Mathematical Methods for Economic Dynamics (ECON 696)

Lecture 3: Linear Algebra II (Ch. 1) and Multivariate Calculus (Ch. 2)

The Rank of a Matrix (Ch. 1.3)

Definition: The **rank** of a matrix A, written rank(A), is the maximum number of linearly independent column vectors in A. If A is the $\mathbf{0}$ matrix, we put rank(A) $= 0$.

$A = (\mathbf{a}_1, \ldots, \mathbf{a}_n)$: $n \times n$ matrix

rank(A) $= n \Rightarrow |A| \neq 0$;

rank(A) $< n \Rightarrow |A| = 0$.

Theorem: Let $A = (\mathbf{a}_1, \ldots, \mathbf{a}_n)$ be an $n \times n$ matrix. Then, $|A| \neq 0$ if and only if $\mathbf{a}_1, \ldots, \mathbf{a}_n$ are linearly independent.

Proof: We skip the proof. ■

Example

$$A = \begin{pmatrix} 1 & 0 & 2 & 1 \\ 0 & 2 & 4 & 2 \\ 0 & 2 & 2 & 1 \end{pmatrix}$$

We can easily see that the third and fourth columns are parallel to each other:

$$\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 2 \\ 4 \\ 2 \end{pmatrix}.$$

So, we now consider the following truncated matrix:

$$B = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 2 & 4 \\ 0 & 2 & 2 \end{pmatrix}$$

We compute the determinant of the truncated matrix:

$$\begin{aligned} \begin{vmatrix} 1 & 0 & 2 \\ 0 & 2 & 4 \\ 0 & 2 & 2 \end{vmatrix} &= 1 \cdot (-1)^{1+1} \begin{vmatrix} 2 & 4 \\ 2 & 2 \end{vmatrix} \\ &= 4 - 8 = -4 \neq 0. \end{aligned}$$

Since $|B| \neq 0$, all the columns vectors in B are linearly independent.

Thus, $\text{rank}(A) = 3$.

Main Results on Linear Systems (Ch. 1.4)

Consider the general **system of linear equations**:

$$\begin{array}{rcl} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n & = & b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n & = & b_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots & . & . \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots & . & . \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots & . & . \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n & = & b_m \end{array} \qquad (*)$$

can be written as $Ax = b$ where

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, \; x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \; b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

Let

$$A_b = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{pmatrix}$$

be the **augmented** matrix of the system $(*)$.

It turns out that the relationship between the ranks of A and A_b is crucial in determining whether system $(*)$ has a solution.

Because all the columns in A occur in A_b, the rank of A_b is certainly greater than or equal to the rank of A. Moreover, because A_b contains only one more column than A, rank$(A_b) \leq$ rank$(A) + 1$.

Theorem: $Ax = b$ has at least one solution if and only if rank$(A) =$ rank(A_b).

Proof: We omit the proof. ■

Remark: When we say $Ax = b$ has at least one solution, there may well be multiple solutions to $Ax = b$.

Theorem: Suppose that system $(*)$ has solutions with $\text{rank}(A) = \text{rank}(A_b) = k$.

1. If $k < m$, then $m - k$ equations are **superfluous** in the sense that if we choose any subsystem of equations corresponding to k linearly independent rows, then any solution of these k equations also satisfies the remaining $m - k$ equations.

2. If $k < n$, there exist $n - k$ variables that can be chosen freely, whereas the remaining k variables are uniquely determined by the choice of these $n - k$ free variables. The system then has $n - k$ **degrees of freedom**.

Proof: We omit the proof. ■

Quadratic Forms (Ch. 1.7, 1.8)

A **quadratic form** in n variables is a function $Q : \mathbb{R}^n \to \mathbb{R}$ of the form:

$$\begin{aligned} Q(x_1, \ldots, x_n) &= \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j \\ &= a_{11} x_1^2 + a_{12} x_1 x_2 + \cdots + a_{ij} x_i x_j + \cdots + a_{nn} x_n^2. \end{aligned}$$

where the a_{ij} are constants.

Let $x = (x_1, \ldots, x_n)^T$ and $A = (a_{ij})$. Then, it follows that

$$Q(x_1, \ldots, x_n) = Q(x) = x^T A x = x \cdot Ax.$$

Definition: A quadratic form $Q(x) = x^T Ax = x \cdot Ax$, as well as its associated matrix A, are said to be **positive definite, positive semidefinite, negative definite**, or **negative semidefinite** according as

$$Q(x) > 0,\ Q(x) \geq 0,\ Q(x) < 0,\ Q(x) \leq 0,$$

for all $x \in \mathbb{R}^n \backslash \{\mathbf{0}\}$. The quadratic form $Q(x)$ is **indefinite** if there exist vectors x^* and y^* in $\mathbb{R}^n \backslash \{\mathbf{0}\}$ such that $Q(x^*) < 0$ and $Q(y^*) > 0$.

Let $A = (a_{ij})$ be any $n \times n$ matrix.

An arbitrary **principal minor** of A of **order** r is the determinant of the matrix obtained by deleting all but r rows and r columns such that if the i-th row (column) is selected, then so is the i-th column (row).

In particular, a principal minor of order r always includes exactly r elements of the main (principal) diagonal.

The determinant $|A|$ itself is called a principal minor (no rows and columns are deleted).

Let Δ_r denote an arbitrary principal minor of A of order r.

A principal minor is said to be a **leading principal minor** of order r $(1 \leq r \leq n)$, if it consists of the first "leading" rows and columns of $|A|$. The leading principal minors of A of order r is

$$D_r = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1r} \\ a_{21} & a_{22} & \cdots & a_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ a_{r1} & a_{r2} & \cdots & a_{rr} \end{vmatrix}.$$

where D_r denotes the leading principal minor of A of order r

Theorem: Consider the quadratic form $Q = x^T Ax$ with the associated symmetric matrix $A = (a_{ij})_{n\times n}$. Then

1. Q is positive definite $\Leftrightarrow D_r > 0$ for $r = 1, \ldots, n$

2. Q is positive semidefinite $\Leftrightarrow \Delta_r \geq 0$ for all Δ_r and $r = 1, \ldots, n$.

3. Q is negative definite $\Leftrightarrow (-1)^r D_r > 0$ for $r = 1, \ldots, n$

4. Q is negative semidefinite $\Leftrightarrow (-1)^r \Delta_r \geq 0$ for all Δ_r and $r = 1, \ldots, n$.

Proof: We only prove this for $n = 2$. Then, the quadratic form is

$$Q(x_1, x_2) = a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2$$

After some manipulation through **perfect square**, we obtain

$$Q(x_1, x_2) = a_{11}\underbrace{\left(x_1 + \frac{a_{12}}{a_{11}}x_2\right)^2}_{>0} + \left(a_{22} - \frac{a_{12}^2}{a_{11}}\right)\underbrace{x_2^2}_{\geq 0}$$

Thus, we obtain

$$Q(x_1, x_2) > 0 \ \forall x_1, x_2 \Longleftrightarrow a_{11} > 0 \text{ and } a_{11}a_{22} - a_{12}^2 > 0.$$

When $a_{11} < 0$, we observe the following:

$$a_{11}a_{22} - a_{12}^2 > 0 \Leftrightarrow a_{22} - \frac{a_{12}^2}{a_{11}} < 0.$$

Therefore, we obtain

$$Q(x_1, x_2) < 0 \ \forall x_1, x_2 \Longleftrightarrow a_{11} < 0 \text{ and } a_{11}a_{22} - a_{12}^2 > 0. \ \blacksquare$$

Quadratic Forms with Linear Constraints (Ch. 1.8)

In constrained optimization theory, the second-order conditions involve the signs of quadratic forms subject to homogeneous linear constraints.

Let $Q = a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2$ and assume that the variables are subject to the homogeneous linear constraint $b_1x_1 + b_2x_2 = 0$, where $b_1 \neq 0$.

$b_1 \neq 0 \Rightarrow x_1 = -b_2x_2/b_1$.

Plugging $x_1 = -b_2x_2/b_1$ into Q, we obtain

$$\begin{aligned} Q &= a_{11}\left(-\frac{b_2x_2}{b_1}\right)^2 + 2a_{12}\left(-\frac{b_2x_2}{b_1}\right)x_2 + a_{22}x_2^2 \\ &= \frac{1}{b_1^2}\left(a_{11}b_2^2 - 2a_{12}b_1b_2 + a_{22}b_1^2\right)x_2^2. \end{aligned}$$

It is easy to see

$$a_{11}b_2^2 - 2a_{12}b_1b_2 + a_{22}b_1^2 = -\begin{vmatrix} 0 & b_1 & b_2 \\ b_1 & a_{11} & a_{12} \\ b_2 & a_{12} & a_{22} \end{vmatrix}.$$

Therefore,

$$Q \text{ is PD subject to } b_1x_1 + b_2x_2 = 0 \;\Leftrightarrow \begin{vmatrix} 0 & b_1 & b_2 \\ b_1 & a_{11} & a_{12} \\ b_2 & a_{12} & a_{22} \end{vmatrix} < 0.$$

$$Q \text{ is ND subject to } b_1x_1 + b_2x_2 = 0 \;\Leftrightarrow \begin{vmatrix} 0 & b_1 & b_2 \\ b_1 & a_{11} & a_{12} \\ b_2 & a_{12} & a_{22} \end{vmatrix} > 0.$$

This is also valid when $b_1 = 0$ but $b_2 \neq 0$.

General Case of Q when there are the linear constraints

$$Q(x) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j \quad (a_{ij} = a_{ji})$$

subject to m linear homogeneous constraints:

$$\begin{array}{rcl} b_{11}x_1 + \cdots + b_{1n}x_n & = & 0 \\ b_{21}x_1 + \cdots + b_{2n}x_n & = & 0 \\ \vdots \qquad\quad \vdots \qquad\quad \vdots & = & \vdots \Leftrightarrow \mathbf{Bx} = \mathbf{0}\ (*) \\ b_{m1}x_1 + \cdots + b_{mn}x_n & = & 0, \end{array}$$

where $\mathbf{B}$ is an $m \times n$ matrix.

PD/ND when there are the linear constraints

Definition:

(1) Q is PD subject to the linear constraints $(*)$ if $Q(x) > 0$ for all $x = (x_1, \ldots, x_n) \neq \mathbf{0}$ that satisfy $(*)$

(2) Q is ND subject to the linear constraints $(*)$ if $Q(x) < 0$ for all $x = (x_1, \ldots, x_n) \neq \mathbf{0}$ that satisfy $(*)$

Define the symmetric determinants

$$B_r = \begin{vmatrix} 0 & \cdots & 0 & b_{11} & \cdots & b_{1r} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & b_{m1} & \cdots & b_{mr} \\ b_{11} & \cdots & b_{m1} & a_{11} & \cdots & a_{1r} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ b_{1r} & \cdots & b_{mr} & a_{r1} & \cdots & a_{rr} \end{vmatrix}$$

The determinant B_r is the $(m+r)$th leading principal minor of the $(m+n) \times (m+n)$ bordered matrix

$$\begin{pmatrix} \mathbf{0}_{m\times m} & \mathbf{B} \\ \mathbf{B}^T & \mathbf{A} \end{pmatrix}.$$

Theorem: Assume that rank$(B) = m$. Then, a necessary and sufficient condition for the quadratic form Q to be positive definite (PD) subject to the linear constraints $(*)$ is

$$(-1)^m B_r > 0,\ \forall r = m+1, \ldots, n.$$

The corresponding necessary and sufficient condition for Q to be negative definite (ND) subject to the linear constraints $(*)$ is

$$(-1)^r B_r > 0,\ \forall r = m+1, \ldots, n.$$

Proof: We omit the proof. ■

Multivariate Calculus

Real-Valued Functions of Several Variables

$f : D \to R$ is a **real-valued** function if D is any nonempty set in $\mathbb{R}^n$ and $R \subseteq \mathbb{R}$.

Rather than having a single slope, a function of n variables can be thought to have n **partial slopes**, each giving only the rate at which y would change if one x_i alone were to change.

Each of these partial slopes is called the **partial derivative**.

Definition: Let $y = f(x_1, \ldots, x_n)$. The **partial** derivative of f with respect to x_i is defined as

$$\frac{\partial f(x)}{\partial x_i} \equiv \lim_{h \to 0} \frac{f(x_1, \ldots, x_i + h, \ldots, x_n) - f(x_1, \ldots, x_i, \ldots, x_n)}{h}$$

$\partial y / \partial x_i$ or $f_i(x)$ are used to denote partial derivatives.

Gradients (Ch. 2.1)

If $z = F(x, y)$ and C is any number, we call the graph of the equation $F(x, y) = C$ a **level curve** for F.

The slope of the level curve $F(x, y) = C$ at a point (x, y) is given by the formula:

$$F(x,y) = C \Rightarrow y' = \frac{dy}{dx} = -\frac{\partial F(x,y)/\partial x}{\partial F(x,y)/\partial y} = -\frac{F_1(x,y)}{F_2(x,y)},$$

where $\partial F(x, y)/\partial y \neq 0$.

If (x_0, y_0) is a particular point on the level curve $F(x, y) = C$, the slope at (x_0, y_0) is $-F_1(x_0, y_0)/F_2(x_0, y_0)$.

The equation for the **tangent hyperplane** T is

$$y - y_0 = -\left[F_1(x_0, y_0)/F_2(x_0, y_0)\right](x - x_0)$$

or, rearranging

$$F_1(x_0, y_0)(x - x_0) + F_2(x_0, y_0)(y - y_0) = 0.$$

Recalling the inner product, the equation can be written as

$$(F_1(x_0, y_0), F_2(x_0, y_0)) \cdot (x - x_0, y - y_0) = 0.$$

The vector $(F_1(x_0, y_0), F_2(x_0, y_0))$ is the **gradient** of F at (x_0, y_0) and often denoted by $\nabla F(x_0, y_0)$ (∇ is pronounced as "nabla").

Then, the previous equation can be rewritten as

$$\nabla F(x^0, y^0) \cdot dx = 0,$$

where $dx = (x - x^0, y - y^0)$.

The vector $(x - x_0, y - y_0)$ is a vector on the tangent hyperplane T which implies that $\nabla F(x_0, y_0)$ is **orthogonal** to the tangent hyperplane T at (x_0, y_0).

Examples of Gradients

(i) $F(x,y) = xy$ and $(x^0, y^0) = (1/2, 1)$.

$$\nabla F(x,y) = (\partial F(x,y)/\partial x, \partial F(x,y)/\partial y) = (y,x).$$

So,

$$\nabla F(x^0, y^0) = (1, 1/2).$$

(ii) $F(x,y,z) = xe^{xy} - z^2$ and $(x^0, y^0, z^0) = (0,0,1)$.

$\nabla F(x,y,z) = (F_x, F_y, F_z) = (e^{xy} + xye^{xy}, x^2e^{xy}, -2z)$ So,

$$\nabla F(x^0, y^0, z^0) = (1, 0, -2).$$

Example: Indifference Curves

Let $u : \mathbb{R}^2_+ \to \mathbb{R}$ be the utility function for a consumer.

Fix $x^0 \in \mathbb{R}^2_+$. If we move from x^0 to $x^0 + dx$, the resulting change of utility becomes

$$du = \frac{\partial u(x^0)}{\partial x_1} dx_1 + \frac{\partial u(x^0)}{\partial x_2} dx_2 \text{ (differentials).}$$

If we satisfy $u(x^0 + dx) = u(x^0)$, we must have $du = 0$, i.e.,

$$\underbrace{-\frac{dx_2}{dx_1}\bigg|_{x=x^0}}_{\text{Marginal Rate of Substitution}} = \frac{\partial u(x^0)/\partial x_1}{\partial u(x^0)/\partial x_2}.$$

Example: Budget Frontier as a Tangent Hyperplane at Optimum

In the standard consumer optimization problem, the optimal choice is determined such that the marginal rate of substitution is equalized to the price ratio. That is,

$$\frac{\partial u(x)/\partial x_1}{\partial u(x)/\partial x_2} = \frac{p_1}{p_2}.$$

This implies that $\nabla u(x^*)$ and p are parallel to each other at the optimal consumption x^*.

Therefore, there exists some number $\lambda \in \mathbb{R}$ such that $\nabla u(x^*) = \lambda p$.

Gradients (Ch. 2.1)

Suppose that $F(x) = F(x_1, \dots, x_n)$ is a function of n variables defined on an open set S in $\mathbb{R}^n$, and let $x^0 = (x_1^0, \dots, x_n^0)$ be a point in S.

The **gradient** of F at x^0 is the vector

$$\nabla F(x^0) = \left(\frac{\partial F(x^0)}{\partial x_1}, \cdots, \frac{\partial F(x^0)}{\partial x_n}\right)$$

of first-order partial derivatives.

T_{x^0}: the tangent hyperplane that passes through $x^0 \in \mathbb{R}^n$. Then,

$$T_{x^0} = \left\{x \in \mathbb{R}^n | \ \nabla F(x^0) \cdot (x - x^0) = 0\right\}.$$

Example: Budget Frontier as a Tangent Hyperplane at Optimum

In the standard consumer optimization problem, the optimal choice is determined such that the marginal rate of substitution is equalized to the price ratio. That is,

$$\frac{\partial u(x)/\partial x_1}{\partial u(x)/\partial x_2} = \frac{p_1}{p_2}.$$

This implies that $\nabla u(x^*)$ and p are parallel to each other at the optimal consumption x^*.

Therefore, there exists some number $\lambda \in \mathbb{R}$ such that $\nabla u(x^*) = \lambda p$.

Definition: A function $f : S \to \mathbb{R}$ is **continuously differentiable** (or C^1) on an open set $S \subseteq \mathbb{R}^n$ if, for each $i = 1, \ldots, n$, $(\partial f / \partial x_i)(x)$ exists for all $x \in S$ and is continuous on S.

Definition: f is k-**times continuously differentiable** or C^k on S if all the derivatives of f of order less than or equal to $k (\geq 1)$ exist and they are continuous on S.

Convex Sets (Ch. 2.2)

Convex sets are basic building blocks in virtually every area of Economics.

Convexity guarantees that the analysis is mathematically tractable and the results are clear-cut and "well-behaved."

Definition: $S \subseteq \mathbb{R}^n$ is a **convex** set if for all $x, y \in S$, we have

$$\alpha x + (1-\alpha)y \in S,$$

for all $\alpha \in [0,1]$.

z is called a **convex combination** of x and y if $z = \alpha x + (1-\alpha)y$ for some $\alpha \in [0,1]$.

A set is convex $\Leftrightarrow$ we can connect any two points in the set by a straight line that lies entirely within the set.

NOTE: Intuitively, a convex set must be "connected" without any "holes," and its boundary must not "bend inwards" at any point.

Example: Suppose that $p_i > 0$ for each $i = 1, \ldots, n$ and $w \geq 0$. Let $B(p, w) = \{x \in \mathbb{R}^n_+ | p \cdot x \leq w\}$ be the budget set of the consumer.

Fix $x, x' \in B(p, w)$. Fix also $\alpha \in [0, 1]$.

Define $x^\alpha = \alpha x + (1 - \alpha)x'$. Since $x, x' \in B(p, w)$, we know

$$p \cdot x \leq w \text{ and } p \cdot x' \leq w.$$

Then,

$$p \cdot x^\alpha = p \cdot (\alpha x + (1 - \alpha)x') = \alpha p \cdot x + (1 - \alpha)p \cdot x' \leq \alpha w + (1 - \alpha)w = w.$$

Thus, $x^\alpha \in B(p, w)$ so that $B(p, w)$ is convex.

Concave/Convex Functions (Ch. 2.3, 2.4)

Concave and Convex Functions in One Variable

Let I be an interval on $\mathbb{R}$.

A C^2 function $f : I \to \mathbb{R}$ is said to be **concave (convex)** on the interval I if $f''(x) \le (\ge)\ 0$ for all $x \in I$.

A C^2 function $f : \mathbb{R} \to \mathbb{R}$ is **strictly concave (strictly convex)** if $f''(x) < (>) 0$ for all $x \in I$.

Concave and Convex Functions in Several Variables (Ch. 2.3)

Let S be a convex, open subset of $\mathbb{R}^n$ throughout.

Definition: A function $f : S \to \mathbb{R}$ is **concave** on S if, $\forall x, x^{'} \in S$, $\forall \lambda \in [0, 1]$,

$$f(\lambda x + (1 - \lambda)x^{'}) \geq \lambda f(x) + (1 - \lambda)f(x^{'}).$$

Similarly, $f : S \to \mathbb{R}$ is **convex** if $-f(\cdot)$ is concave.

Definition: A function $f : S \to \mathbb{R}$ is **strictly concave** on S if, $\forall x, x^{'} \in S$ with $x \neq x^{'}$, $\forall \lambda \in (0, 1)$,

$$f(\lambda x + (1 - \lambda)x^{'}) > \lambda f(x) + (1 - \lambda)f(x^{'}).$$

Similarly, $f : S \to \mathbb{R}$ is **strictly convex** if $-f(\cdot)$ is strictly concave.

Concavity/Convexity via Second Derivatives (Ch. 2.3)

Suppose that $f : S \to \mathbb{R}$ is a C^2 function. The **Hessian** matrix of $f(\cdot)$ at x:

$$D^2 f(x) = \Big(f_{ij}(x)\Big)_{n\times n}.$$

For each $r = 1, \ldots, n$,

$$D^2_{(r)} f(x) = \begin{vmatrix} f_{11}(x) & f_{12}(x) & \cdots & f_{1r}(x) \\ f_{21}(x) & f_{22}(x) & \cdots & f_{2r}(x) \\ \vdots & \vdots & \ddots & \vdots \\ f_{r1}(x) & f_{r2}(x) & \cdots & f_{rr}(x) \end{vmatrix}$$

is the **leading principal minors** of $D^2 f(x)$ of order r.

$f_{ij}(x) = \partial^2 f(x)/\partial x_i \partial x_j$ for any $i, j = 1, \ldots, r$.

$\Delta^2_{(r)} f(x)$: a principal minor of $D^2 f(x)$ of order r.

Theorem (Characterization of Concave (Convex) Functions via Second Derivatives): Let $f : S \to \mathbb{R}$ be a C^2 function. Then

(1) $f(\cdot)$ is convex in $S \Longleftrightarrow \forall x \in S,\ \forall r = 1, \ldots, n,\ \forall \Delta^2_{(r)} f(x)$, it follows that $\Delta^2_{(r)} f(x) \geq 0 \Longleftrightarrow D^2 f(x)$ is positive semidefinite for any $x \in S$.

(2) $f(\cdot)$ is concave in $S \Longleftrightarrow \forall x \in S,\ \forall r = 1, \ldots, n,\ \forall \Delta^2_{(r)} f(x)$, it follows that $(-1)^r \Delta^2_{(r)} f(x) \geq 0 \Longleftrightarrow D^2 f(x)$ is negative semidefinite for any $x \in S$.

Proof: We skip the proof. $\blacksquare$

Example of Cobb-Douglas Utility Function: Define $u : \mathbb{R}^2_+ \to \mathbb{R}$ be the utility function of the consumer: for any $x \in \mathbb{R}^2_+$, $u(x) = x_1^\alpha x_2^\beta$ where α and β are positive real numbers.

$$\begin{aligned} u_1 &= \alpha x_1^{\alpha-1} x_2^\beta \\ u_{11} &= -\alpha(1-\alpha) x_1^{\alpha-2} x_2^\beta \\ u_{12} = u_{21} &= \alpha\beta x_1^{\alpha-1} x_2^{\beta-1} \\ u_2 &= \beta x_1^\alpha x_2^{\beta-1} \\ u_{22} &= -\beta(1-\beta) x_1^\alpha x_2^{\beta-2}. \end{aligned}$$

Then,

$$D^2u(x) = \begin{pmatrix} -\alpha(1-\alpha) x_1^{\alpha-2} x_2^\beta & \alpha\beta x_1^{\alpha-1} x_2^{\beta-1} \\ \alpha\beta x_1^{\alpha-1} x_2^{\beta-1} & -\beta(1-\beta) x_1^\alpha x_2^{\beta-2} \end{pmatrix}.$$

$$
\begin{aligned}
u_{11}u_{22} - u_{12}u_{21} &= \alpha\beta(1-\alpha)(1-\beta)x_1^{2\alpha-2}x_2^{2\beta-2} - \alpha^2\beta^2x_1^{2\alpha-2}x_2^{2\beta-2} \\
&= \alpha\beta x_1^{2\alpha-2}x_2^{2\beta-2}\{(1-\alpha)(1-\beta) - \alpha\beta\} \\
&= \alpha\beta x_1^{2\alpha-2}x_2^{2\beta-2}(1-\alpha-\beta).
\end{aligned}
$$

Then, $u(\cdot)$ is concave if and only if $0 < \alpha + \beta \leq 1$.

“Strict” Concavity/Convexity via Second Derivatives (Ch. 2.3)

Theorem: Let $f : S \to \mathbb{R}$ be a C^2 function. Then

(1) $D^2 f(x)$ is positive definite for any $x \in S \Longleftrightarrow D^2_{(r)} f(x) > 0$ for all $x \in S$ and all $r = 1, \ldots, n \Rightarrow f(\cdot)$ is strictly convex.

(2) $D^2 f(x)$ is negative definite for any $x \in S \Longleftrightarrow (-1)^r D^2_{(r)} f(x) > 0$ for all $x \in S$ and all $r = 1, \ldots, n \Rightarrow f(\cdot)$ is strictly concave.

Proof: We skip the proof. ■

Example of Cobb-Douglas Utility Function Revisited: Define $u : \mathbb{R}^2_+ \to \mathbb{R}$ as the utility function of a consumer such that for any $x \in \mathbb{R}^2_+$, $u(x) = x_1^\alpha x_2^\beta$ where α and β are positive real numbers.

Then, $u(\cdot)$ is strictly concave if $0 < \alpha + \beta < 1$.

Concavity/Convexity via First Derivatives (Ch. 2.4)

The following result is extremely important in both static and dynamic optimization.

Theorem: Suppose that $f : S \to \mathbb{R}$ is a C^1 function. Then

(1) $f(\cdot)$ is concave in $S \Leftrightarrow$

$$f(x) - f(x^0) \le \nabla f(x^0) \cdot (x - x^0) = \sum_{i=1}^{n} \frac{\partial f(x^0)}{\partial x_i}(x_i - x_i^0)$$

for all $x, x^0 \in S$.

(2) $f(\cdot)$ is strictly concave $\Leftrightarrow$ the above inequality is always strict when $x \ne x^0$.

Remark: Geometrically, this result says that the tangent at any point on the graph will lie above the graph.

Proof: We skip the proof. ■

Quasiconcave and Quasiconvex Functions (Ch. 2.5)

Definition: A function $f : S \to \mathbb{R}$ is **quasiconcave** if the upper level set $P_\alpha = \{x \in S | f(x) \geq \alpha\}$ is convex for each $\alpha \in \mathbb{R}$. We say that f is **quasiconvex** if $-f$ is quasiconcave. So, f is quasiconvex if the lower level set $P^\alpha = \{x \in S | f(x) \leq \alpha\}$ is convex for each $\alpha \in \mathbb{R}$.

There are equivalent definitions of quasiconcavity.

Theorem: $f : S \rightarrow \mathbb{R}$ is quasiconcave if and only if either of the following conditions is satisfied for all $x, x^{'} \in S$ and all $\lambda \in [0, 1]$,

(1) $f(\lambda x + (1 - \lambda)x^{'}) \geq \min\{f(x), f(x^{'})\}$

(2) $f(x^{'}) \geq f(x) \Rightarrow f(\lambda x + (1 - \lambda)x^{'}) \geq f(x)$.

Proof: We omit the proof. ■

Concavity $\Rightarrow$ Quasiconcavity

Proposition: If $f : S \to \mathbb{R}$ is concave, then it is quasiconcave. Similarly, if $f(\cdot)$ is convex, then it is quasiconvex.

Proof: We omit the proof. ■

Quasiconcavity is preserved under positive monotone transformation

A function $F : \mathbb{R} \to \mathbb{R}$ is said to be **strictly increasing** if $F(x) > F(y)$ whenever $x > y$.

Theorem: Let $f : S \to \mathbb{R}$ and let $F : D \to \mathbb{R}$ where $f(S) \subseteq D \subseteq \mathbb{R}$. If $f(\cdot)$ is quasiconcave (quasiconvex) and F is strictly increasing, then $F(f(\cdot))$ is quasiconcave (quasiconvex).

Proof: Suppose $f(\cdot)$ is quasiconcave. Using the previous theorem, we must have

$$f\left(\lambda x + (1-\lambda)x^{'}\right) \geq \min\{f(x), f(x^{'})\}.$$

Since $F(\cdot)$ is strictly increasing,

$$F\left(f\left(\lambda x + (1-\lambda)x^{'}\right)\right) \geq F\left(\min\{f(x), f(x^{'})\}\right) = \min\{F(f(x)), F(f(x^{'}))\}.$$

It follows that $F \circ f$ is quasiconcave. The argument in the quasiconvex case is entirely similar, replacing $\geq$ with $\leq$ and min with max. ■

Definition: A function $f : S \to \mathbb{R}$ is **strictly quasiconcave** if

$$f(\lambda x + (1-\lambda)x^{'}) > \min\{f(x), f(x^{'})\}$$

for all $x, x^{'} \in S$ with $x \neq x^{'}$ and all $\lambda \in (0,1)$.

The function $f(\cdot)$ is **strictly quasiconvex** if $-f(\cdot)$ is strictly quasiconcave.

The Cobb-Douglas Function

Example: Let $f(x_1, \ldots, x_n) = Ax_1^{a_1} \cdots x_n^{a_n}$, where $x_1, \ldots, x_n > 0$, $A > 0$, and $a_1, \ldots, a_n > 0$. Set $a = a_1 + \cdots + a_n$.

- $f(\cdot)$ is quasiconcave for all $a_1, \ldots, a_n$;

- $f(\cdot)$ is concave for $a \leq 1$;

- $f(\cdot)$ is strictly concave for $a < 1$.