

ECON20110 (W25): The Elements of Economic Analysis II Honors

Lecturer: Ryan Fang
Notes by: Aden Chen

Wednesday 15th January, 2025

Contents

1	On Mathematics	3
2	Production Technology	9
3	Cost Minimization	12

1 On Mathematics

1.1 Constrained Maximization

E.g.,

$$\max_{\mathbf{x}} U(\mathbf{x}, \theta) \quad \text{s.t.} \quad G(\mathbf{x}, \theta) \geq 0.$$

Solving a whole class of optimization problems parameterized by $\tilde{\theta}$ generates two functions:

- The solution function
- The Value function

Results like the envelope theorem relates these two functions.

1.2 The Kuhn-Tucker Theorem

Consider the maximization function $\max_x f(x)$. The first order condition gives $f'(x^*) = 0$. Now suppose that x_1 is such that $f'(x_1) > 0$. We may be tempted to argue that x_1 is not a solution since we can increase f by increasing the value of x , but this assumes that x is in the interior of the domain. Thus the first order condition considers only interior solutions. The Kuhn-Tucker theorem addresses this issue.

Theorem 1.1 (Kuhn-Tucker). *The FOCs for the constrained optimization problem*

$$\max_{\mathbf{x}} U(\mathbf{x}, \theta) \quad \text{s.t.} \quad G(\mathbf{x}, \theta) \geq 0.$$

are:

- for each i , $\partial \mathcal{L} / \partial x_i \leq 0$ and $x_i \geq 0$, with complementary slackness; That is, at most one of the two conditions can be a strict inequality.
- $\partial \mathcal{L} / \partial \lambda \geq 0$ and $\lambda \geq 0$, with complementary slackness.

Remark 1.2.

- For the direction of the inequalities on $\partial \mathcal{L} / \partial x_i \geq 0$ and $\partial \mathcal{L} / \partial \lambda \geq 0$, remember the picture. In minimization problems they are flipped.
- Often, we can rule out many of these cases. For example, when u is strictly increasing, we have that $\lambda > 0$; and $\lim_{x_i \rightarrow 0} \partial u / \partial x_i = \infty$ gives $x_i > 0$.
- Negative sign in front so that we have a positive parameter.

1.3 Elasticity of Substitution

Elasticities are of the form

$$-\frac{d \log y}{d \log x} = -\frac{dy/y}{dx/x}.$$

- Elasticities gives the proportion response of x as y changes proportionately.
- Knowing the elasticities gives information on how the product xy changes as y changes. For example, if $\sigma > 1$, then xy decreases as y increases.

The elasticity of substitution captures how the (optimal) relative consumption level between two goods responds to changes of the corresponding price ratio:

$$\sigma_{ij} = -\frac{d \log(x_i^*/x_j^*)}{d \log(p_i/p_j)} = \frac{d \log(x_j/x_i)}{d \log(MU_i/MU_j)}.$$

Remark 1.3.

- We think of relative prices as exogenous. The last formula is often used as the definition because it can be computed straight from definition.
- If $\sigma_{ij} > 1$, then relative expenditure $(p_i x_i)/(p_j x_j)$ decreases as p_i/p_j increases, etc.
- Larger values of σ_{ij} means it is “easier to substitute i for j ”.

1.4 Sets and Mapping

Notation 1.4. We write $\mathbb{R}_+ = [0, \infty)$ and $\mathbb{R}_{++} = (0, \infty)$.

Definition 1.5. For two vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ we write:

- $\mathbf{x} \geq \mathbf{y}$ if $\forall i : x_i \geq y_i$;
- $\mathbf{x} > \mathbf{y}$ if $\mathbf{x} \geq \mathbf{y}$ and $\mathbf{x} \neq \mathbf{y}$.
- $\mathbf{x} \gg \mathbf{y}$ (read strongly greater than) if $\forall i : x_i > y_i$;

Definition 1.6. A function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is **strictly increasing** if $f(\mathbf{x}) > f(\mathbf{y})$ for all $\mathbf{x} \gg \mathbf{y}$. It is **strongly increasing** if $f(\mathbf{x}) > f(\mathbf{y})$ for all $\mathbf{x} > \mathbf{y}$.

Example 1.7. Strongly increasing implies strictly increasing.

- Cobb-Douglas is strongly increasing when $\mathbf{x} \gg \mathbf{0}$ but is only strictly increasing when $x_i = 0$ for some i .
- The linear production function $f(\mathbf{x}) = \sum x_i$ is strongly increasing.
- The Leontief production function $f(\mathbf{x}) = \min x_i$ is strictly increasing but not strongly increasing.

Definition 1.8. The $N \times N$ matrix M is **negative semidefinite** (NSD) if

$$\forall z \in \mathbb{R}^N : z \cdot Mz \leq 0$$

and **positive semidefinite** (PSD) if

$$\forall z \in \mathbb{R}^N : z \cdot Mz \geq 0.$$

If the inequality is strict for all $z \neq 0$, then M is **negative definite** (ND) (resp., **positive definite** (PD)).

Proposition 1.9.

- (i) M is PSD (PD) $\iff -M$ is NSD (ND).
- (ii) M is ND (PD) $\iff M$ is NSD (PSD), but the converse is not true.
- (iii) M is ND (PD) $\iff M'$ is ND (PD).
- (iv) M is ND (PD) $\iff M^{-1}$ is ND (PD).

Proof. The first three statements are immediate. For the last, note that

$$z'Mz = (z'Mz)' = z'M'z = z'MM^{-1}M'z = (M'z)'M^{-1}M'z.$$

□

1.5 Concave and Convex Functions

Notation 1.10. Let $\mathbf{x}^1, \mathbf{x}^2 \in X$ and $t \in [0, 1]$. We often denote $\mathbf{x}^t = t\mathbf{x}^1 + (1-t)\mathbf{x}^2$.

Definition 1.11. A function $f : X \rightarrow \mathbb{R}$ is convex (resp., strictly convex) if $\text{epi } f$ is convex (resp., strictly convex). The function f is concave (resp., strictly concave) if $-f$ is convex (resp., strictly convex).

Proposition 1.12. A function $f : X \rightarrow \mathbb{R}$ is convex if and only if for all $x_1, \dots, x_k \in X$ and $\alpha_1, \dots, \alpha_n$ such that $\sum \alpha_i = 1$, we have $f(\sum \alpha_i x_i) \leq \sum \alpha_i f(x_i)$.

We may think of α_i as probability masses. The following result generalizes this to probability densities:

Proposition 1.13 (Jensen's Inequality). If $f : \mathbb{R} \rightarrow \mathbb{R}$ is convex and differentiable, and X is a random variable such that $E[X]$ and $f(E X)$ exist, then $f(E X) \leq E f(X)$.

Proof. From convexity of f we know $f(x) \geq f(y) + f'(y)(x - y)$ for any x and y . Setting $y = E X$ gives

$$f(X) \geq f(E X) + f'(E X)(X - \mu), \quad \forall x.$$

Taking expectation on both sides gives the desired result. □

Proposition 1.14. The C^1 function $f : X \rightarrow \mathbb{R}$ is convex if and only if

$$f(x + t) \geq f(x) + \nabla f(x) \cdot t$$

for all $x \in X$ and $t \in \mathbb{R}^N$ such that $x + t \in X$.

Proof. Suppose f is convex. For any $\alpha \neq 0$,

$$f(\alpha(x + t) + (1 - \alpha)x) \leq \alpha f(x + t) + (1 - \alpha)f(x),$$

giving

$$f(x + \alpha t) - f(x) \leq \alpha(f(x + t) - f(x))$$

and then

$$f(x) + \frac{f(x + \alpha t) - f(x)}{\alpha} \leq f(x + t).$$

Taking $\alpha \rightarrow 0$ gives the desired result.

For the reverse direction, consider arbitrary $x, y \in X$ and $\lambda \in [0, 1]$. Write $z = \lambda x + (1 - \lambda)y$. By assumption we have

$$\begin{aligned} f(x) &\geq f(z) + \nabla f(z)(1 - \lambda)(x - y), \\ f(y) &\geq f(z) + \nabla f(z)\lambda(y - x), \end{aligned}$$

together giving

$$\lambda f(x) + (1 - \lambda)f(y) \geq f(z).$$

□

Proposition 1.15. *The C^2 function $f : X \rightarrow \mathbb{R}$ is convex if and only if $D^2 f(x)$ is PSD for every $x \in X$.*

Proof. Suppose f is convex. Fix some $x \in X$. For any $t \neq 0$, the second-order Taylor expansion is

$$f(x + \alpha t) = f(x) + \nabla f(x) \cdot (\alpha t) + \frac{\alpha^2}{2} t \cdot D^2 f(x + \beta t) t.$$

By Proposition 1.14,

$$\frac{\alpha^2}{2} t \cdot D^2 f(x + \beta t) t \geq 0.$$

And conversely.

□

1.6 Quasi-Concavity and Quasi-Convexity

Definition 1.16. Let $X \subset \mathbb{R}^n$ be convex. A function $f : X \rightarrow \mathbb{R}$ is **quasi-concave** if for all $\mathbf{x}^1, \mathbf{x}^2 \in X$ and $t \in [0, 1]$, we have $f(\mathbf{x}^t) \geq \min\{f(\mathbf{x}^1), f(\mathbf{x}^2)\}$. The function f is **strictly quasi-concave** if the inequality is strict for all $\mathbf{x}^1 \neq \mathbf{x}^2$.

Remark 1.17.

- Thus (strictly) concave functions are (strictly) quasi-concave. The converse is not true; consider $x \mapsto x^2, x > 0$.
- Quasi-concavity is a ordinal property that is preserved by monotone transformations, by the following result:

Proposition 1.18. *Let $g : \mathbb{R} \rightarrow \mathbb{R}$ be an increasing function. If $f : X \rightarrow \mathbb{R}$ is quasi-concave, so is $g \circ f$. If, in addition, g is strictly increasing, then $g \circ f$ is strictly quasi-concave for all f that is strictly quasi-concave.*

Definition 1.19. For $f : X \rightarrow \mathbb{R}$ and $\mathbf{x}^0 \in X$, the **level set** relative to $f(\mathbf{x}^0)$ is the set $L(\mathbf{x}^0) := \{\mathbf{x} \in X : f(\mathbf{x}) = f(\mathbf{x}^0)\}$; the **superior set** (or the upper contour set) is the set $S(\mathbf{x}^0) := \{\mathbf{x} \in X : f(\mathbf{x}) \geq f(\mathbf{x}^0)\}$; the **inferior set** (or the lower contour set) is the set $I(\mathbf{x}) := \{\mathbf{x} \in X : f(\mathbf{x}) \leq f(\mathbf{x}^0)\}$.

The following results is more or less immediate:

Proposition 1.20. *The function $f : X \rightarrow \mathbb{R}$ is quasi-concave if and only if for all $\mathbf{x}^0 \in X$, $S(\mathbf{x}^0)$ is convex.*

Proposition 1.21. *If $f : X \rightarrow \mathbb{R}$ is (strictly) quasi-concave then $-f$ is (strictly) quasi-convex.*

Just like convexity, quasi-convexity of a function can be related to its Hessian, using the following results:

Lemma 1.22. *The C^1 function $f : X \rightarrow \mathbb{R}$ is quasi-convex if and only if for each $x, y \in X$ such that $f(y) \geq f(x)$ we have*

$$\nabla f(x) \cdot (y - x) \geq 0.$$

Proof. Similar to Proposition 1.14. □

Proposition 1.23. *The C^2 function $f : X \rightarrow \mathbb{R}$ is quasiconvex if and only if for each $x \in X$, the Hessian matrix $D^2f(x)$ is PSD in the subspace $\{x \in \mathbb{R}^N : \nabla f(x) \cdot y = 0\}$.*

Proof. Similar to Proposition 1.15. □

2 Production Technology

“Firm” simply refers to actors on the supply side. They transform resources (inputs) into goods and services (outputs), while constrained by the production technology. The only difference between firms and consumers is the problems they are solving: the former solves a profit maximization problem, the latter a utility maximization problem.

We identify the firm’s input choices with members of $X \subset \mathbb{R}_+^m$ and output choices with members of $Y \subset \mathbb{R}_+^n$. We can describe a firm’s technology by specifying its **production possibility set**, $F \subset X \times Y$, each member of which is called a production plan. In this course we assume $n = 1$. The upper contour of the production possibility set is called the production possibility frontier, which can be described by a production function. Most of the times there is no loss of generality in considering only the production function. Think when.

Definition 2.1. Let $F \subset \mathbb{R}_+^m \times \mathbb{R}_+$ be a production possibility set. The **production function** $f : \mathbb{R}_+^m \rightarrow \mathbb{R}_+$ is defined by

$$f(\mathbf{x}) := \sup\{y \in \mathbb{R}_+ : (\mathbf{x}, y) \in F\}.$$

2.1 Placing Structure on the Production Function

Assumption 2.2. We typically assume that the production function $f : \mathbb{R}_+^m \rightarrow \mathbb{R}_+$ is continuous, strictly increasing, and strictly quasiconcave on \mathbb{R}_+^m and $f(\mathbf{0}) = 0$.

This assumes that the firm’s production optimization (cost minimization) problem is well-defined and has a unique solution.

Definition 2.3. The **marginal product** of input i at input vector \mathbf{x} is

$$MP_i(\mathbf{x}) := \frac{\partial f(\mathbf{x})}{\partial x_i}.$$

The **marginal rate of technical substitution (MRTS)** between inputs i and j is

$$MRTS_{ij}(\mathbf{x}) := \frac{MP_i(\mathbf{x})}{MP_j(\mathbf{x})}.$$

For $y \geq 0$, the y -level **isoquant** of $f : \mathbb{R}_+^m \rightarrow \mathbb{R}_+$ is

$$Q(y) := \{\mathbf{x} \in \mathbb{R}_+^m : f(\mathbf{x}) = y\}.$$

Remark 2.4. Recall the notions of marginal utility, marginal rate of substitution, and indifference curves in consumer theory.

Proposition 2.5. *Under Assumption 2.2 and when $n = 2$,*

- (i) *The slope of isoquant (at a point \mathbf{x}) is given by the MRTS.*
- (ii) *Isoquant are always downward sloping.*
- (iii) *We have diminishing MRTS.*

Proof. (i) Clear.

(ii) From f being strictly increasing, we know $f_i > 0$.

(iii) From f being strictly quasiconcave, isoquants bend towards the origin.

more detail?

□

2.2 Return to Scale

Definition 2.6. We say the production function f exhibits (globally)

- **constant return to scale** if $f(t\mathbf{x}) = tf(\mathbf{x})$ for all $x \in \mathbb{R}_+^n$ and all $t > 0$.
- **increasing return to scale** if $f(t\mathbf{x}) > tf(\mathbf{x})$ for all $x \in \mathbb{R}_+^n$ and all $t > 1$.
- **decreasing return to scale** if $f(t\mathbf{x}) < tf(\mathbf{x})$ for all $x \in \mathbb{R}_+^n$ and all $t > 1$.

Example 2.7. The Cobb-Douglas production function $f(\mathbf{x}) = A \prod x_i^{\alpha_i}$ with $\sum \alpha_i = 1$ is homogeneous of degree one and thus exhibits constant return to scale.

Example 2.8. A firm with increasing return to scale will enjoying decreasing average cost. The reverse is not always true.

Definition 2.9. For a production function $f(\mathbf{x})$, the **elasticity of Substitution between inputs i and j** (at point \mathbf{x}) is defined as

$$\begin{aligned}\sigma_{ij}(\mathbf{x}) &:= -\frac{d \log(x_i/x_j)}{d \log(MP_i/MP_j)} \\ &= \frac{d \log(x_j/x_i)}{d \log(MP_i/MP_j)}.\end{aligned}$$

Remark 2.10. A larger σ_{ij} means it is easier to substitute i for j . To see this, consider the CES production function

$$f(x_1, x_2) = \left(\alpha x_1^{\frac{\sigma-1}{\sigma}} + (1-\alpha) x_2^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}.$$

As $\sigma \rightarrow \infty$, $f(\mathbf{x}) \rightarrow \alpha x_1 + (1-\alpha)x_2$; as $\sigma \rightarrow 0$, $f(\mathbf{x}) \rightarrow \min\{x_1, x_2\}$.

Example 2.11. Cobb-Douglas has constant and unit elasticity of substitution, this directly gives constant expenditure share. In general, if $u(\mathbf{x}) = x_1^\alpha x_2^\beta$ and $\alpha, \beta > 0$, then

$$p_1 x_1 = \frac{\alpha m}{\alpha + \beta}, \quad p_2 x_2 = \frac{\beta m}{\alpha + \beta}$$

and thus

$$x_1^* = \frac{\alpha}{\alpha + \beta} \frac{m}{p_1}, \quad x_2^* = \frac{\beta}{\alpha + \beta} \frac{m}{p_2}.$$

3 Cost Minimization

We assume both the product market and the factor (input) markets are perfectly competitive. The firm thus solves the following profit-maximizing problem:

$$\max_{y, \mathbf{x}} py - \omega \mathbf{x} \quad \text{s.t.} \quad y = f(\mathbf{x}),$$

where ω contains the prices of the inputs.

We may rewrite it as a two part problem:

$$\max_y \max_{\mathbf{x}} py - \omega \mathbf{x} \quad \text{s.t.} \quad y = f(\mathbf{x}),$$

which is equivalent to

$$\max_y py - \max_{\mathbf{x}} \omega \mathbf{x} \quad \text{s.t.} \quad y = f(\mathbf{x}).$$

We consider first the cost minimization problem, that is:

$$\min_{\mathbf{x}} \omega \mathbf{x} \quad \text{s.t.} \quad f(\mathbf{x}) \geq y.$$

The Lagrangian is

$$\mathcal{L} = \omega \mathbf{x} + \lambda(y - f(\mathbf{x})).$$

The Kuhn-Tucker FOCs are

- for $i = 1, \dots, m$: $\partial \mathcal{L} / \partial x_i = \omega_i - \lambda \partial f(\mathbf{x}) / \partial x_i \geq 0$ and $x_i \geq 0$, with C.S.;
- $\partial \mathcal{L} / \partial \lambda = y - f(\mathbf{x}) \leq 0$ and $\lambda \geq 0$, with C.S.

Assuming interior solution and $f(\mathbf{x}) = y$, the FOCs reduce to

$$\omega = \lambda \mathbf{MP}, \quad f(\mathbf{x}) = y$$

We have thus for all i, j that

$$\frac{\omega_i}{\omega_j} = \frac{MP_i}{MP_j},$$

that is, the MRTS between i and j equals their price ratio.

Definition 3.1. The conditional input demand function is

$$\mathbf{x}(\omega, y) \equiv \arg \min_{\mathbf{x} \in f^{-1}(\{y\})} \omega \cdot \mathbf{x},$$

and the **cost function** is the minimized value function.

Remark 3.2.

- The continuity of the production function guarantees the existence of a solution to the cost-minimization problem (when $\omega \gg 0$).
- Strict quasi-concavity of the production function guarantees that the solution to the cost-minimization problem is unique.
- The conditional input demand function is thus well defined.

Appendix A

The Cobb-Douglas production function

$$f(\mathbf{x}) = Ax_1^\alpha x_2^{1-\alpha}$$

has conditional input demand functions

$$x_1 = \frac{y}{A} \left(\frac{\omega_2}{\omega_1} \frac{\alpha}{1-\alpha} \right)^{1-\alpha}, \quad x_2 = \frac{y}{A} \left(\frac{\omega_1}{\omega_2} \frac{1-\alpha}{\alpha} \right)^\alpha$$