T A\mathbf{x} = \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} 4 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 4x_1^2 + 4x_1x_2 + x_2^2.$$

Since $4x_1^2 + 4x_1x_2 + x_2^2 = (2x_1 + x_2)^2$, it follows that $\mathbf{x}^T A\mathbf{x} \geq 0$ for all $\mathbf{x}$. In addition, $\mathbf{x}^T A\mathbf{x} = 0$ when $\mathbf{x} = (-1, 2) \neq (0, 0)$. Therefore A is positive semi-definite. ▲

In the last example it was quite easy to address the definiteness of matrix A. In higher dimensions, checking this property is often quite involved. Fortunately, there are techniques based on specific determinants that help us to do so.

Characterization through Determinants

We next characterize the definiteness of a matrix through its principal minors.

Definition 3. **(Principals and principal minors)** Let A be an $n \times n$ matrix. A $k \times k$ submatrix of A formed by deleting $n-k$ columns and **the same** $n-k$ rows from A is called a kth order principal submatrix of A. The determinant of a $k \times k$ principal submatrix is called a kth order principal minor of A.

(Leading principals and leading principal minors) Let A be an $n \times n$ matrix. The kth order principal submatrix of A, with $1 \leq k \leq n$, obtained by deleting the last $n-k$ rows and the last $n-k$ columns from A is called the kth leading principal submatrix of A. Its determinant is called the kth leading principal minor of A. We denote the former by A_k, and the latter by $|A_k|$.

The next theorem uses Definition 3 to characterize the definiteness of a matrix.

Theorem 4. Let A be an $n \times n$ matrix. Then,

(a) A is positive definite if and only if its n leading principal minors are strictly positive;

(b) A is negative definite if and only if, for $k = 1, ..., n$, the kth leading principal minors have the same sign as $(-1)^k$, that is,

$$|A_1| < 0, |A_2| > 0, |A_3| < 0, \text{ etc.};$$

(c) A is positive semi-definite if and only if all its principal minors are ≥ 0; and

(d) A is negative semi-definite if and only if every principal minor of odd order is ≤ 0 and every principal minor of even order is ≥ 0.

Remark. The condition in part (b) can be written as $(-1)^k |A_k| > 0$ for $k = 1, ..., n$.

Proof. See SB pp. 394-395. ■

Example 4. Consider the following matrix

$$A = \begin{pmatrix} -3 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Its leading principal submatrices are

$$A_1 = -3,\ A_2 = \begin{pmatrix} -3 & 0 \\ 0 & 0 \end{pmatrix} \text{ and } A_3 = \begin{pmatrix} -3 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Therefore, its leading principal minors are $|A_1| = -3$, $|A_2| = 0$, and $|A_3| = 0$. It follows that A is neither negative nor positive definite. (**Check whether the matrix is negative semi-definite.**) ▲

(**Offer a precise characterization of positive and negative definite diagonal matrices. Repeat it for** 2×2 **matrices.**)

Quadratic Forms with Linear Constraints

Determining the definiteness of a quadratic form $Q(\mathbf{x})$ is equivalent to determining whether $\mathbf{x} = \mathbf{0}$ is a maximum, a minimum or neither for the real-valued function Q. Specifically, $\mathbf{x} = \mathbf{0}$ is the unique global minimum of a quadratic form Q if and only if Q is positive definite. Similarly, $\mathbf{x} = \mathbf{0}$ is the unique global maximum if and only if Q is negative definite.

The characterization of quadratic forms in Theorem 4 works only when there are no constraints in the problem under consideration, that is, if $\mathbf{x}$ can take on any value on $\mathbb{R}^n$. If there are constraints the analysis becomes more delicate.

Example 5. Let us consider again the quadratic form $Q_3(x_1, x_2) = x_1^2 - x_2^2$, which takes on both positive and negative values and is thereby indefinite; the origin is neither a maximum nor a minimum. But if we restrict attention to the x_1-axis, that is, if we impose $x_2 = 0$, then $Q_3(x_1, 0) = x_1^2$ has a strict global minimum at $x_1 = 0$. On the constraint set $\left\{(x_1, x_2) \in \mathbb{R}^2 : x_2 = 0\right\}$ the quadratic form Q_3 is positive definite. Alternatively, if we restrict attention to the x_2-axis, that is, if we impose $x_1 = 0$, then $Q_3(0, x_2) = -x_2^2$ has a strict global maximum at $x_2 = 0$. On the constraint set $\{(x_1, x_2) \in \mathbb{R}^2 : x_1 = 0\}$ the quadratic form Q_3 is negative definite. ▲

Let us extend Example 5. Suppose we want to determine the definiteness of a general quadratic form of two variables

$$Q(x_1, x_2) = ax_1^2 + 2bx_1x_2 + cx_2^2 = (x_1\ x_2)\begin{pmatrix} a & b \\ b & c \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \tag{2}$$

on the general linear subspace

$$Ax_1 + Bx_2 = 0. \tag{3}$$

The simplest approach is to solve (2) for x_1 in terms of x_2. By (3) we know that $x_1 = -(B/A)x_2$. Substituting the latter in (2) and simplifying we get

$$Q\left(-\frac{B}{A}x_2, x_2\right) = \frac{aB^2 - 2bAB + cA^2}{A^2}x_2^2. \tag{4}$$

We conclude that Q is positive definite on the constraint if and only if $aB^2 - 2bAB + cA^2 > 0$, and negative definite on the constraint if and only if $aB^2 - 2bAB + cA^2 < 0$. There is a convenient way of rewriting this expression

$$aB^2 - 2bAB + cA^2 = -\left|\begin{pmatrix} 0 & A & B \\ A & a & b \\ B & b & c \end{pmatrix}\right|. \tag{5}$$

The matrix in (5) is obtained by "bordering" the 2×2 matrix in (2) on the top and left by the coefficients of the linear constraint (3). The next theorem captures these observations.

Theorem 5. The quadratic form $Q(x_1, x_2) = ax_1^2 + 2bx_1x_2 + cx_2^2$ is positive (negative) definite on the constraint set $Ax_1 + Bx_2 = 0$ if and only if

$$\left|\begin{pmatrix} 0 & A & B \\ A & a & b \\ B & b & c \end{pmatrix}\right| \tag{6}$$

is strictly negative (positive).

A similar result holds for the general problem of determining the definiteness of

$$Q(\mathbf{x}) = \mathbf{x}^T A \mathbf{x} = \begin{pmatrix} x_1 & \dots & x_n \end{pmatrix} \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \tag{7}$$

on the linear constraint

$$B\mathbf{x} = \begin{pmatrix} B_{11} & \dots & B_{1n} \\ \vdots & \ddots & \vdots \\ B_{m1} & \dots & B_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{8}$$

We border the matrix of the quadratic form (7) on the top and on the left by the coefficient matrix of the linear constraint (8)

$$H = \left(\begin{array}{ccc|ccc} 0 & \dots & 0 & B_{11} & \dots & B_{1n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & B_{m1} & \dots & B_{mn} \\ \hline B_{11} & \dots & B_{m1} & a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ B_{1n} & \dots & B_{mn} & a_{n1} & \dots & a_{nn} \end{array}\right)$$

and address the definiteness of the quadratic form on the linear constraint by studying specific features of H, that we describe in the next theorem.

Theorem 6. To determine the definiteness of a quadratic form of n variables $Q(\mathbf{x}) = \mathbf{x}^T A \mathbf{x}$, when restricted to the constraint set given by m linear equations $B\mathbf{x} = \mathbf{0}$, you should proceed as follows. First, construct the $(n+m) \times (n+m)$ symmetric matrix H by bordering the matrix A above and to the left by the coefficient matrix B of the linear constraints

$$H = \begin{pmatrix} \mathbf{0} & B \\ B^T & A \end{pmatrix}.$$

Then, check the signs of the last $n-m$ leading principal minors of H, starting with the determinant of H itself. There are three possibilities

(a) if $|H|$ has the same sign as $(-1)^n$ and if these last $n-m$ leading principal minors alternate in sign, then Q is negative definite on the constraint set $B\mathbf{x} = \mathbf{0}$, and $\mathbf{x} = \mathbf{0}$ is a strict global maximum of Q on this constraint set;

(b) if $|H|$ and the last $n-m$ leading principal minors all have the same sign as $(-1)^m$, then Q is positive definite on the constraint set $B\mathbf{x} = \mathbf{0}$, and $\mathbf{x} = \mathbf{0}$ is a strict global minimum of Q on this constraint set; and

(c) if (a) and (b) are violated by nonzero leading principal minors, then Q is indefinite on the constraint set $B\mathbf{x} = \mathbf{0}$, and $\mathbf{x} = \mathbf{0}$ is neither a maximum nor a minimum of Q on this constraint set.