Lecture 4
Quadratic Forms and Convexity

Eivind Eriksen

BI Norwegian School of Management
Department of Economics

September 17, 2010

Quadratic forms

Definition

A quadratic form *in the variables* $x_1, x_2, \ldots, x_n$ *is a polynomial function* Q *where all terms in the functional expression* $Q(x_1, x_2, \ldots, x_n)$ *have order two.*

Example

- $Q(x_1, x_2, x_3) = x_1^2 + x_2^2 + 3x_2x_3 - x_3^2$ *is a quadratic form in three variables*
- $f(x_1, x_2) = x_1^2 - x_2^2 + 3x_1 + 2x_2$ *is polynomial of degree two but not a quadratic form. It can be written as a sum* $f = L + Q$ *of a linear form* $L(x_1, x_2) = 3x_1 + 2x_2$ *and a quadratic form* $Q(x_1, x_2) = x_1^2 - x_2^2$.

The symmetric matrix of a quadratic form

Lemma

A function $f(x_1, x_2, \ldots, x_n)$ in n variables is a quadratic form if and only if it can be written as

$$f(x_1, x_2, \ldots, x_n) = \mathbf{x}^T A\,\mathbf{x}, \qquad \text{where } \mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

for a symmetric $n \times n$ matrix A.

The matrix A is uniquely determined by the quadratic form, and is called the symmetric matrix associated with the quadratic form.

Example

Compute the product $x^T A\,\mathbf{x}$ when A is the symmetric 2×2 matrix $A = \left(\begin{smallmatrix} 1 & 2 \\ 2 & -1 \end{smallmatrix}\right)$.

Quadratic forms: An example

Solution

We compute the matrix product

$$\mathbf{x}^T A\,\mathbf{x} = \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 + 2x_2 & 2x_1 - x_2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$
$$= x_1^2 + 2x_2x_1 + 2x_1x_2 - x_2^2 = x_1^2 + 4x_1x_2 - x_2^2$$

Note that the result of the matrix multiplication $\mathbf{x}^T A\,\mathbf{x}$ is, strictly speaking, the 1×1 matrix with entry $x_1^2 + 4x_1x_2 - x_2$.

Example

Find the symmetric matrix associated with the quadratic form $Q(x_1, x_2, x_3) = x_1^2 + 2x_1x_2 + x_2^2 + 4x_2x_3 - x_3^2$.

Quadratic forms: Another example

Solution

We note that the diagonal entry a_{ii} in the matrix A is the coefficient in front of the term x_i^2, and that the sum $a_{ij} + a_{ji}$ of corresponding entries off the diagonal in A is the coefficient in front of the term $x_i x_j$. Since A is symmetric, $a_{ij} = a_{ji}$ is half the coefficient in front of $x_i x_j$. Therefore, the quadratic form $Q(x_1, x_2, x_3) = x_1^2 + 2x_1x_2 + x_2^2 + 4x_2x_3 - x_3^2$ has symmetric matrix

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 2 \\ 0 & 2 & -1 \end{pmatrix}$$

Definiteness

For any quadratic form, we clearly have $Q(\mathbf{0}) = 0$. We want to find out if $\mathbf{x} = \mathbf{0}$ is a max or min point for Q.

Definition

A quadratic form $Q(\mathbf{x}) = \mathbf{x}^T A\, \mathbf{x}$ and its symmetric matrix A is

- *positive definite if $Q(\mathbf{x}) > 0$ when $\mathbf{x} \neq \mathbf{0}$*
- *positive semidefinite if $Q(\mathbf{x}) \geq 0$ when $\mathbf{x} \neq \mathbf{0}$*
- *negative definite if $Q(\mathbf{x}) < 0$ when $\mathbf{x} \neq \mathbf{0}$*
- *negative semidefinite if $Q(\mathbf{x}) \leq 0$ when $\mathbf{x} \neq \mathbf{0}$*
- *indefinite if $Q(\mathbf{x})$ takes both positive and negative values*

Question: Given a quadratic form, how do we determine if it is positive or negative (semi)definite?

Definiteness and eigenvalues

Proposition

Let $Q(\mathbf{x}) = \mathbf{x}^T A\, \mathbf{x}$ be a quadratic form, let A be its symmetric matrix, and let $\lambda_1, \lambda_2, \dots, \lambda_n$ be the eigenvalues of A. Then:

- *Q is positive definite $\Leftrightarrow \lambda_1, \lambda_2, \dots, \lambda_n > 0$*
- *Q is positive semidefinite $\Leftrightarrow \lambda_1, \lambda_2, \dots, \lambda_n \geq 0$*
- *Q is negativ definite $\Leftrightarrow \lambda_1, \lambda_2, \dots, \lambda_n < 0$*
- *Q is negative semidefinite $\Leftrightarrow \lambda_1, \lambda_2, \dots, \lambda_n \leq 0$*
- *A is indefinite $\Leftrightarrow$ there exists $\lambda_i > 0$ and $\lambda_j < 0$*

Note that a symmetric matrix is diagonalizable; hence it has n real eigenvalues (but there may be repetitions).

Definiteness and eigenvalues

Idea of proof: If A is diagonal, then $Q(\mathbf{x}) = \lambda_1 x_1^2 + \lambda_2 x_2^2 + \cdots + \lambda_n x_n^2$ and it is not difficult to see that the result is correct. In general, A may be non-diagonal, but it is always diagonalizable, and this can be used to show that the result is correct in general.

Example

Determine the definiteness of $Q(\mathbf{x}) = -x_1^2 + 6x_1x_2 - 9x_2^2 - 2x_3^2$.

Solution

The symmetric matrix associated with the quadratic form Q is given by

$$A = \begin{pmatrix} -1 & 3 & 0 \\ 3 & -9 & 0 \\ 0 & 0 & -2 \end{pmatrix}$$

Definiteness: An example

Solution (Continued)

We compute the eigenvalues of A using the characteristic equation:

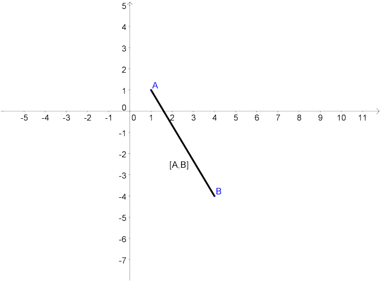
$$\det(A-\lambda I)=\begin{vmatrix}-1-\lambda & 3 & 0\\ 3 & -9-\lambda & 0\\ 0 & 0 & -2-\lambda\end{vmatrix}=(-2-\lambda)(\lambda^2+10\lambda)=0$$

We see that the eigenvalues are $\lambda=-2, \lambda=0, \lambda=-10$. Hence Q is negative semidefinite.

This example shows that $\mathbf{x}=\mathbf{0}$ is a maximum point for the quadratic form $Q(\mathbf{x})=-x_1^2+6x_1x_2-9x_2^2-2x_3^2$, since $Q(\mathbf{x})\leq 0$ for all $\mathbf{x}\neq\mathbf{0}$ and $Q(\mathbf{0})=0$.

Convex subsets

Let $\mathbb{R}^n$ be the n-dimensional space, and let $A = (a_1, a_2, \ldots, a_n)$ and $B = (b_1, b_2, \ldots, b_n)$ be two point in $\mathbb{R}^n$ described by their coordinates. We denote the straight line between A and B by $[A, B]$.

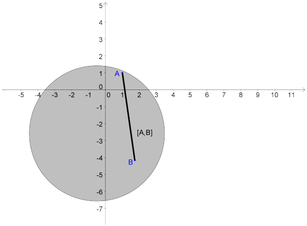


We call $[A, B]$ the line segment from A to B.

Definition: Convex subsets

Definition

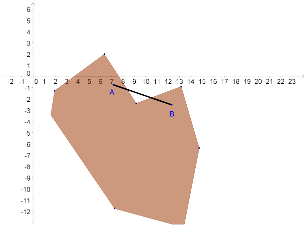
A subset $S \subseteq \mathbb{R}^n$ is convex if the following condition holds: Whenever $A, B \in S$ are point in S, the line segment $[A, B]$ is entirely in S.



Example

Let $S \subseteq \mathbb{R}^2$ be the marked circular disk in the plane. This is a convex subset of the plane, since the line segment between any two points in the circular disk is entirely within the disk.

Convex subsets: An example



Example

Let $S \subseteq \mathbb{R}^2$ be the marked area in the plane. This is not a convex subset of the plane, since there is a line segment between two points in S that is not entirely within S.

Concrete description of line segments

Let $A = (a_1, a_2, \ldots, a_n)$ and $B = (b_1, b_2, \ldots, b_n)$ be points in $\mathbb{R}^n$. We may consider the coordinates as vectors

$$\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

Parametrization of line segments

The line segment $[A, B]$ can be parametrized as

$$[A, B] : \qquad (1 - t)\mathbf{a} + t\mathbf{b} \qquad \text{for } t \in [0, 1]$$

Convex and concave functions in one variable

We shall generalize the notion of convex and concave functions from one variable to several variables. We recall the situation in the case of one variable:

Functions of one variable

A function f in one variable defined on an interval $I \subseteq \mathbb{R}$ is convex if $f''(x) \geq 0$ for all $x \in I$, and concave if $f''(x) \leq 0$ for all $x \in I$. The graph of convex and concave function have the following shapes:

$$\text{Convex: } \bigcup \qquad \text{Concave: } \bigcap$$

If f is a quadratic form in one variable, it can be written as $f(x) = ax^2$. In this case, f is convex if $a \geq 0$ and concave if $a \leq 0$.

Convex and concave functions in several variables

When $f(x_1, x_2, \ldots, x_n)$ is a function in n variables, its graph is given by the equation $x_{n+1} = f(x_1, x_2, \ldots, x_n)$ and it can be drawn in a coordinate system of dimension $n+1$.

Definition

Let f be a function in n variables defined on a convex subset $S \subseteq \mathbb{R}^n$. Then we define that

- *f is convex if the line segment joining any two points of the graph of f is never under the graph.*
- *f is concave if the line segment joining any two points of the graph of f is never over the graph.*

Example: A convex function in two variables

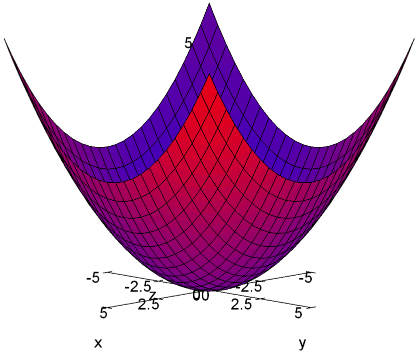


Figure: The graph of the function $f(x, y) = x^2 + y^2$

Concrete conditions for convexity

Let $A = (a_1, \dots, a_n)$ and $B = (b_1, \dots, b_n)$ be any two points in S, and let $\mathbf{a}, \mathbf{b}$ be the corresponding column vectors. Then we have the following concrete conditions of convexity:

Concrete conditions

- f is convex $\Leftrightarrow f((1-t)\mathbf{a} + t\mathbf{b}) \leq (1-t)f(\mathbf{a}) + tf(\mathbf{b})$ for all $\mathbf{a}, \mathbf{b} \in S$ and all $t \in [0,1]$
- f is concave $\Leftrightarrow f((1-t)\mathbf{a} + t\mathbf{b}) \geq (1-t)f(\mathbf{a}) + tf(\mathbf{b})$ for all $\mathbf{a}, \mathbf{b} \in S$ and all $t \in [0,1]$

Idea of proof: We know that $\mathbf{v} = (1-t)\mathbf{a} + t\mathbf{b}$ is a point on the line segment $[A, B]$. So the first expression in each inequality is $f(\mathbf{v})$, the x_{n+1} coordinate of a point on the graph of f, while the second expression is the x_{n+1} coordinate of the line segment between $(A, f(A))$ and $(B, f(B))$.

Strictly convex and concave functions

Definition

Let f be a function in n variables defined on a convex subset $S \subseteq \mathbb{R}^n$. Then we define that

- *f is strictly convex if the line segment joining any two points of the graph of f is entirely over the graph between A and B, i.e.*

$$f((1-t)\mathbf{a}+t\mathbf{b}) < (1-t)f(\mathbf{a})+tf(\mathbf{b}) \text{ for } t \in (0,1)$$

- *f is strictly concave if the line segment joining any two points of the graph of f is entirely under the graph between A and B, i.e.*

$$f((1-t)\mathbf{a}+t\mathbf{b}) > (1-t)f(\mathbf{a})+tf(\mathbf{b}) \text{ for } t \in (0,1)$$

Concrete conditions: An example

Example

Let $f(x_1, x_2) = x_1 + x_2$. Is f convex or concave?

Solution

In this case, f is defined on $S = \mathbb{R}^2$. For any points $A = (a_1, a_2)$ and $B = (b_1, b_2)$ in $\mathbb{R}^2$, we have that

$$f((1-t)\mathbf{a} + t\mathbf{b}) = f\begin{pmatrix}(1-t)a_1 + tb_1 \\ (1-t)a_2 + tb_2\end{pmatrix} = (1-t)(a_1 + a_2) + t(b_1 + b_2)$$

and

$$(1-t)f(\mathbf{a}) + tf(\mathbf{b}) = (1-t)(a_1 + a_2) + t(b_1 + b_2)$$

We see that these expressions coincide, so f is both convex and concave.

Properties of convex functions

Theorem

Let f be a function in n variables defined on a convex subset $S \subseteq \mathbb{R}^n$. Then we have

1. *If f is constant or a linear function (polynomial function of degree one), then f is both convex and concave.*
2. *f is convex $\Leftrightarrow$ $-f$ is concave*
3. *If $f = a_1 f_1 + a_2 f_2$, where $a_1, a_2 \geq 0$ and f_1, f_2 are convex functions defined on S, then f is convex.*
4. *If $f = a_1 f_1 + a_2 f_2$, where $a_1, a_2 \geq 0$ and f_1, f_2 are concave functions defined on S, then f is concave.*

Convex functions: Another example

Example

Show that $f(x, y) = 2x^2 + 3y^2$ and $g(x, y) = 2x^2 + 3y^2 + x - y + 3$ are convex functions.

Solution

We know that x^2 is a convex function in one variable. So the functions x^2 and y^2 are also convex functions in two variables. Hence f is a convex function by the second part of the theorem. We also see that $x - y + 3$ is a convex functions; it follows from part one of the theorem since it is linear. Hence $g(x, y) = f(x, y) + x - y + 3$ is also a convex function.

What can we say about the convexity of more general functions? The first difficult case is the convexity of quadratic forms.

Convexity of quadratic forms

Proposition

Let $Q(x_1, x_2, \dots, x_n) = \mathbf{x}^T A\, \mathbf{x}$ be a quadratic form in n variables, with associated symmetric matrix A. Then we have:

- *Q is convex $\Leftrightarrow$ A is positive semidefinite*
- *Q is concave $\Leftrightarrow$ A is negative semidefinite*
- *Q is strictly convex $\Leftrightarrow$ A is positive definite*
- *Q is strictly concave $\Leftrightarrow$ A is negative definite*

Idea of proof: If A is diagonal with eigenvalues $\lambda_1, \dots, \lambda_n$, then $Q(x_1, x_2, \dots, x_n) = \lambda_1 x_1^2 + \lambda_2 x_2^2 + \cdots + \lambda_n x_n^2$. In this situation, one may use the sign of the eigenvalues to decide the convexity of Q using the theorem. If A is not diagonal, it can be diagonalized, and the same type of argument can be used also in this case.