ECON 508-A

(Quasi)Concavity and (Quasi)Convexity.

Juan Ignacio Vizcaino

Washington University in St. Louis

July 29 - August 16 2019

The material here is based on the slides on Concavity, Quasiconcavity, Convexity and Quasiconvexity prepared by Carmen Aston-Figari.

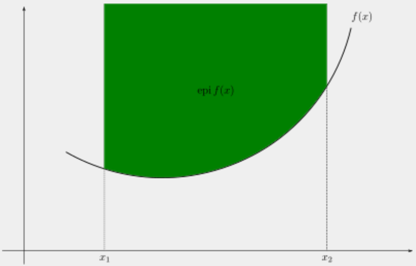
All errors are mine.

Section 1: Concavity and Quasi-concavity

Let $C \subseteq \mathbb{R}^N$ be a convex set, and $f : C \to \mathbb{R}$.

Definition

The hypograph of f (the epigraph of f), denoted *hyp f* (denoted *epi f*), is the set of points lying on and below the graph of f:

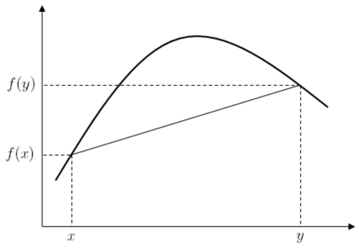


$$hyp\, f = \left\{(x, y) \in \mathbb{R}^{N+1} : x \in \mathbb{R}^N \text{ and } y \leq f(x)\right\}$$

$$epi\, f = \left\{(x, y) \in \mathbb{R}^{N+1} : x \in \mathbb{R}^N \text{ and } y \geq f(x)\right\}$$

Definition
Let $x, y \in C$, and $\alpha \in [0,1]$. f is concave iff

$$f(\alpha x + (1-\alpha)y) \geq \alpha f(x) + (1-\alpha)f(y)$$



Theorem 1

f is concave iff the set *hyp* f is convex.

Proof: Omitted. □

Definition f is convex iff $-f$ is concave.

Theorem 2

The following are equivalent:

1. f is convex.
2. *epi* f is a convex set.
3. For any $x, y \in C$, and $\alpha \in [0, 1]$
   $f(\alpha x + (1 - \alpha)y) \leq \alpha f(x) + (1 - \alpha) f(y)$.

Proof: Omitted. □

Definition

1. f is **strictly concave** iff for any $x, y \in C$, and $\alpha \in [0, 1]$
   $f(\alpha x + (1 - \alpha)y) > \alpha f(x) + (1 - \alpha)f(y)$
2. f is **strictly convex** iff $-f$ is strictly concave.

Theorem 3

Let $f_1, \ldots f_k$ be concave functions, each defined on C. Let $a_1, \ldots a_k \in \mathbb{R}_+$. Then $a_1 f_1 + \ldots + a_k f_k$ is a concave function on C.

Proof: Omitted. □

Theorem 4

Let $f : C \to \mathbb{R}$ where C is open and convex. If f is concave on C, it is continuous on C. More generally f is continuous on the interior of C.

Proof: Omitted. □

Useful Properties

1. If f is concave and g is concave and increasing, then $U(x) = g(f(x))$ is concave.
2. If f is convex and g is convex and increasing, then $U(x) = g(f(x))$ is convex.
3. Let f and g be concave functions. Then $h(x) = min\{f(x), g(x)\}$ is concave.
4. Let f and g be convex functions. Then $h(x) = max\{f(x), g(x)\}$ is convex.

Definition

1. f is quasiconcave iff $\forall\, x, y \in C, z = \alpha x + (1-\alpha)y, \alpha \in [0,1]$.

$$f(z) \geq \min\{f(x), f(y)\}$$

2. f is strictly quasiconcave iff
   $\forall\, x, y \in C, z = \alpha x + (1-\alpha)y, \alpha \in (0,1)$.

$$f(z) > \min\{f(x), f(y)\}$$

The function f is quasiconvex iff $-f$ is quasiconcave.
The function f strictly quasiconvex iff $-f$ is strictly quasiconcave.

Quasiconcavity and Upper Contour Sets

Given a point a *in* $\mathbb{R}$, the upper contour set of f is the set

$$U_a = \{x \in C : f(x) \geq a\}$$

Theorem 5

f is quasiconcave iff the upper contour sets are convex.

Proof: Omitted. □

Useful Theorems

Theorem 6

let f be **quasiconcave**, and $g : \mathbb{R} \rightarrow \mathbb{R}$ be a weakly increasing function defined on an interval I that contains $f(C)$. Then the composite function $g \circ f(x)$ is **quasiconcave** in C.

Proof: Omitted. □

Theorem 7

If f is concave, then f is quasiconcave.

Proof: Omitted. □

Section 2: Definite Matrices

Definition 1

A quadratic form on $\mathbb{R}^N$ is a real valued function of the form

$$Q(x_1, \dots, x_N) = \sum_{i=1}^{N} \sum_{j=1}^{N} a_{ij} x_i x_j$$

where each term is a monomial of degree two.

Matrix Representation of a Quadratic Form

Let $x = (x_1, \ldots, x_N)$

Then $Q(x)$ can also be represented in matrix form:

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

where A is a symmetric $N \times N$ matrix.

The general quadratic form in $\mathbb{R}^2$:

$$a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2$$

can be written in matrix terms:

$$\begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Quadratic Forms: Example 2

The general quadratic form in $\mathbb{R}^3$:

$$a_{11}x_1^2 + a_{22}x_2^2 + a_{33}x_3^2 + 2a_{12}x_1x_2 + 2a_{13}x_1x_3 + 2a_{23}x_2x_3$$

can be written in matrix terms:

$$\begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

And so forth

Definition 2

Let A be a symmetric $N \times N$ matrix. Then A is:

1. Positive Definite (PD) iff $Q(x) = x'Ax > 0 \quad \forall x \in \mathbb{R}^N \backslash \{0\}$
2. Positive Semidefinite (PSD) iff $Q(x) = x'Ax \geq 0 \quad \forall x \in \mathbb{R}^N \backslash \{0\}$
3. Negative Definite (ND) iff $Q(x) = x'Ax < 0 \forall x \in \mathbb{R}^N \backslash \{0\}$
4. Negative Semidefinite (NSD) iff $Q(x) = x'Ax \leq 0 \forall x \in \mathbb{R}^N \backslash \{0\}$
5. Indefinite if $Q(x) = x'Ax > 0$ for some $x \in \mathbb{R}^N$, and $Q(x) = x'Ax < 0$ for some other $x \in \mathbb{R}^N$

A matrix that is $PD(ND)$ is also PSD (NSD).

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I$$

Pick any $x \in \mathbb{R}^2 \backslash \{0\}$

$$Q(x) = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1^2 + x_2^2 > 0$$

So A is PD

Definite Matrices: Example 2

$$C = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

Pick any $x \in \mathbb{R}^2 \backslash \{0\}$

$$Q(x) = x'Cx = 2x_1x_2$$

$$x^T Cx = 2 > 0 \text{ for } x = (1, 1)$$

$$x^T Cx = -2 < 0 \text{ for } x = (-1, 1)$$

So C is Indefinite

Identifying Definitness and Semidefinitness

Two different but equivalent approaches:

1. Eigenvalues
2. Principal minors

Eigenvalues

Let A be a $N \times N$ square matrix.

The following equation is called characteristic equation:

$$\det[A - \lambda I] = 0$$

The solutions to the characteristic equation are called **characteristic roots** or **eigenvalues**.

$$A = \begin{bmatrix} 2 & -6 \\ -6 & -7 \end{bmatrix}$$

The characteristic equation is:

$$\det[A - \lambda I] = 0$$

$$\left| \begin{bmatrix} 2 & -6 \\ -6 & -7 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} \right| = 0$$

$$\left| \begin{bmatrix} 2 - \lambda & -6 \\ -6 & -7 - \lambda \end{bmatrix} \right| = 0$$

Eigenvalues: Example 1 continued

$$(2-\lambda)(-7-\lambda)-(-6)(-6)=0$$

$$\lambda^2+5\lambda-50=0$$

$$(\lambda-5)(\lambda+10)=0$$

$$\lambda_1 = 5$$
$$\lambda_2 = -10$$

Identifying Definitness Using Eigenvalues

Theorem

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

- A is **PD** iff all its eigenvalues are positive.
- A is **PSD** iff all its eigenvalues are nonnegative.
- A is **ND** iff all its eigenvalues are negative.
- A is **NSD** iff all its eigenvalues are nonpositive.
- A is **indefinite** iff it has at least one positive eigenvalue and at least one negative eigenvalue.

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

The characteristic equation is

$$|I - \lambda I| = \left| \begin{bmatrix} 1-\lambda & 0 \\ 0 & 1-\lambda \end{bmatrix} \right| = (1-\lambda)^2$$

$$\lambda_1 = \lambda_2 = 1$$

So A is PD

Example 2

$$B = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

The characteristic equation is

$$\left| \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} - \lambda I = \left| \begin{bmatrix} 1-\lambda & 0 \\ 0 & -\lambda \end{bmatrix} \right| \right| = (1-\lambda)(-\lambda) = 0$$

$\lambda_1 = 0$
$\lambda_2 = 1$

So B is PSD

Example 3

$$C = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

The characteristic equation is

$$\left| \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} - \lambda I = \left| \begin{bmatrix} -\lambda & 1 \\ 1 & -\lambda \end{bmatrix} \right| \right| = (-\lambda)^2 - 1 = 0$$

$$\lambda_1 = -1$$
$$\lambda_2 = 1$$

So C is Indefinite

Definition 3

Let A be an $N \times N$ square matrix.

1. The $K \times K$ submatrix obtained from A by deleting any $(N-K)$ columns of A and the corresponding $(N-K)$ rows of A is called K -order principal submatrix of A.
2. The determinant of a K -order principal submatrix of A is called a K -order principal minor (principal minor (**PM**) of order K).
3. The K -order principal submatrix of A obtained by deleting the last $(N-K)$ columns of A and the last $(N-K)$ rows of A is called the K -order leading principal submatrix of A, denoted A_K. (An $N \times N$ matrix has N leading principal submatrices).
4. The determinalnt of A_K is called the K -order leading principal minor **LPM**, denoted $|A_K|$.

Example

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

Leading principal submatrices and leading principal minors:

$$A_1 = [a_{11}]$$

$$A_2 = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

$$A_3 = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

$$|A_1| = a_{11}$$

$$|A_2| = a_{11}a_{22} - a_{12}a_{21}$$

Principal submatrices and principal minors:
Of order 1:

$$[a_{22}] \quad a_{22}$$
$$[a_{33}] \quad a_{33}$$

Of order 2:

$$\begin{bmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{bmatrix} \quad \left| \begin{bmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{bmatrix} \right| = a_{11}a_{33} - a_{13}a_{31}$$

$$\begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix} \quad \left| \begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix} \right| = a_{22}a_{33} - a_{23}a_{32}$$

Identifying Definitness Using Principal Minors

Theorem

Let A be an $N \times N$ symmetric matrix. Then

1. A is **PD** iff all its **LPM**'s are positive.
2. A is **PSD** iff all its **PM**′ s are nonnegative.
3. A is **ND** iff its LPM's alternate in signs with $(-1)^K |A_K| > 0$:
   - every LPM of odd order is negative
   - every LPM of even order is positive
4. A is **NSD** iff every **PM** of odd order is nonpositive every **PM** of even order is nonnegative.
5. A is **indefinite** iff some **LPM**'s of order k are nonzero but do not fit into (1) or (3).

Example 1

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Look at the A_K' s and LPM's first:

$A_1 = \begin{bmatrix} 1 \end{bmatrix}, \qquad |A_1| = 1$

$A_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad |A_2| = 1$

So A is **PD**

Example 2

$$B = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

Look at the B_K' s and LPM's first:

$$B_1 = \begin{bmatrix} 1 \end{bmatrix}, \qquad |B_1| = 1$$

$$B_2 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \quad |B_2| = 0$$

So we knbw that B is not PD. Is it PSD?
Check all the remaining principal submatrices and PM's
Of order 1 :

$$\begin{bmatrix} 0 \end{bmatrix} \quad |[0]| = 0$$

So B is PSD

$$C = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

Look at the C_K' s and LPM's first:

$C_1 = \begin{bmatrix} 0 \end{bmatrix}, \qquad |C_1| = 0$

$C_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad |C_2| = -1$

Since the second order LPM is negative, C doesn't fit into any category

So C is Indefinite

Section 3: Concavity and D^2f

CONCAVITY AND D^2f

Theorem

Let $C \subseteq \mathbb{R}^N$ be convex and $f : C \to \mathbb{R}$ be C^2.
Then

1. $D^2f(x)\ \forall\, x \in C$ is NSD iff f is concave.
2. $D^2f(x)\ \forall\, x \in C$ is PSD iff f is convex.
3. $D^2f(x)\ \forall\, x \in C$ is ND then f is strictly concave.
4. $D^2f(x)\ \forall\, x \in C$ is PD then f is strictly convex.

Remark: in (3) and (4), negative and positive definiteness areonly sufficient.