Supplementary Lecture 2. Quadratic forms

S2.1. Linear and quadratic forms

A linear equation in the n variables $x_1, x_2, ..., x_n$ is one that can be expressed in the form

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$$

The left hand side of this equation is a function of n variables called a linear form. In a linear form, all variables occur to the first power (ie. there are no terms x_i^n for $n > 1$), and there are no products of variables in the expression.

In functions called quadratic forms, all the terms are either squares of variables, or products of two variables. A quadratic form in two variables, x and y, is defined to be an expression that can be written as $ax^2 + 2bxy + cy^2$. This expression can be written in matrix terms as follows:

$$ax^2 + 2bxy + cy^2 = \begin{bmatrix} x & y \end{bmatrix}\begin{bmatrix} a & b \\ b & c \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$$

Note that the (2×2) matrix is symmetric; the diagonal entries are the coefficients of the squared terms, and the entries off the main diagonal are each equal to half the coefficient of the product term xy.

Quadratic forms are not limited to two variables. A general quadratic form in $x_1, x_2, ..., x_n$ is an expression that can be written in matrix terms as

$$\begin{bmatrix} x_1 & \dots & x_n \end{bmatrix} A \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

where A is a symmetric (n×n) matrix. It can also be written more compactly as x'Ax. If these matrices are multiplied out, the resulting expression has the form

$$x'Ax = a_{11}x_1^2 + \cdots + a_{nn}x_n^2 + \sum_{i \neq j} a_{ij}x_ix_j$$

where $\sum_{i \neq j} a_{ij}x_ix_j$ denotes a sum of terms of the form $a_{ij}x_ix_j$, where x_i and x_j are different variables. They are called the cross-product terms of the quadratic form.

Example: The following expression is a quadratic form in x_1, x_2 and x_3:

$$x_1^2 + 7x_2^2 - 3x_3^2 + 4x_1x_2 - 2x_1x_3 + 6x_2x_3 = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix}\begin{bmatrix} 1 & 2 & -1 \\ 2 & 7 & 3 \\ -1 & 3 & -3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

Notice that the coefficients of the squared terms appear on the main diagonal of the (3×3) matrix, and the coefficients of the cross-product terms are each split in half, and appear in the off-diagonal positions as follows:

Coefficient of:	Positions in matrix A:
x_1x_2	a_{12} and a_{21}
x_1x_3	a_{13} and a_{31}
x_2x_3	a_{23} and a_{32}

So in terms of a general (3×3) matrix A we have

$$x'Ax = a_{11}x_1^2 + a_{22}x_2^2 + a_{33}x_3^2 + (a_{12} + a_{21})x_1x_2 + (a_{13} + a_{31})x_1x_3 + (a_{23} + a_{32})x_2x_3$$

S2.2. Definite matrices

Consider the following three symmetric matrices:

$$A = \begin{bmatrix} 4 & 1 \\ 1 & 2 \end{bmatrix} \quad B = \begin{bmatrix} -3 & 1 \\ 1 & -2 \end{bmatrix} \quad C = \begin{bmatrix} 3 & 0 \\ 0 & -3 \end{bmatrix}$$

Given a two-dimensional vector $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ in E^2, matrix A yields the following quadratic form:

$$x'Ax = 4x_1^2 + 2x_2^2 + 2x_1x_2$$

Now, it just so happens that the value of this quadratic form is <u>always</u> strictly positive, irrespective of the values of x_1 and x_2, provided that x_1 and x_2 are not both zero ie. provided that x is not the null vector. Intuitively, this is because the weights on the x_i^2 are large and positive, and hence 'outweigh' the cross-product term $2x_1x_2$, which can be negative (of course, $x'Ax = 0$ when x is the null vector in E^2). If we think of the quadratic form x'Ax as a <u>function</u> from E^2 to E^1 (ie. $f(x_1, x_2) = x'Ax = 4x_1^2 + 2x_2^2 + 2x_1x_2$), then the graph of the function is a 'bowl' or 'valley-shaped' parabola in three dimensions. The value of f (ie. the value of the quadratic form $x'Ax = 4x_1^2 + 2x_2^2 + 2x_1x_2$) is strictly positive, no matter what the values of x_1 and x_2 are (as long as they are not both zero).

Now consider the matrix B above. Given a two-dimensional vector $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ in E^2, matrix B yields the following quadratic form:

$$x'Bx = -3x_1^2 - 2x_2^2 + 2x_1x_2$$

Here, the opposite is true: the terms on the x_i^2 are large and negative, so the value of the quadratic form $f(x_1, x_2) = x'Bx = -3x_1^2 - 2x_2^2 + 2x_1x_2$ is strictly negative for any $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \neq \begin{bmatrix} 0 \\ 0 \end{bmatrix}$. The quadratic form in this case has a graph which is an 'upside-down' bowl, or hill shape in three dimensions. Finally, consider the matrix C above. Given a two-dimensional vector $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ in E^2, matrix C yields the following quadratic form:

$$x'Cx = 3x_1^2 - 3x_2^2$$

In this case, the value of the function $f(x_1, x_2) = x'Cx = 3x_1^2 - 3x_2^2$ can be <u>both</u> positive and negative on E^2. It is positive when $x_1^2 > x_2^2$, and negative when $x_1^2 < x_2^2$. The three-dimensional graph of $f(x_1, x_2)$ in this case has a 'saddle' shape which I can never draw properly, so I'm not going to try!

I chose the three examples above very carefully, because it turns out that all quadratic forms defined on E^n for $n \geq 1$ must have one of these three basic ‘shapes’. We use this fact to characterise symmetric matrices as follows:

(i). An (n×n) symmetric matrix A is positive definite if $x'Ax > 0$ for all $x \neq \underset{(n\times 1)}{\mathbf{0}}$.

(ii). An (n×n) symmetric matrix A is positive semidefinite if $x'Ax \geq 0$ for all x.

(iii). An (n×n) symmetric matrix A is negative definite if $x'Ax < 0$ for all $x \neq \underset{(n\times 1)}{\mathbf{0}}$.

(iv). An (n×n) symmetric matrix A is negative semidefinite if $x'Ax \leq 0$ for all x.

If a symmetric matrix is not definite, then it is said to be indefinite. In the three examples I gave you above, matrix A was positive definite, matrix B was negative definite, and matrix C was indefinite.

As we shall see later in the course, it is extremely important in the application of multivariate calculus to optimisation problems to have a criterion for deciding whether a symmetric matrix is positive or negative (semi)definite, or indefinite. There are two basic ‘tests’ for establishing this: the eigenvalue test (which we do not have time to study in this course), and the principal minors test, to which we now turn.

S2.3. The principal minors test for sign definiteness

The kth principal minor of an (n×n) symmetric matrix A is the determinant of the submatrix formed from the first k rows and columns of A. For example, if

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

is a (3×3) symmetric matrix, then the first principal minor is just

$$a_{11}$$

the second principal minor is

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

and the third principal minor is

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = [a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{32}a_{21}] - [a_{13}a_{22}a_{31} + a_{12}a_{21}a_{33} + a_{11}a_{32}a_{23}]$$

Obviously, the nth principal minor of an (n×n) matrix A is just the determinant of A, |A|.

We have the following:

(a). An (n×n) symmetric matrix A is <u>positive definite</u> if and only if

$$a_{11} > 0$$

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0$$

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} > 0$$

$$\vdots$$

$$|A| > 0$$

ie. if and only if <u>all</u> the principal minors of A are <u>strictly positive</u>.

(b). An (n×n) symmetric matrix A is <u>positive semidefinite</u> if and only if

$$a_{11} \geq 0$$

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \geq 0$$

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \geq 0$$

$$\vdots$$

$$|A| \geq 0$$

ie. if and only if all the principal minors of A are <u>nonnegative</u> (some zero determinants are allowed).

(c). An (n×n) symmetric matrix A is <u>negative definite</u> if and only if the principal minors alternate in sign, starting negative ie. if and only if

$$a_{11} < 0$$

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0$$

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} < 0$$

$$\vdots$$

$$\text{sign}(|A|) = \text{sign}((-1)^n)$$

Notice that sign(|A|) must be positive if n is an <u>even</u> number, and negative if n is an <u>odd</u> number.

(d). An (n×n) symmetric matrix A is <u>negative semidefinite</u> if and only if

$$\mathbf{a_{11} \le 0}$$

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \ge 0$$

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \le 0$$

$$\vdots$$

$$\mathbf{sign(|A|) = sign((-1)^n) \text{ or zero}}$$

Again, some zero determinants are allowed for negative semidefiniteness.

<u>Example</u>: The (3×3) matrix

$$\begin{bmatrix} -3 & 0 & 0 \\ 0 & -2 & 1 \\ 0 & 1 & -7 \end{bmatrix}$$

is <u>negative definite</u>, because

$$-3 < 0$$

$$\begin{vmatrix} -3 & 0 \\ 0 & -2 \end{vmatrix} = 6 > 0$$

$$\begin{vmatrix} -3 & 0 & 0 \\ 0 & -2 & 1 \\ 0 & 1 & -7 \end{vmatrix} = -39 < 0$$

S2.4. <u>Another worked example</u>

Here is a worked example of the sort of problem <u>you will definitely be asked to solve</u> in a test:

Test the following matrix for definiteness using the principal minors test:

$$\begin{bmatrix} 1 & -1 & 0 \\ -1 & 3 & 1 \\ 0 & 1 & 1 \end{bmatrix}$$

<u>Solution</u>: The matrix is positive definite, since

$$1 > 0$$

$$\begin{vmatrix} 1 & -1 \\ -1 & 3 \end{vmatrix} = 2 > 0$$

$$\begin{vmatrix} 1 & -1 & 0 \\ -1 & 3 & 1 \\ 0 & 1 & 1 \end{vmatrix} = 1 > 0$$

(End of Supplementary Lecture 2)