Quadratic Forms

We consider the unconstrained optimization for the case of functions with many variables:

$$max_x \ f\ (x) \ subject \ to \ x \ \in S$$

where x is a vector

To face this topic we need some preliminary notions:

- Quadratic forms
- Concavity and convexity of functions of many variables

Definition of quadratic forms

A form is a polynomial function in which each component has the same sum of the exponents:

- a linear form is f(x,y,z) = 4x – 9y + z

(each term has exponents that add to one (the “first degree”)

- a quadratic form is f(x y z) = $4x^2 + 2zy - xz + 2z^2$

(each term has exponents that add to two (the “second degree”)

A polynomial equation in which ***each term*** is of the 2nd degree (sum of the integer exponents = 2) is a quadratic form

Definition

A **quadratic form** in *n* variables is a function

$$Q(x_1, \ldots . x_n) = b_{11}x_1^2 + b_{12}x_1\ x_2\ + \cdots + b_{ij}\ x_i\ \ x_j\ \ + \cdots + b_{nn}x_n^2 =$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} b_{ij}x_i\ \ x_j$$

where b_{ij} for $i = 1, \ldots n$ and $j = 1, \ldots n$ *n* are constants.

Example

The function

$$Q(x_1, x_2) = {x_1}^2 + 2x_1x_2 - 3x_2x_1 + 5{x_2}^2$$

is a quadratic form in two variables.

We can write it using matrices

$$Q(x_1, x_2) = \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ -3 & 5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

Note: we can simplify this function

$$Q(x_1, x_2) = {x_1}^2 - x_2x_1 + 5{x_2}^2$$

And write it as

$$Q(x_1, x_2) = \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} 1 & -0.5 \\ -0.5 & 5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

Where the matrix is symmetric.

In general we can write any quadratic form as

$$Q(x) = x'Ax$$

where

- x is the column vector of x_i's and

- A is a symmetric $n \times n$ matrix for which the (i, j)th element is

$$a_{ij} = (1/2)(b_{ij} + b_{ji})$$

note that $x_i x_j = x_j x_i$ for any i and j, so that

$$b_{ij} x_i \ x_j + b_{ji} x_j \ x_i$$

can be written as

$$(b_{ij} + b_{ji}) x_i \ x_j$$

or

$$\frac{1}{2}(b_{ij} + b_{ji}) x_i \ x_j + \frac{1}{2}(b_{ij} + b_{ji}) x_j \ x_i$$

Example

$$Q(x_1, x_2) = {x_1}^2 \ + a x_1 x_2 + b x_2 x_1 - c x_1 x_3 + 5{x_2}^2$$

$$Q(x_1, x_2) = \begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix} \begin{pmatrix} 1 & \frac{a+b}{2} & -\frac{c}{2} \\ \frac{a+b}{2} & 5 & 0 \\ -\frac{c}{2} & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x2 \\ x_3 \end{pmatrix}$$

Conditions for definiteness

With quadratic forms there are ways of establishing whether their signs are positive or negative and this will help determine whether the function of interest is concave or convex

Definition

Let $Q(x)$ be a quadratic form, and let A be the symmetric matrix that represents it (i.e. $Q(x) = x'Ax$).

Then the associated matrix A (and the quadratic form) is:

1. positive definite if $x'Ax > 0$ for all $x \neq 0$
2. negative definite if $x'Ax < 0$ for all $x \neq 0$
3. positive semidefinite if $x'Ax \geq 0$ for all x
4. negative semidefinite if $x'Ax \leq 0$ for all x
5. indefinite if it is neither positive nor negative semidefinite (i.e. if $x'Ax > 0$ for some x and $x'Ax < 0$ for some x).

Examples

1) $ax_1^{\,2} + cx_2^{\,2} = \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} a & 0 \\ 0 & c \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ is positive definite for $a, c > 0$ because $ax_1^{\,2} + cx_2^{\,2} > 0$ for $a, c > 0$ and $\begin{pmatrix} x_1 & x_2 \end{pmatrix} \neq 0$

2) $x_1^{\,2} + 2x_1x_2 + x_2^{\,2} = \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ is positive semidefinite because we can write it as $(x_1 + x_2)^2$ that is non negative for all x_1, x_2

It is not positive definite because for $x_1 = 1, x_2 = -1$ its value is 0.

Positive or Negative definite matrices

Definition:

The ***leading principal matrices*** of a nxn square matrix are the matrices found by deleting

1. The last n-1 rows and columns – to give D_1
2. The last n-2 rows and columns – to give D_2
3. …
4. and the original matrix D_n

Definition:

The ***leading principal minors*** of a matrix are the determinants of these leading principal matrices.

Example:

$$A = \begin{matrix} 1 & 2 \\ -1 & 3 \end{matrix}$$

the leading principal matrices are then

D_2 = A and (D_1 = 1)

and the determinants (leading principal minors) are $|D_2| = 5$

and $|D_1| = 1$

Example 2. Find D_1 D_2 and D_3 of the following matrix

$$A = \begin{matrix} 1 & 0 & -1 \\ 0 & 2 & 0 \\ 2 & -1 & 0 \end{matrix} \rightarrow \quad D_1 = 1, D_2 = \begin{matrix} 1 & 0 \\ 0 & 2 \end{matrix}, \quad D_3 = A$$

$|D_3| = 4, |D_2| = 2$ and $|D_1| = 1$

If a square matrix is negative definite then the leading principal minors have the following signs

$$|D_1| < 0; |D_2| > 0; |D_3| < 0...$$

a positive definite matrix requires leading principal minors are **all** positive, i.e.

$$|D_1| > 0; |D_2| > 0; |D_3| > 0...$$

To check if a square matrix is negative semi-definite we have to compute all principal minors (not only the leading principal minors)

Positive or Negative semidefinite matrices

To obtain conditions for an *n*-variable quadratic form to be positive or negative semidefinite, we need to examine the determinants of some of its submatrices.

Definition:

The ***principal matrices*** of a nxn square matrix are the matrices found by deleting

1. n-1 rows and columns – in all possible combinations
2. n-2 rows and columns – – in all possible combinations
3. …
4. and the original matrix

Definition:

The ***principal minors*** of a matrix are the determinants of the principal matrices.

Let

$$A = \begin{matrix} a & b \\ b & c \end{matrix}$$

The **first-order principal minors** of A are a and c, and the **second-order principal minor** is the determinant of A, namely $ac - b^2$.

Let

$$A = \begin{matrix} 3 & 1 & 2 \\ 1 & -1 & 3 \\ 2 & 3 & 2 \end{matrix}$$

This matrix has 3 **first-order principal minors**, obtained by deleting

- the last two rows and last two columns
- the first and third rows and the first and third columns
- the first two rows and first two columns

which gives us simply the elements on the main diagonal of the matrix: 3, −1, and 2.

The matrix $A = \begin{matrix} 3 & 1 & 2 \\ 1 & -1 & 3 \\ 2 & 3 & 2 \end{matrix}$

also has 3 **second-order principal minors**, obtained by deleting

- the last row and last colunm
- the second row and second column
- the first row and first column

which gives us −4, 2, and −11.

The matrix has one **third-order principal minor**, namely its determinant, −19.

Let A be an $n \times n$ symmetric matrix. Then:

A is **positive semidefinite** if and only if **all** the principal minors of A are nonnegative.

A is **negative semidefinite** if and only if all the k^{th} order principal minors of A are ≤ 0 if k is odd and ≥ 0 if k is even.

Example

$$\begin{matrix} -2 & 4 \\ 4 & -8 \end{matrix}$$

The two first-order principal minors and -2 and -8, and the second-order principal minor is 0. Thus the matrix is negative semidefinite.

Procedures for checking the definiteness of a matrix

1. Find the **leading principal minors** and check if the conditions for positive or negative definiteness are satisfied. If they are, you are done.
2. the conditions are not satisfied, check if they are ***strictly* violated**. If they are, then the matrix is indefinite.
3. If the conditions are **not strictly violated**, find all its principal minors and check if the conditions for positive or negative semidefiniteness are satisfied.

Note: if matrix is positive definite, it is certainly positive semidefinite, and if it is negative definite, it is certainly negative semidefinite

An intuition on quadratic forms

Example with quadratic form in 3 variables

$q = d_{11}x^2 + d_{12}xy + d_{13}xz + d_{21}yx + d_{22}y^2 + d_{23}yz + d_{31}zx + d_{32}zy + d_{33}z^2$

Can be written in matrix form x'Ax where x = (x, y, z) and A is a symmetric 3 by 3 matrix

$$x'Ax = \begin{bmatrix} x & y & z \end{bmatrix} \begin{bmatrix} d_{11} & d_{12} & d_{13} \\ d_{21} & d_{22} & d_{23} \\ d_{31} & d_{32} & d_{33} \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

There are 3 leading principal minors from the discriminants of A

$$|D_1| = |d_{11}|; |D_2| = \begin{vmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \end{vmatrix}; |D_3| = \begin{vmatrix} d_{11} & d_{12} & d_{13} \\ d_{21} & d_{22} & d_{23} \\ d_{31} & d_{32} & d_{33} \end{vmatrix}$$

$$|D_1| = |d_{11}| = d_{11}$$

$$|D_2| = \begin{vmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \end{vmatrix} = d_{11}d_{22} - d_{21}d_{12} = d_{11}d_{22} - d_{21}{}^2$$

$$|D_3| = \begin{vmatrix} d_{11} & d_{12} & d_{13} \\ d_{21} & d_{22} & d_{23} \\ d_{31} & d_{32} & d_{33} \end{vmatrix} =$$

$$= d_{11}d_{22}d_{33} + d_{12}d_{23}d_{31} + d_{13}d_{21}d_{32} - d_{31}d_{22}d_{13} - d_{23}d_{32}d_{11} - d_{33}d_{12}d_{21}$$

$$= d_{11}d_{22}d_{33} + 2d_{12}d_{23}d_{13} - d_{22}d_{13}{}^2 - d_{11}d_{23}{}^2 - d_{33}d_{12}{}^2$$

Once again can convert into an expression where the 3 variables appear only as squared terms

$$q = d_{11}\left(x + \frac{d_{12}}{d_{11}} y + \frac{d_{13}}{d_{11}} z \right)^2 + \frac{d_{11}d_{22} - d_{12}^2}{d_{11}}\left(y + \frac{d_{11}d_{23} - d_{12}d_{13}}{d_{11}d_{22} - d_{12}^2} z \right)^2 +$$

$$\frac{d_{11}d_{22}d_{33} - d_{11}d_{23}^2 - d_{22}d_{13}^2 - d_{33}d_{12}^2 + 2d_{12}d_{13}d_{23}}{d_{11}d_{22} - d_{12}^2}(z)^2$$

And can show that q <0 (>0) iff the terms outside the brackets are all negative (positive)

and these terms are respectively:

$$|D_1|; \frac{|D_2|}{|D_1|}; \frac{|D_3|}{|D_2|}$$

If

$|D_1| < 0, |D_2| > 0, |D_3| < 0$

the matrix is said to be negative definite

if

$|D_1| > 0, |D_2| > 0, |D_3| > 0$

the matrix is said to be positive definite

A second test to check definiteness

Characteristic root test

Given an $n\ x\ n$ matrix D, we find a scalar r and an $n\ x$ 1vector $x \neq 0$ such that:

$$D\,x = r\,x$$

$\boldsymbol{r}$ is the **characteristic root** of matrix D (or eigenvalue)

$\boldsymbol{x}$ is the **characteristic vector** of matrix D (or eigenvector)

This equation is rewritten as:

$$(D - rI)\,x = 0$$

The condition that satisfies this is when the matrix $(D - rI)$ is singular; i.e., its determinant is zero

The idea is to solve for r and then x

Example

$$D = \begin{bmatrix} 2 & 2 \\ 2 & -1 \end{bmatrix} \qquad D - rI = \begin{bmatrix} 2 & 2 \\ 2 & -1 \end{bmatrix} - r\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2-r & 2 \\ 2 & -1-r \end{bmatrix}$$

$$\left| D - rI \right| = \begin{vmatrix} 2-r & 2 \\ 2 & -1-r \end{vmatrix} = r^2 - r - 6 = 0$$

So the characteristic roots are $r_1 = 3$ and $r_2 = -2$

For $r_1 = 3$

$$(D - rI)x = 0 = \begin{bmatrix} 2-3 & 2 \\ 2 & -1-3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -1 & 2 \\ 2 & -4 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Note that the rows of the matrix are linearly dependent – as expected for a singular matrix – giving an infinite number of solutions $x_1 = 2x_2$

To force out a unique solution, we need to ***normalise*** by imposing a restriction:

$$x_1{}^2 + x_2{}^2 = 1$$

and in general for n unknowns $\sum_{i=1}^{n} x_i{}^2 = 1$

This is arbitrary but whichever rule is chosen, all subsequent values will be related

Then $$x_1^2 + x_2^2 = (2x_2)^2 + x_2^2 = 5x_2^2 = 1$$

and $$x_2 = 1/\sqrt{5}; x_1 = 2/\sqrt{5}$$

Thus, the 1st characteristic vector (eigenvector) is $$x1 = \begin{pmatrix} \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{pmatrix}'$$

and for $r = -2$, the 2nd characteristic vector (eigenvector) is

$$x2 = \begin{pmatrix} \frac{-1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{pmatrix}'$$

Properties:

1) normalisation implies that the product of characteristic vectors, i.e. $x1'x1 = 1$
2) Each pair of characteristic vectors are orthogonal,

   i.e. $x1' x2 = 0$

Characteristic root test for the sign definiteness of a matrix D

1. D is positive definite if and only if every characteristic root is positive, i.e. > 0
2. D is negative definite if and only if every characteristic root is negative , i.e. < 0
3. D is positive semidefinite if and only if every characteristic root is nonnegative, i.e. ≥ 0
4. D is negative semidefinite if and only if every characteristic root is nonpositive, i.e. ≤ 0

Finding if a function with more variables is concave: an intuition

Let be
$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad (x' - x_0) = \begin{pmatrix} x_1' - x_{10} \\ \vdots \\ x_n' - x_{n0} \end{pmatrix}$$

We also need the ***vector*** of first partial derivatives of f, and the ***matrix*** of second order partial derivatives, H

$$\nabla f = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{pmatrix}; \quad H = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{pmatrix}$$

J is called ***Jacobian*** of the function *f*

H is called the ***Hessian*** of the function *f*

The concavity condition is now:

$$f(x_0) + (x' - x_0)'\nabla f \geq f(x')$$

The Taylor approximation of $f(x')$ is now

$$f(x') \approx f(x_0) + (x' - x_0)'\nabla f + \frac{1}{2}(x' - x_0)'H(x' - x_0) + \cdots$$

Replacing in the first equation we get

$$f(x_0) + (x' - x_0)'\nabla f$$

$$\geq f(x_0) + (x' - x_0)'\nabla f + \frac{1}{2}(x' - x_0)'H(x' - x_0) + \cdots$$

Simplifying we get

$$0 \geq (x' - x_0)'H(x' - x_0)$$

Then matrix H has to be a negative semi-definite matrix

Conditions for concavity / convexity

Let *f* be a function of many variables with continuous partial derivatives of first and second order on the convex open set *S* and denote the Hessian of *f* at the point *x* by *H*(*x*). Then *f* is:

- concave **if and only if** *H*(*x*) is negative semidefinite $\text{for } \forall x \in S$
- convex **if and only if** *H*(*x*) is positive semidefinite $\text{for } \forall x \in S$

if *H*(*x*) is:

- negative definite $\text{for } \forall x \in S$ then *f* is strictly concave
- positive definite $\text{for } \forall x \in S$ then *f* is strictly convex.

Putting it all together

So given a function f(x)

To find out whether the function is concave we need to know if

$$0 \geq (x' - x_0)' H (x' - x_0)$$

i.e. whether H is negative semi-definite

1. Find the Hessian matrix of second order derivatives, H

2. From H find the leading principal matrices by eliminating:
    1. The last n-1 rows and columns – written as D_1
    2. The last n-2 rows and columns – written as D_2
    3. …
    4. The original matrix D_n

3. Compute the determinants of these leading principal matrices

4. if the determinants have the following pattern (with not all zero): $|D_1| < 0, |D_2| > 0,\ |D_3| < 0$ ... ., then f is (strictly) concave; if the determinants are all strictly positive then f is (strictly) convex

5. if some condition is violated by equality you need to check the sign of all principal minors (condition or semidefiniteness)

6. if these conditions do not hold you've proved that the function is not concave or convex

Example: Find whether the function $f(x) = -x_1x_2^2$ is concave

We need the Hessian matrix of second order derivatives, H

- The Jacobian is

$$\begin{pmatrix} \dfrac{df}{dx_1} \\ \dfrac{df}{dx_2} \end{pmatrix} = \begin{pmatrix} -x_2^2 \\ -2x_1x_2 \end{pmatrix}$$

- The Hessian is

$$H = \begin{pmatrix} \dfrac{\partial^2 f}{\partial x_1^2} & \dfrac{\partial^2 f}{\partial x_1 \partial x_2} \\ \dfrac{\partial^2 f}{\partial x_2 \partial x_1} & \dfrac{\partial^2 f}{\partial x_2^2} \end{pmatrix} = \begin{pmatrix} 0 & -2x_2 \\ -2x_2 & -2x_1 \end{pmatrix}$$

From H find the leading principal matrices by eliminating:

1. The last n-1 rows and columns – written as $D_1 = (0)$
2. The last n-2 rows and columns – written as $D_2 = H$

Compute the determinants of these leading principal matrices.

1. Det. $D_1 = 0$
2. Det. $H = -4x_2^2$ which is negative

f is concave if the leading principal minors are $|D_1| < 0; |D_2| > 0;$

f is convex if the leading principal minors are $|D_1| > 0; |D_2| > 0;$

Leading principal minors do not have one of this patterns so f is not concave, not convex