Mathematical Methods for Economic Dynamics (ECON 696)

Lecture 4: Multivariate Calculus II (Ch. 2) and Static Optimization (Ch. 3)

“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 $x \in S \iff D^2_{(r)} f(x) > 0$ 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 $x \in S \iff (-1)^r D^2_{(r)} f(x) > 0$ 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 $S \Leftrightarrow$

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

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

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

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: $f : S \to \mathbb{R}$ 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. $\blacksquare$

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 $a < 1$.

Characterization of Quasiconcavity via First Derivatives (Ch. 2.5)

Theorem: Let $f : S \to \mathbb{R}$ be a C^1 function. Then $f(\cdot)$ is quasi-concave on S if and only if for all $x, x^0 \in S$,

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

Proof: We skip the proof. ■

The content of the above theorem is that for any quasiconcave function $f(\cdot)$ and any pair of points x and x^0 with $f(x) \geq f(x^0)$, the gradient vector $\nabla f(x^0)$ and the vector $(x - x^0)$ must form an acute angle.

A Determinant Criterion for Quasiconcavity (Ch. 2.5)

Theorem: Let $S \subseteq \mathbb{R}^2$ be an open, convex set and $f : S \to \mathbb{R}$ be a C^2 function. Define the **bordered Hessian determinant**

$$B_2(x,y) = \begin{vmatrix} 0 & f_1'(x,y) & f_2'(x,y) \\ f_1'(x,y) & f_{11}''(x,y) & f_{12}''(x,y) \\ f_2'(x,y) & f_{21}''(x,y) & f_{22}''(x,y) \end{vmatrix}.$$

1. A necessary condition for f to be **quasiconcave** in S is that $B_2(x,y) \geq 0$ for all $(x,y) \in S$.

2. A sufficient condition for f to be **strictly quasiconcave** in S is that $f_1'(x,y) \neq 0$ and $B_2(x,y) > 0$ for all $(x,y) \in S$.

Proof: We omit the proof. ■

We move on to the general case. Define the bordered Hessian determinants

$$B_r(\mathbf{x}) = \begin{vmatrix} 0 & f_1'(\mathbf{x}) & \cdots & f_r'(\mathbf{x}) \\ f_1'(\mathbf{x}) & f_{11}''(\mathbf{x}) & \cdots & f_{1r}''(\mathbf{x}) \\ \vdots & \vdots & \ddots & \vdots \\ f_r'(\mathbf{x}) & f_{r1}''(\mathbf{x}) & \cdots & f_{rr}''(\mathbf{x}) \end{vmatrix}.$$

for $r = 1, \ldots, n$.

Theorem: Let $S \subseteq \mathbb{R}^n$ be an open, convex set and $f : S \to \mathbb{R}$ be a C^2 function. Then,

1. A necessary condition for f to be **quasiconcave** is that $(-1)^r B_r(\mathbf{x}) \geq 0$ for all $\mathbf{x} \in S$ and all $r = 1, \ldots, n$.

2. A sufficient condition for f to be **strictly quasiconcave** is that $(-1)^r B_r(\mathbf{x}) > 0$ for all $\mathbf{x} \in S$ and all $r = 1, \ldots, n$.

Proof: We omit the proof. ■

Unconstrained Optimization (Ch. 3.1, 3.2)

An **optimization problem** is one where the values of a given function $f : \mathbb{R}^n \to \mathbb{R}$ are to be maximized or minimized over a given set $S \subseteq \mathbb{R}^n$.

This $f(\cdot)$ is called the **objective function** and the set S is called the **constraint set**.

Objectives of Optimization Theory

(1) we identify a set of conditions on $f(\cdot)$ and S under which the **existence** of the solutions to optimization problem is guaranteed. This is already achieved by Weierstrass (the extreme value) theorem.

(2) we obtain a **characterization** of the set of optimal points.

Extreme Points (Ch. 3.1)

Suppose that the point $x^* = (x_1^*, \ldots, x_n^*)$ belongs to S and

$$f(x^*) \geq f(x) \text{ for all } x \in S \ (*)$$

Then, x^* is called a (global) **maximum point** for $f(\cdot)$ in S and $f(x^*)$ is called the **maximum value**.

If the inequality $(*)$ is strict for all $x \neq x^*$, then x^* is a **strict** maximum point for $f(\cdot)$ in S.

We can define **minimum point** and **minimum value** by reversing the inequality sign in $(*)$.

As collective names, we use **extreme points** and **extreme values** to indicate both maxima or minima.

A Necessary Condition for Extreme Points

Theorem: Let $f(\cdot)$ be defined on a set S in $\mathbb{R}^n$ and let $x^* = (x_1^*, \ldots, x_n^*)$ be an interior point in S at which $f(\cdot)$ has partial derivatives. A necessary condition for x^* to be an extreme point for f is that x^* is a **stationary point** for $f(\cdot)$ – that is, it satisfies the equations

$$\nabla f(x) = \mathbf{0} \Longleftrightarrow \frac{\partial f(x)}{\partial x_i} = 0, \text{ for } i = 1, \ldots, n$$

Proof: Suppose, on the contrary, that x^* is a maximum point but not a stationary point for $f(\cdot)$. Then, there is no loss of generality to assume that there exists at least i such that $\partial f(x^*)/\partial x_i > 0$. Define $x^{**} = (x_1^*, \ldots, x_i^* + \varepsilon, \ldots, x_n^*)$. Since x^* is an interior point in S, one can make sure that $x^{**} \in S$ by choosing $\varepsilon > 0$ sufficiently small. Then,

$$f(x^{**}) \approx f(x^*) + \nabla f(x) \cdot (0, \ldots, 0, \underbrace{\varepsilon}_{i}, 0, \ldots, 0) > f(x^*).$$

However, this contradicts the hypothesis that x^* is a maximum point for $f(\cdot)$. ■

Theorem: Suppose that the function $f(\cdot)$ is defined in a convex set $S \subseteq \mathbb{R}^n$ and let x^* be an **interior** point of S. Assume that $f(\cdot)$ is C^1 in a ball around x^*.

1. If $f(\cdot)$ is concave in S, then x^* is a (global) maximum point for $f(\cdot)$ in S if and only if x^* is a stationary point for $f(\cdot)$.

2. $f(\cdot)$ is convex in S, then x^* is a (global) minimum point for $f(\cdot)$ in S if and only if x^* is a stationary point for $f(\cdot)$.

Proof: We focus on the first part of the theorem. The second part follows once we take into account that $-f$ is concave. ($\Rightarrow$) This follows from the previous theorem.

($\Leftarrow$) Suppose that x^* is a stationary point for $f(\cdot)$ and that $f(\cdot)$ is concave. We use the following characterization of concave functions:

Theorem: If $f : S \to \mathbb{R}$ is concave, for any $x, x^{'} \in S$,

$$f(x^{'}) - f(x) \leq \nabla f(x) \cdot (x^{'} - x).$$

Setting $x^{'} = x$ and $x = x^*$,

$$f(x) - f(x^*) \leq \nabla f(x^*) \cdot (x - x^*) = 0 \; (\because \nabla f(x^*) = 0)$$

Thus, we have $f(x) \leq f(x^*)$ for any $x \in S$ as desired. ■

Local Extreme Points (Ch. 3.2)

$x^* \in S$ is a **local maximum point** of $f(\cdot)$ in S if there exists an $\varepsilon > 0$ such that $f(x) \leq f(x^*)$ for all $x \in B_\varepsilon(x^*) \cap S$.

If x^* is the unique local maximum point for $f(\cdot)$, then it is a **strict local maximum point** for $f(\cdot)$ in S.

A **(strict) local minimum point** is defined in the obvious way, and the meaning of the following terms should be clear: **local maximum and minimum values, local extreme points**, and **local extreme values**.

A stationary point x^* of $f(\cdot)$ that is neither a local maximum point nor a local minimum point is called a **saddle point** of $f(\cdot)$.

Theorem (Sufficient Conditions for Local Extreme Points): Suppose that $f(x) = f(x_1, \ldots, x_n)$ is defined on a set $S \subseteq \mathbb{R}^n$ and that $x^* \in S$ is an interior stationary point. Assume also that $f(\cdot)$ is C^2 in an open ball around x^*. Then,

1. $D^2 f(x^*)$ is positive definite $\Rightarrow x^*$ is a local minimum point.

2. $D^2 f(x^*)$ is negative definite $\Rightarrow x^*$ is a local maximum point.

Proof: We skip the proof. ■

The next lemma establishes a sufficient condition for saddle points.

Lemma: If x^* is an interior stationary point of $f(\cdot)$ such that $|D^2 f(x^*)| \neq 0$ and $D^2 f(x^*)$ is neither positive definite nor negative definite, then x^* is a saddle point.

Proof: We skip the proof. ■

Theorem (Necessary Conditions for Local Extreme Points): Suppose that $f(x) = f(x_1, \ldots, x_n)$ is defined on a set $S \subseteq \mathbb{R}^n$, and x^* is an interior stationary point in S. Assume that f is C^2 in a ball around x^*. Then,

1. x^* is a local minimum point $\Rightarrow D^2 f(x^*)$ is positive semidefinite.

2. x^* is a local maximum point $\Rightarrow D^2 f(x^*)$ is negative semidefinite.

Proof: We skip the proof. ■

Constrained Optimization: Equality Constraints (Ch. 3.3, 3.4)

Equality Constraints as a Tangent Hyperplane

If the constraints do bite at an optimum x, we need to have some knowledge of what the constraint set looks like in a neighborhood of x in order to characterize the behavior of the objective function $f(\cdot)$ around x.

A set of equality constraints in $\mathbb{R}^n$, $g(x) = \mathbf{0}$, i.e.,

$$\begin{aligned} g^1(x) &= 0 \\ g^2(x) &= 0 \\ \vdots\quad & \quad\vdots \\ g^m(x) &= 0 \end{aligned}$$

defines a subset of $\mathbb{R}^n$ which is best viewed as a hypersurface.

We write

$$\underbrace{Dg(x)}_{m \times n \text{ matrix}} = \begin{pmatrix} \nabla g^1(x) \\ \nabla g^2(x) \\ \vdots \\ \nabla g^m(x) \end{pmatrix} = \begin{pmatrix} \frac{\partial g^1(x)}{\partial x_1} & \frac{\partial g^1(x)}{\partial x_2} & \cdots & \frac{\partial g^1(x)}{\partial x_n} \\ \frac{\partial g^2(x)}{\partial x_1} & \frac{\partial g^2(x)}{\partial x_2} & \cdots & \frac{\partial g^2(x)}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g^m(x)}{\partial x_1} & \frac{\partial g^m(x)}{\partial x_2} & \cdots & \frac{\partial g^m(x)}{\partial x_n} \end{pmatrix}.$$

Equality Constraints: The Lagrange Problem (Ch. 3.3)

Let $S \subseteq \mathbb{R}^n$. A general maximization problem with equality constraints is of the form

$$\max_{x=(x_1,\ldots,x_n)\in S} f(x_1,\ldots,x_n) \text{ s.t. } g^j(x) = 0 \; \forall j = 1,\ldots,m \; (m < n) \; (*)$$

We say $x \in S$ is **feasible** if $g^j(x) = 0$ for each $j \in \{1,\ldots,m\}$.

Define the **Lagrangian**,

$$\mathcal{L}(x,\lambda) = f(x) - \lambda_1 g^1(x) - \cdots - \lambda_m g^m(x)$$

where $\lambda_1, \ldots, \lambda_m$ are called **Lagrange multipliers**.

The necessary first-order conditions for optimality are then:

$$\nabla\mathcal{L}(x,\lambda) = \nabla f(x) - \sum_{j=1}^{m} \lambda_j \nabla g^j(x) = \mathbf{0}$$

$$\iff \frac{\partial \mathcal{L}(x)}{\partial x_i} = \frac{\partial f(x)}{\partial x_i} - \sum_{j=1}^{m} \lambda_j \frac{\partial g^j(x)}{\partial x_i} = 0,\ \forall i = 1, \ldots, n \ (**)$$

Theorem (Conditions for Extreme Points with Equality Constraints):

(**Necessity**) Suppose that the functions $f(\cdot)$ and $g^1(\cdot), \ldots, g^m(\cdot)$ are defined on a set S in $\mathbb{R}^n$ and $x^* = (x_1^*, \ldots, x_n^*)$ is an interior point of S that solves the maximization problem $(*)$. Assume further that $f(\cdot)$ and $g^1(\cdot), \ldots, g^m(\cdot)$ are C^1 in an open ball around x^*, and that $\text{rank}(Dg(x^*)) = m$. Then, there exist unique numbers $\lambda_1, \ldots, \lambda_m$ such that the first-order conditions $(**)$ are valid.

(**Sufficiency**) Suppose that there exist numbers $\lambda_1, \ldots, \lambda_m$ and a feasible x^* which together satisfy the first-order conditions $(**)$. Then, if the Lagrangian $\mathcal{L}(x, \lambda_1, \ldots, \lambda_m)$ is concave in x, then x^* solves the maximization problem $(*)$.

Proof: (**Necessity**) We content ourselves with a heuristic argument based on the simplest formulation.

Consider

$$\max_{(x,y)\in\mathbb{R}^2} f(x,y) \text{ subject to } g(x,y)=c.$$

Let (x^*,y^*) be a local maximum point of f to the above constrained optimization problem. So, we must have $g(x^*,y^*)=c$.

We consider a pair of "small" numbers $(\Delta x,\Delta y)\in\mathbb{R}^2$ such that $g(x^*+\Delta x,y^*+\Delta y)=g(x^*,y^*)$. Then, we have

$$\begin{aligned}\Delta g &= g(x^*+\Delta x,y^*+\Delta y)-g(x^*,y^*)\\ &\underbrace{\approx}_{\text{linear approx}} g_1'(x^*,y^*)\Delta x+g_2'(x^*,y^*)\Delta y=0.\end{aligned}$$

Assuming that $g_1^{'}(x^*, y^*) \neq 0$ (i.e., rank$(Dg(x^*, y^*)) = m$), we derive

$$\Delta x = -\frac{g_2^{'}(x^*, y^*)}{g_1^{'}(x^*, y^*)} \Delta y. \quad (*)$$

Since (x^*, y^*) is a local maximum point to the constrained optimization problem,

$$\begin{aligned} 0 &\geq f(x^* + \Delta x, y^* + \Delta y) - f(x^*, y^*) \\ &\underbrace{\approx}_{\text{linear approx}} f_1^{'}(x^*, y^*)\Delta x + f_2^{'}(x^*, y^*)\Delta y \\ &= \left(-\frac{f_1^{'}(x^*, y^*)}{g_1^{'}(x^*, y^*)} g_2^{'}(x^*, y^*) + f_2^{'}(x^*, y^*) \right) \Delta y. \quad (\because \ (*)) \end{aligned}$$

Since Δy could be positive or negative, we must have

$$-\frac{f_1'(x^*, y^*)}{g_1'(x^*, y^*)} g_2'(x^*, y^*) + f_2'(x^*, y^*) = 0. \qquad (**)$$

Define

$$\lambda^* \equiv \frac{f_1'(x^*, y^*)}{g_1'(x^*, y^*)}.$$

Then, $(**)$ can be translated into:

$$\begin{aligned} f_1'(x^*, y^*) &= \lambda^* g_1'(x^*, y^*), \\ f_2'(x^*, y^*) &= \lambda^* g_2'(x^*, y^*). \end{aligned}$$

(**Sufficiency**) Suppose that the Lagrangian $\mathcal{L}(x, \lambda)$ is concave in x. The first-order necessary conditions imply that x^* is a stationary point of the Lagrangian. Then, by the sufficiency result for unconstrained maximization with $\mathcal{L}(x, \lambda)$ being the objective function,

$$\mathcal{L}(x^*, \lambda) = f(x^*) - \sum_{j=1}^{m} \lambda_j g^j(x^*) \geq f(x) - \sum_{j=1}^{m} \lambda_j g^j(x) = \mathcal{L}(x, \lambda) \ \forall x \in S$$

But for all feasible x, we have $g^j(x) = 0$ and of course, $g^j(x^*) = 0$ for all $j = 1, \ldots, m$. This implies that $f(x^*) \geq f(x)$ for all feasible x. Thus, x^* solves the maximization problem $(*)$. ■