Note on Math for Microeconomics

Weikai Chen

March 27, 2019

This is a supplementary note for the discussion session of Econ 701 Spring 2019 at Umass Amherst, covering almost all the topics in the Appendix of MWG but with different order and more details[1]. Solutions to homework are attached at the end, and the updated version can be found here.

[1]

1 Continuous Function on Compact Set

In this section, we will discuss some basic concepts in $\mathbb{R}^n$, the continuity of a function, and the extreme value theorem (EVT) stating that a continuous function defined on a compact set achieves its minimum and maximum on the set [2].

[2] See Section M.F in ?, pp. 943–6

Some basic concepts.

Definition 1.1 (open ball). An *open ball* of radius $\varepsilon > 0$ about a point $x \in \mathbb{R}^n$ is $\{y \in \mathbb{R}^n \mid \|x - y\| < \varepsilon\}$ and is denoted $B_\varepsilon(x)$.

An open ball is also called a (open) neighborhood and denoted by $N_\varepsilon(x)$.

Definition 1.2 (open set). A subset X of $\mathbb{R}^n$ is *open* if for every $x \in X$, there is an $\varepsilon > 0$ such that $B_\varepsilon(x) \subseteq X$.

A set is open if every point in it is surrounded by other points in the set.

Example 1.1. $B_\varepsilon(x) \subseteq \mathbb{R}^n$ is open.

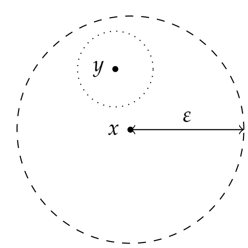

Figure 1: Illustration of Example 1.1

Definition 1.3 (sequence). A sequence in a set X is a function $x : \mathbb{Z}_+^* \to X$ denoted by $\{x_n\}$ or simply x_n.

Definition 1.4 (convergence). A sequence x_n in $\mathbb{R}^n$ converges to x in $\mathbb{R}^n$ if for every $\varepsilon > 0$, there exists an integer N such that $\|x_n - x\| < \varepsilon$ for all $n \geq N$, denoted by

$$\lim_{n\to\infty} x_n = x$$

Definition 1.5 (closed set). A subset X of $\mathbb{R}^n$ is closed if and only if every sequence in X that is convergent in $\mathbb{R}^n$ converges to a point in X.

A set is closed if it is "closed" with respect to the action "taking limit".

The following facts on the sequence x_n are important in showing the closedness.

1. $\forall n, x_n \leq M$ and $\lim_{n\to\infty} x_n = x \Rightarrow x \leq M$.
2. $\forall n, x_n < M$ and $\lim_{n\to\infty} x_n = x \Rightarrow x \leq M$.

Exercise 1.1. Show that the set $\{x \in \mathbb{R}^n \mid Ax \leq b\}$ is closed in $\mathbb{R}^n$.

Continuity.

A function is continuous if and only if the limit of the function values equal to the function value of the limit.

Definition 1.6 (continuous function)**.** The function $f : X \to Y$ is continuous if and only if

$$\lim_{n\to\infty} f(x_n) = f(\lim_{n\to\infty} x_n)$$

whenever x_n is a sequence in X that converges to a point in X.

Definition 1.7 (inverse image)**.** Let $f : X \to Y$, for any subset U of Y, the inverse image of V under f is defined as $f^{-1}(V) = \{x \in X \mid f(x) \in V\}$.

Theorem 1.1. *Let $X \subseteq \mathbb{R}^n$, $Y \subseteq \mathbb{R}^m$, and $f : X \to Y$. Then f is continuous if and only if for every subset V of Y that is closed (open) in Y, $f^{-1}(V)$ is closed (open) in X.*

A function is continuous if and only if the inverse image of any open (closed) set is open (closed).

Proof. We show that if f is continuous and V is closed, then $f^{-1}(V)$ is closed. Let $x_n \in f^{-1}(U)$ and $x_n \to x$ as $n \to \infty$. By continuity, we have

$$\lim_{n\to\infty} f(x_n) = f(x) \in U$$

since U is closed and $f(x_n) \in U$. Therefore $x \in f^{-1}(U)$. □

Example 1.2. Let $u : \mathbb{R}^n_+ \to \mathbb{R}$ is a continuous function, then the upper contour set $\{x \in \mathbb{R}^n_+ \mid u(x) \geq \bar{u}\}$ is a closed set as the set $[\bar{u}, \infty)$ is closed in $\mathbb{R}$. Similarly, $\{x \in \mathbb{R}^n_+ \mid u(x) > \bar{u}\}$ is open.

Below is the argument using the continuity that you really need to master[3].

If f is continuous, $f(x) > u$, then $\exists \varepsilon > 0$ such that $\forall y \in B_\varepsilon(x), f(y) > u$.

[3] It is used in Homework 1.

Compactness.

Definition 1.8 (open cover)**.** A collection $\mathcal{U}$ of open sets is an open cover of X if X is contained in the union of the sets in $\mathcal{U}$, i.e.,

$$X \subseteq \bigcup_{U \in \mathcal{U}} U$$

Definition 1.9 (subsover)**.** If $\mathcal{U}$ is a open cover of X, then a collection of sets $U \in \mathcal{U}$ whose union contains X is called a subcover.

Definition 1.10 (compact set)**.** A set is compact if every open cover of it has a finite subcover.

From this viewpoint, compact sets are 'the next best thing, after singletons and finite sets'.

Exercise 1.2. Give two open covers of $[0, 1)$: one contains a finite sub-cover, the other does not.

Theorem 1.2. *If $X \subset \mathbb{R}^n$ is a compact set and $f : X \to \mathbb{R}^m$ is continuous, then $f(X) = \{f(x) \mid x \in X\}$ is compact.*

The continuous image of a compact set is compact.

Proof. Let $\mathcal{U}$ be an open cover of $f(X)$. By continuity, $f^{-1}(U)$ is open for all $U \in \mathcal{U}$. Therefore, $\{f^{-1}(U) \mid U \in \mathcal{U}\}$ is an open cover of X. By compactness of X, there exists a finite subcover $\{f^{-1}(U_1), \dots, f^{-1}(U_k)\}$. Then $\{U_1, \dots, U_k\}$ is an open cover of $f(X)$ and thus $\mathcal{U}$ has a finite subcover. □

Definition 1.11 (bounded set)**.** A subset X of $\mathbb{R}^n$ is bounded if there is a positive number M such that $\|x\| \leq M$ for all $x \in X$.

Four facts you should know about compact set in $\mathbb{R}^n$.

1. A set $X \subset \mathbb{R}^n$ is compact if and only if it is closed and bounded. [4]
2. A set $X \subset \mathbb{R}^n$ is compact if and only if for any sequence $\{x_n\}$ in X, there exists a convergent subsequence $\{x_{n_k}\}$, $x_{n_k} \to x^* \in X$. [5]
3. For any continuous function f from a compact set X to $\mathbb{R}$, f takes maximum and minimum on X.[6]
4. The simplex $\Delta = \{p \in \mathbb{R}^n \mid \sum_i p_i = 1, p_i \geq 0, \forall i\}$ is compact.

[4] This is called Heine-Borel Theorem

[5] This is true for any metric space and called Bolzano-Weierstrass Theorem

[6] This is called Extreme Value Theorem, discussed below. It can be shown as a corollary of Theorem 1.2.

Extreme Value Theorem. A compact set in $\mathbb{R}^n$ is defined as a bounded and closed set, therefore the EVT could be expressed as

$$\text{closedness} + \text{boundedness} + \text{continuity} \Rightarrow \text{ Max and Min}$$

You could try to provide counterexamples to see why none of them can be omitted on the one hand, and that they are not necessary conditions on the other.

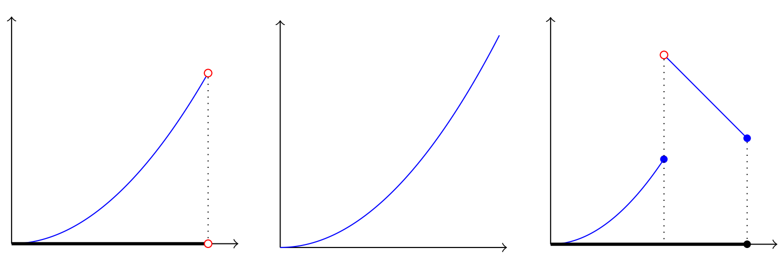

Figure 2: None of the conditions can be omitted when applying EVT.

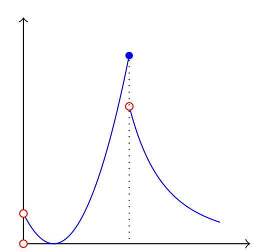

Figure 3: Those conditions in EVT are sufficient but not necessary: The function is not continuous, and the domain is neither bounded nor closed, but it has maximum and minimum.

The EVT is used in microeconomics to show the existence of solutions to the utility maximization/ expenditure minimization/ profit maximization/ cost minimization problems. For example, to show the existence of solution to the following problem

$$\begin{aligned} \max \quad & u^i(x_i) \\ \text{s.t.} \quad & p \cdot x_i \leqq p \cdot \omega^i \end{aligned} \tag{MP^i}$$

where $u^i : \mathbb{R}^n_+ \to \mathbb{R}$ is continuous, and $p = (p_1, \dots, p_n) > 0$, it is sufficient to show that the set $\{x_i \in \mathbb{R}^n_+ \mid p \cdot x_i \leqq p \cdot \omega^i\}$ is closed and bounded. (What if $p \geq 0$?)

2 *Multivariate Calculus*

See Section M.A, M.E, and M.J in ?, pp. 926-7, 940-3, 954-5

THE MAIN IDEA of calculus is to approximate a function by an affine one locally. An affine function is a linear function plus a constant.

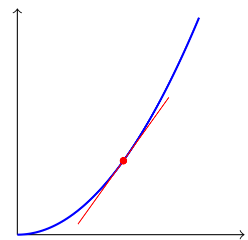

Figure 4: Approximation by an affine function locally.

Definition 2.1 (affine function)**.** A function $f : \mathbb{R}^n \to \mathbb{R}^m$ is *affine* if it is of the form $f(x) = T(x) + b$, where $T : \mathbb{R}^n \to \mathbb{R}^m$ is a linear function and $b \in \mathbb{R}^m$ is a constant.

Definition 2.2 (differentiability)**.** Let U be an open subset of $\mathbb{R}^n$ and let $f : U \to \mathbb{R}^m$. The function f is *differentiable* at $\bar{x} \in U$ if there exists a linear transformation $Df(\bar{x}) : \mathbb{R}^n \to \mathbb{R}^m$, called the derivative of f at $\bar{x}$, such that

$$\lim_{x \to \bar{x}, x \neq \bar{x}} \frac{\|f(x) - f(\bar{x}) - Df(\bar{x})(x - \bar{x})\|}{\|x - \bar{x}\|} = 0$$

Matrix Notation. By the definition, we have f is approximated at $\bar{x}$ by the following affine function $f(\bar{x}) + Df(\bar{x})(x - \bar{x})$. That is for any $x \in B_\varepsilon(\bar{x})$

$$f(x) = \begin{bmatrix} f_1(x) \\ \vdots \\ f_m(x) \end{bmatrix} \approx \begin{bmatrix} f_1(\bar{x}) \\ \vdots \\ f_m(\bar{x}) \end{bmatrix} + \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix} \begin{bmatrix} x_1 - \bar{x}_1 \\ \vdots \\ x_n - \bar{x}_n \end{bmatrix}$$

JACOBIAN MATRIX: The matrix represents the derivative $Df(\bar{x})$ with respect to the standard bases is called the *Jacobian* of f at $\bar{x}$.

When f is a real-valued function, the derivative is sometimes referred to the *gradient* of f at $\bar{x}$ as a column vector, denoted by $\nabla f(\bar{x})$.

Example 2.1. The derivatives in some useful special cases.

- If $f(x) = a^T x$ where $a \in \mathbb{R}^n$ is a constant vector, then $Df(x) = a^T$, or $\nabla f(x) = a$.
- If $f(x) = Ax$ where A is a $m \times n$ matrix, then $Df(x) = A$.
- If $f(x) = x^T A x$ where A is a $n \times n$ matrix, then $Df(x) = x^T A^T + x^T A$, or $\nabla f(x) = Ax + A^T x$. If A is symmetric, then $Df(x) = 2x^T A$, or $\nabla f(x) = 2Ax$.

Inverse and Implicit Function Theorem. We know that for an affine function from $\mathbb{R}^n$ to $\mathbb{R}^n$, $y = f(x) = Ax + b$, if the matrix A is invertible, then it has an inverse function

$$x = f^{-1}(y) = A^{-1}(y - b)$$

For any continuous differentiable function $f : \mathbb{R}^n \to \mathbb{R}^n$, it behaves locally like the derivative, therefore, we have the following theorem.

Theorem 2.1 (Inverse Function Theorem). *Let U be an open subset of $\mathbb{R}^n$ and let $f : U \to \mathbb{R}^n$ be a continuously differentiable function. Suppose $Df(\bar{x})$ is invertible, where $\bar{x} \in U$. Then there exists an open set V such that $\bar{x} \in V$ and a function $g : f(V) \to V$ that is continuously differentiable. Moreover, for any $v \in V$,*

$$Dg(f(v)) = (Df(v))^{-1}.$$

Note that the inverse function theorem is actually a special case of the Implicit Function Theorem. For the function $y = f(x)$, let

$$F(x,y) = f(x) - y.$$

Then from the equation $F(x,y) = 0$ and a solution $(\bar{x},\bar{y})$, can we find an implicit function $x = g(y)$ such that $(g(y),y)$ locally solves the equation? The inverse function theorem tells us that the answer is positive if $D_x F(\bar{x}) = Df(\bar{x})$ is invertible. This idea can be generalized as follows.

Theorem 2.2 (Implicit Function Theorem). *For a continuously differentiable function $F(x,y) : \mathbb{R}^{n+m} \to \mathbb{R}^m$. Consider a solution $(\bar{x},\bar{y})$ to the system of equations*

$$F(x,y) = 0$$

if the Jacobian matrix of F with respect to x evaluated at $(\bar{x},\bar{y})$ is invertible, that is

$$\det D_x F(\bar{x},\bar{y}) = \begin{vmatrix} \frac{\partial f_1(\bar{x},\bar{y})}{\partial x_1} & \cdots & \frac{\partial f_1(\bar{x},\bar{y})}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m(\bar{x},\bar{y})}{\partial x_1} & \cdots & \frac{\partial f_m(\bar{x},\bar{y})}{\partial x_n} \end{vmatrix} \neq 0$$

then the system can be locally solved at $(\bar{x},\bar{y})$ by implicitly defined function $x = g(y)$ such that $\bar{x} = g(\bar{y})$. Moreover,

$$Dg(\bar{y}) = -[D_x F(\bar{x},\bar{y})]^{-1} D_y F(\bar{x},\bar{y})$$

For some change in y, in order to keep $F(x,y) = 0$, x should change to counter the effect of the change in y on F.

Second Derivative. For the real-valued function f, consider the second derivative of f at x, defined as the derivative of the gradient,

$$D^2 f(x) = D\nabla f(x) = \begin{bmatrix} \frac{\partial^2 f(x)}{\partial x_1 \partial x_1} & \cdots & \frac{\partial^2 f(x)}{\partial x_N \partial x_1} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f(x)}{\partial x_1 \partial x_N} & \cdots & \frac{\partial^2 f(x)}{\partial x_N \partial x_N} \end{bmatrix}$$

is called the *Hessian* matrix. If for all n and m, $\frac{\partial^2 f(x)}{\partial x_n \partial x_m}$ exists and is a continuous function of x, then

Hessian Matrix of f at x: The matrix of second partial derivative is the Jacobian matrix of the gradient of f at x.

$$\frac{\partial^2 f(x)}{\partial x_n \partial x_m} = \frac{\partial^2 f(x)}{\partial x_m \partial x_n}$$

then f is said to be *twice continuous differentiable* and therefore the Hessian matrix is symmetric.

Unconstrained Optimization. Next we apply the notations defined above to the problem of unconstrained optimization of a differentiable function.

Definition 2.3. The function $f : \mathbb{R}^n \to \mathbb{R}$ has a *local maximum* at $\bar{x} \in U$ where U is an open subset of $\mathbb{R}^n$, if for $\varepsilon > 0$, $f(\bar{x}) \geq f(x)$, if $\|x - \bar{x}\| < \varepsilon$.

Theorem 2.3. *If $f : U \to \mathbb{R}$ is differentiable at $\bar{x} \in U$, where U is an open subset of $\mathbb{R}^n$ and if f has a local maximum (or minimum) at $\bar{x}$, then*

$$\nabla f(\bar{x}) = 0 \tag{1}$$

Any vector $\bar{x} \in \mathbb{R}^n$ such that $\nabla f(\bar{x}) = 0$ is called a *critical point*. The following theorem provides a criterion to tell whether a critical point is local maximizer (minimizer) or not.

Theorem 2.4 (Second order condition)**.** *If $f : \mathbb{R}^n \to \mathbb{R}$ is twice differentiable and $\nabla f(\bar{x}) = 0$,*

1. *If $\bar{x}$ is a local maximizer, then the Hessian matrix $D^2 f(\bar{x})$ is negative semidefinite. That is, for any $z \in \mathbb{R}^n$,*

$$z^T D^2 f(\bar{x}) z \leq 0$$

2. *If $D^2 f(\bar{x})$ is negative semidefinite, then $\bar{x}$ is a local maximizer.*

Replacing 'negative' by 'positive', the same is true for local minimizers.

Below let's look at an example of its application to the least squares estimation in your Econometrics class[7].

[7] See, for example, in ?, Ch. 3.

Example 2.2. For the model

$$y = \sum_{k=1}^{K} \beta_k x_k + \varepsilon$$

we have N observation of the variable x and of the corresponding y,

$$X = \begin{bmatrix} x_{11} & \cdots & x_{1K} \\ \vdots & \ddots & \vdots \\ x_{N1} & \cdots & x_{NK} \end{bmatrix}, y = \begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix}$$

and would like to choose some estimate of β, say $\hat{\beta}$ to minimize the sum of squared errors. That is

$$\min_{\beta} \quad f(\beta) = (y - X\beta)^T (y - X\beta) = y^T y - 2y^T X\beta + \beta^T X^T X\beta$$

Then by $\nabla f(\hat{\beta}) = 0$ we have

$$-2X^T y + 2X^T X\hat{\beta} = 0 \Longrightarrow \hat{\beta} = (X^T X)^{-1} X^T y.$$

Please check the second order condition. (Hint: $D^2 f(\hat{\beta}) = 2X^T X$ and for any $z \in \mathbb{R}^N$, $z^T (X^T X) z = (Xz) \cdot (Xz) \geq 0$) [8]

[8] Indeed, we have $D^2 f(\beta)$ is positive definite for any β, therefore, it achieves a global minimum at $\hat{\beta}$.

3 Constrained Optimization.

Equality Constraints. Now consider the following problem

$$\begin{aligned} \max_{x\in X} \quad & f(x) \\ \text{s.t.} \quad & g_k(x) = \bar{a}_k, \quad k = 1, \ldots, K \end{aligned} \tag{2}$$

where X is an open subset of $\mathbb{R}^N$, f and each of the g_k are functions from U to $\mathbb{R}$, and $\bar{a}_k$ are constant numbers. Denote the *constraint set* by

$$C = \{x \in X \mid g_k(x) = \bar{a}_k, k = 1, \ldots, K\}$$

The following condition is called the *constraint qualification*: the constraints are independent at $\bar{x}$, i.e., the vectors $\nabla g_1(\bar{x}), \ldots, \nabla g_K(\bar{x})$ are independents. It implies that the matrix $Dg(\bar{x})$ has rank K.

Theorem 3.1. *Assume that f and g_k are continuous differentiable and that f achieves at $\bar{x}$ a local maximum (or minimum) on the constraint set C with constraint qualification. Then there exists $\lambda = (\lambda_1, \ldots, \lambda_K) \in \mathbb{R}^K$ such that*

$$\nabla f(\bar{x}) = \sum_{k=1}^{K} \lambda_k \nabla g_k(\bar{x}) \tag{3}$$

The numbers $\lambda_1, \ldots, \lambda_k$ are called Lagrange multipliers.

This necessary condition can be shown using the implicit function theorem (Theorem 2.2), which is omitted here. The intuition of this theorem is illustrated in Figure 5. Take x' for example, there exists v such that $\nabla g(x') \cdot v = 0$ but $\nabla f(x') \cdot v > 0$. Then for any small $\varepsilon > 0$, we have

$$g(x' + \varepsilon v) \approx g(x') + \varepsilon \nabla g(x') \cdot v = g(x') = \bar{a}$$

while

$$f(x + \varepsilon v) \approx f(x') + \varepsilon \nabla f(x') \cdot v > f(x')$$

so x' can not be the maximizer.

Actually this is also the intuition behind the first order conditions (1) of the unconstrained problem. Indeed,

$$\nabla f(\bar{x}) = 0 \Leftrightarrow \nabla f(\bar{x}) \cdot v = 0, \forall v$$

and

$$\nabla f(\bar{x}) = \sum_{k=1}^{K} \lambda_k \nabla g_k(\bar{x}) \Leftrightarrow \nabla f(\bar{x}) \cdot v = 0, \forall v \in Z(\bar{x})$$

where $Z(\bar{x}) = \{z \in \mathbb{R}^n \mid \nabla g_k(\bar{x}) \cdot z = 0, \forall k\}$.

Construct the following function called the *Lagrangian* whose unconstrained optimization problem is equivalent to the constrained optimization problem (2)

$$\mathcal{L}(x, \lambda, \bar{a}) = f(x) - \lambda \cdot [g(x) - \bar{a}]$$

then we have the following second order conditions:

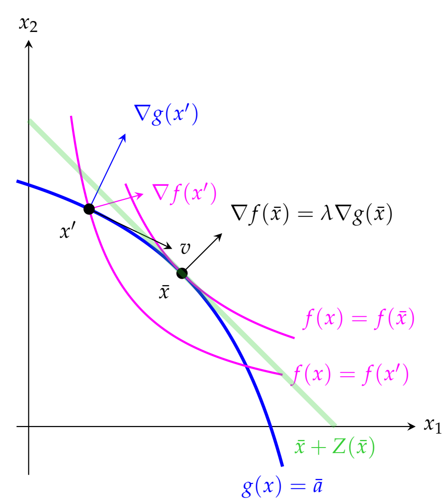

Figure 5: Necessity of the first order condition.

1. If $z \cdot D_x^2 \mathcal{L}(\bar{x}, \lambda, \bar{a}) z < 0$ for all $z \in Z(\bar{x})$ such that $z \neq 0$, then f achieves a local maximum at $\bar{x}$ on C.
2. If $z \cdot D_x^2 \mathcal{L}(\bar{x}, \lambda, \bar{a}) z > 0$ for all $z \in Z(\bar{x})$ such that $z \neq 0$, then f achieves a local minimum at $\bar{x}$ on C.

The Krause-Kuhn-Tucker Theory. Next we extend the idea of Lagrange theory to constrained optimization problems with inequality constraints[9]. Consider the following problem

[9] See Section M.K in ?, pp. 958-63

$$\begin{aligned} \max_{x \in X} \quad & f(x) \\ \text{s.t.} \quad & g_k(x) \leqq \bar{a}_k, \quad \forall k = 1, \ldots, K \end{aligned} \tag{4}$$

Suppose that $\bar{x}$ is a local maximizer of the problem (4), and if $g_k(\bar{x}) = \bar{a}_k$ for all k, then it is equivalent to the problem with equality constraints (2), and we should have the same first order condition (3). If there are some constraints that are not binding, i.e., $g_k(\bar{x}) < \bar{a}_k$, then it is equivalent to the equality constrained problem after deleting both unbinding constraints. Therefore, we have the following KKT conditions.

Theorem 3.2 (KKT conditions). *Suppose that $\bar{x}$ is a local maximizer of problem (4), and assume that the constraint qualification is satisfied, i.e., $\{\nabla g_k(\bar{x}) \mid g_k(\bar{x}) = \bar{a}_k\}$ is linear independent. Then there exists $\lambda_k \in \mathbb{R}_+$ such that*

$$\nabla f(\bar{x}) = \sum_{k=1}^{K} \lambda_k \nabla g_k(\bar{x}) \tag{5}$$

and

$$\lambda_k [g_k(\bar{x}) - \bar{a}_k] = 0, \forall k \tag{6}$$

i.e., $\lambda_k = 0$ if $g_k(\bar{x}) < \bar{a}_k$. $\lambda_k \geq 0$.

Again, we can construct the Lagrangian function $\mathcal{L}$ as before and (5) becomes the first order condition to maximize $\mathcal{L}$.

NOW WE TURN TO the conditions for global optima, i.e., the solutions to the programming problems, which have to do with the shapes of the domains and graphs of the objective and constraint functions.

Convex Sets and Concave Function.

See Section M.C and M.G in ?, pp. 930-3; 946-7

Definition 3.1 (Convex set). A subset X of $\mathbb{R}^n$ is *convex* if for any $x, x' \in X$ and $\alpha \in [0, 1]$, $\alpha x' + (1 - \alpha) x \in X$.

Definition 3.2 (Concave function). If the subset A of $\mathbb{R}^n$ is convex and $f : A \to \mathbb{R}$, then f is *concave* if for every pair of points x and x' in A, and all $\alpha \in [0, 1]$,

The function $f : A \to \mathbb{R}$ is concave if and only if the set $\{(x, y) \mid x \in A, y \leq f(x)\}$ is convex.

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

If the inequality is strict for all $x \neq x'$ and $\alpha \in (0,1)$, then f is strictly concave.

Theorem 3.3. *If A is a convex subset of $\mathbb{R}^n$ and $f : A \to \mathbb{R}$ is concave, then a local maximum of f is a global maximum.*

Proof. Let $\bar{x}$ be a local maximizer so that there exists $B_\varepsilon(\bar{x})$ such that

$$f(\bar{x}) > f(x), \forall x \in B_\varepsilon(\bar{x})$$

Suppose that $\bar{x}$ is not a global maximizer, then there exists $x' \in A$ such that $f(x') > f(\bar{x})$. Take a convex combination of x' and $\bar{x}$, $z = (1-\alpha)\bar{x} + \alpha x'$, then $z \in A$ since A is convex, and $f(z) \geq (1-\alpha)f(\bar{x}) + \alpha f(x') > f(\bar{x})$ for any $\alpha \in (0,1)$ since f is concave. However, for $0 < \alpha < \varepsilon$, we have $z \in B_\varepsilon(\bar{x})$ and $f(z) < f(\bar{x})$, contradicted. □

Theorem 3.4. *If A is a convex subset of $\mathbb{R}^n$ and $f : A \to \mathbb{R}$ is strictly concave, then f achieves a global maximum at at most one point.*

Proof. Suppose that f achieves a global maximum M at both x and x' where $x \neq x'$. Then for $\alpha \in (0,1)$, let $\bar{x} = \alpha x + (1-\alpha)x' \in A$ since A is convex, and $f(\bar{x}) > \alpha f(x) + (1-\alpha)f(x') = M$ since f is strictly concave. Then M is not a global maximum, contradicted. □

Note that a concave function may not be differentiable, and for differentiable concave functions, we have the following facts.

1. If A is a convex subset of $\mathbb{R}^n$ and $f : A \to \mathbb{R}$ is concave and differentiable, then f achieves a global maximum at $\bar{x}$ if and only if $\nabla f(\bar{x}) = 0$
2. If A is a convex and open subset of $\mathbb{R}^n$ and $f : A \to \mathbb{R}$ is everywhere twice differentiable. Then
   (a) f is *concave* if and only if $D^2 f(x)$ is negative semi-definite for all x.
   (b) f is *strictly concave* if and only if $D^2 f(x)$ is negative definite for all x

Corresponding concepts and assertions apply to minima and convex function, whose statements are left as exercises for the reader.

Convex Optimization. Now we revisit the constrained optimization problem with inequality constraints with additional assumptions on the shapes of the domains and the graphs of the functions.

$$\begin{aligned} \max_{x \in X} \quad & f(x) \\ \text{s.t.} \quad & g_k(x) \leqq \bar{a}_k, \quad \forall k = 1, \ldots, K \end{aligned} \tag{7}$$

where X is a non-empty convex subset of $\mathbb{R}^n$, $f : X \to \mathbb{R}$ is concave, $g : X \to \mathbb{R}^K$ and $g_k : X \to \mathbb{R}$ is convex for all $k = 1, \ldots, K$. Denote the feasible set by $C = \{x \in X \mid g(x) \leqq \bar{a}\}$.

Then the following KKT theorem converts the constrained problem (7) to the unconstrained one.

$$\max_{x \in X} \quad \mathcal{L}(x, \lambda) = f(x) - \lambda \cdot [g(x) - \bar{a}] \tag{8}$$

Theorem 3.5 (KKT Theorem). *Given the convex optimization problem (7),*

1. *Suppose that $\bar{x} \in C$ satisfies the* complementary slackness conditions

$$\lambda \cdot [g(\bar{x}) - \bar{a}] = 0 \tag{9}$$

and solves the problem (8), then it solves (7).

2. *If the* constraint qualification *applies, i.e.,*

$$\exists x \in X \text{ such that } g(x) < \bar{a} \tag{10}$$

constraint and $\bar{x}$ solves (7), then there exists $\lambda \in \mathbb{R}_+^K$ satisfies the complementary slackness constraints (9) such that $\bar{x}$ solves (8).

Note that this theorem holds in general even when the functions are not differentiable. The proof of the first part is straightforward as shown below, while for the second part we need the so-called separating hyperplane theorem which will be discussed in next section.

Proof. Suppose that $\bar{x}$ is not a solution to (7), then there exists $x' \in C$ such that $f(x') > f(\bar{x})$. Then we have

$$\mathcal{L}(x', \lambda) = f(x') - \lambda \cdot [g(x') - a] \geq f(x')$$

while

$$\mathcal{L}(\bar{x}, \lambda) = f(\bar{x}) - \lambda \cdot [g(\bar{x}) - a] = f(\bar{x})$$

therefore $\mathcal{L}(x', \lambda) > \mathcal{L}(\bar{x}, \lambda)$, contradicted. □

Next we look at an application of the KTT Theorem in the proof of the Second Fundamental Welfare Theorem in the pure exchange economy.

Example 3.1 (Second Fundamental Welfare Theorem). Consider a pure exchange economy $\mathcal{E} = \langle \mathcal{N}; (u^i)_{i \in \mathcal{N}}; (\omega^i)_{i \in \mathcal{N}} \rangle$ with private ownership of commodities $(\omega_i)_{i \in \mathcal{N}}$, where each of utility functions is concave and differentiable. Let us take a *Pareto efficient allocation* $x^* \in F(\mathcal{E}) \cap \mathbb{R}_{++}^{mn}$ for this economy $\mathcal{E}$. Then there exists a suitable price vector $p^* > 0$ such that (p^*, x^*) is *competitive equilibrium* for $\mathcal{E}_{x^*} \equiv \langle \mathcal{N}; (u^i)_{i \in \mathcal{N}}; (\omega^{*i})_{i \in \mathcal{N}} \rangle$ where $\omega^{*i} \equiv x_i^* (\forall i \in \mathcal{N})$.

Outline of proof. The idea of proof is to construct a maximization problem such that derive the price from the multiplier.

1. Using the characterization of the Pareto efficient outcome, we can define a social planner's maximization problem (see Homework 1).

$$\begin{aligned} \max_{(x_i)_{i\in\mathcal{N}}\in F(\mathcal{E})} \quad & u^1(x_1) \\ \text{s.t.} \quad & u^i(x_i) \geqq u^i(x_i^*), \forall i \neq 1 \end{aligned} \tag{SP}$$

   which is a convex optimization problem (why?).

2. Construct the Lagrangian function and apply the KTT theorem.

$$\max_{(x_i)_{i\in\mathcal{N}}\in\mathbb{R}_+^{mn}} \quad \mathcal{L} = u^1(x_1) + \sum_{i=2}^{n} \lambda_i[u^i(x_i) - u^i(x_i^*)] - \lambda_m \cdot (\sum_i x_i - \Omega)$$

3. Let $p^* = \lambda_m$, and show that (p^*, x^*) is a competitive equilibrium. Indeed, we need to show that x_i^* is the solution to the utility maximization for agent i given price p^*. This can be proved by showing that for all $i \in \mathcal{N}$

$$u^i(x) > u^i(x_i^*) \Rightarrow p^* \cdot x > p^* \cdot x_i^*$$

□

4 *Separating Hyperplane Theorems*

See Section M.G in ?, pp. 947-9

In this section, we will introduce different versions of the separating hyperplane theorems and apply them to the proof of the 2nd fundamental welfare theorem with production economy.

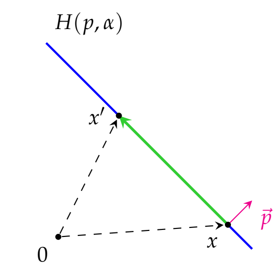

Figure 6: Hyperplane: p is perpendicular to $H(p, \alpha)$.

Definition 4.1 (Hyperplane)**.** Given $p \in \mathbb{R}^n$ with $p \neq 0$ and $\alpha \in \mathbb{R}$, the *hyperplane generated by* p *and* α is the set

$$H(p, \alpha) = \{x \in \mathbb{R}^n \mid p \cdot x = \alpha\}$$

The set $\{x \in \mathbb{R}^n \mid p \cdot x \geq \alpha\}$ and $\{x \in \mathbb{R}^n \mid p \cdot x \leq \alpha\}$ are called, respectively, the *half-space above* and the *half-space below* the hyperplane $H(p, \alpha)$.

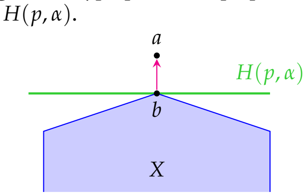

Figure 7: Separating a point and a closed convex set.

Lemma 4.1. *Let X be a nonempty closed convex subset of $\mathbb{R}^n$. Then there exists a unique vector in X of minimum norm. That is, $\exists x_0 \in X$, such that*

$$|x_0| \leq |x|, \forall x \in X$$

Theorem 4.2 (Separating Hyperplane Theorem-I)**.** *Suppose that $X \subset \mathbb{R}^n$ is convex and closed, and $a \notin X$. Then there is $p \in \mathbb{R}^n$ with $p \neq 0$, and a value $\alpha \in \mathbb{R}$ such that $p \cdot a > \alpha$ and $p \cdot x \leqq \alpha$ for every $x \in X$.*

Outline of Proof. First we show that there exists a point $b \in X$ that is closest to a, i.e., $d(a, b) = min_{x\in X} d(a, x)$. Then let $p = b - a$ and $\alpha = p \cdot b$, and show that $H(p, \alpha)$ is separates the point a and the set X. □

Corollary 4.1 (Supporting Hyperplane Theorem)**.** *Let X be a non-empty convex set. If a is a boundary point of X, then there exists a supporting hyperplane that passes through a, i.e., $\exists p \neq 0$ such that $p \cdot a \geqq p \cdot x$ for all $x \in X$.*

SEPARATING HYPERPLANE THEOREM-I: If a is not an interior point of a convex set X, there exists a hyperplane passing through a and including X in a half-space on one side of it.

Separation Two Disjoint Nonempty Convex Sets.

Theorem 4.3 (Separating Hyperplane Theorem-II)**.** *Let X and Y be two nonempty convex subsets in $\mathbb{R}^n$ with $X \cap Y = \emptyset$, then there exists a nonzero vector p and a real number α such that for all $x \in X$, $p \cdot x \leq \alpha$ and for all $y \in Y$, $p \cdot y \geq \alpha$. That is, the hyperplane $H(p, \alpha)$ separates X and Y.*

SEPARATING HYPERPLANE THEOREM-II: There exists a hyperplane separates two disjoint nonempty convex sets.

Next, we apply the separating hyperplane theorem to the proof of the 2nd Fundamental Welfare Theorem with private ownership production economy defined below[10].

Let $\mathcal{N}$ be the set of agents. Let $X_i = \mathbf{R}^m_+$ be individual i's consumption set. Let $\Omega \in \mathbb{R}^m_+$ be a vector of social endowments of commodities. Let $u_i : X_i \to \mathbb{R}$ be individual i's utility function defined on X_i, which is continuous, quasi-concave, and strongly monotonic. Let $Y_j \subseteq \mathbb{R}^m$ be firm j's production possibility set, which is closed, convex. Suppose that a strictly positive production is possible; that is, there are $y_j \in Y_j$ such that $\sum_j y_j + \Omega > 0$. For each firm $j = 1, \ldots, J$, let $(\theta_{ij})_{i \in \mathcal{N}} \in \mathbb{R}^N_+$ be a profile of shares of the profits of firm j, where $\sum_{i \in \mathcal{N}} \theta_{ij} = 1$ and $0 \leq \theta_{ij} \leq 1$. Let $\theta_i \equiv (\theta_{ij})_{j=1,\ldots,J}$ be a private ownership of firms by individual i. For each individual $i \in \mathcal{N}$, let $\omega_i \in \mathbb{R}^m_+$ be individual i's initial endowments of commodities. Note that $\sum_{i \in \mathbb{N}} \omega_i = \Omega$. Then a *private ownership production economy* is an economic environment

$$\mathcal{E}^{PO} = \langle \mathcal{N}; (X_i)_{i \in \mathcal{N}}; (u_i)_{i \in \mathcal{N}}; (\omega_i)_{i \in \mathcal{N}}; (\theta_i)_{i \in \mathcal{N}}; (Y_j)_{j=1,\ldots,J} \rangle.$$

[10] Note that the assumptions in color are different from the standard assumptions in the proof with price quasi-equilibrium.

Example 4.1 (2nd Fundamental Welfare Theorem)**.** Let (x^*, y^*) be a Pareto efficient allocation for $\mathcal{E}^{PO}$. Then there exists a price vector $p^* \in \mathbb{R}^m_{++}$ such that $((x^*, y^*), p^*)$ constitutes a competitive equilibrium with transfers[11].

[11] See Exercise 16.D.3 in ?, p. 574

Outline of Proof. Let $U_i(x_i^*) \equiv \{x_i \in X_i \mid u_i(x_i) > u_i(x_i^*)\}$, $U(x^*) = \sum_{i \in \mathcal{N}} U_i(x_i^*)$ and $Y \equiv \sum_j Y_j$. We would like to show that there exists a hyperplane separates the set $U(x^*)$ and $Y + \{\Omega\}$, and then use the coefficients of the hyperplane to construct the equilibrium price vector.

Step 1: Both $U(x^*)$ and $Y + \{\Omega\}$ are nonempty and convex.

Step 2: $U(x^*) \cap (Y + \{\Omega\}) = \emptyset$ since (x^*, y^*) is Pareto efficient.

Step 3: Apply the Separating Hyperplane Theorem-II, there exists $H(p, r)$ such that $p \cdot z \geq r$ for all $z \in U(x_i^*)$ and $p \cdot z \leq r$ for all $z \in Y + \{\Omega\}$.

Step 4: If $u_i(x_i) \geq u_i(x_i^*)$ for all $i \in \mathcal{N}$, then $p \cdot (\sum_{i\in\mathcal{N}} x_i) \geq r$.

Step 5: $p \cdot (\sum_{i\in\mathcal{N}} x_i^*) = r = p \cdot (\sum_j y_j^* + \Omega)$.

Step 6: Profit Maximization. $\forall j, p \cdot y_j \leq p \cdot y_j^*, \forall y_j \in Y_j$.

Step 7: $\forall i \in \mathcal{N}, u_i(x_i) \geq u_i(x_i^*) \Rightarrow p \cdot x_i \geq p \cdot x_i^*$

Step 8: $p \geq 0$ by free disposal.

Step 9: Let $m_i = p \cdot x_i^*$, then $m_i > 0$ for some i by strictly positive production.

Step 10: If $m_i > 0$ then x_i^* is the utility maximizer. (Cheaper consumption bundle exists since $X_i = \mathbb{R}_+^m$)

Step 11: $p > 0$ by strongly monotone preferences.

Step 12: Utility Maximization.

1) If $x_i^* = 0$, then it is the only feasible bundle.
2) If $x_i^* \neq 0$, then $m_i = p \cdot x_i^* > 0$ and apply Step 10.

□

Appendix: Solutions to Homework

Homework 1. Show the following proposition.

Proposition (Characterization of P-efficient allocation)**.** *Consider a pure exchange economy $\mathcal{E} = \langle \mathcal{N}; (u^i)_{i\in\mathcal{N}}; \Omega \rangle$. A feasible allocation $x^* \in F(\mathcal{E})$ is P-efficient if and only if x^* is a solution to the following program:*

$$\begin{aligned} \max_{x \in F(\mathcal{E})} \quad & u^1(x_1) \\ \text{s.t.} \quad & u^i(x_i) \geq u^i(x_i^*) \quad \forall i \in \mathcal{N}, i \neq 1 \end{aligned} \tag{SP^*}$$

Proof.

(⇒) Let x^* be P-efficient, to show that it is a solution to (SP^*), we assume, on the contrary, that x^* is not a solution to (SP^*). Then, there exists a feasible allocation $x' \in F(\mathcal{E})$ such that

$$\begin{aligned} u^1(x_1') &> u^1(x_1^*) \\ u^i(x_i') &\geq u^i(x_i^*) \quad \forall i \in \mathcal{N}/\{1\} \end{aligned}$$

This implies that x' P-dominate x^* and then x^* cannot be P-efficient. We have a contradiction.

(⇐) Let x^* be the solution to (SP^*), to show that it is P-efficient, we assume, on the contrary, that x^* is not P-efficient. Then, there exists a feasible allocation $x' \in F(\mathcal{E})$ such that

$$\begin{aligned} u^i(x_i') &\geq u^i(x_i^*) \quad \forall i \in \mathcal{N} \\ u^j(x_j') &> u^j(x_j^*) \quad \exists j \in \mathcal{N} \end{aligned}$$

First, consider that

$$\begin{aligned} u^i(x_i') &\geq u^i(x_i^*) \quad \forall i \in \mathcal{N} \\ u^1(x_1') &> u^1(x_1^*) \end{aligned}$$

This is immediately a contradiction from that x^* is a solution to (SP^*). Therefore, we should have

$$\begin{aligned} u^i(x_i') &\geq u^i(x_i^*) \quad \forall i \in \mathcal{N} \\ u^j(x_j') &> u^j(x_j^*) \quad \exists j \in \mathcal{N}/\{1\} \\ u^1(x_1') &= u^1(x_1^*) \end{aligned}$$

Now, take a significantly small positive vector $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_m) > 0$, such that

$$u^j(x_j^\varepsilon) \geq u^j(x_j^*)$$

where $x^\varepsilon \equiv (x_1^* + \varepsilon, x_2^*, \ldots, x_{j-1}^*, x_j^* - \varepsilon, x_{j+1}^*, \ldots, x_n^*)$. Then, $x^\varepsilon \in F(\mathcal{E})$ and by the monotonicity of u^1, we have

$$u^1(x_1^\varepsilon) > u^1(x_1^*)$$

Therefore, x^* cannot be a solution to (SP^*), contradicted.

□

Homework 2. Show the following Theorem.

Theorem (The 1st Fundamental Welfare Theorem). *Consider a private ownership pure exchange economy* $\mathcal{E} = \langle \mathcal{N}; (u^i)_{i\in\mathcal{N}}; (\omega^i)_{i\in\mathcal{N}} \rangle$, *let* (p^*, x^*) *be a competitive equilibrium for* $\mathcal{E}$. *Then, the* x^* *is P-efficient for* $\mathcal{E}$.

We first prove the following two observations about the competitive equilibrium (x^*, p^*):

1. If j prefers some other bundle to the bundle she chooses to buy, that alternative bundle must be not affordable. Otherwise, she would have brought that alternative bundle to maximize her utility. Formally, we have the following statement:
$$u^j(x'_j) > u^j(x^*_j) \Rightarrow p^* \cdot x'_j > p^* \cdot \omega^j$$
2. If i likes some other bundle at least as much as the bundle she chooses to buy, that alternative bundle can not be less valuable than the one she starts with. This depends on the assumption of monotonicity. For if the alternative is actually less valuable than the one she starts with, she could afford a bundle that is slightly more expensive than that alternative by increasing slightly every good, which would give her higher utility than the original one. Formally, we have:
$$u^i(x'_i) \geq u^i(x^*_i) \Rightarrow p^* \cdot x'_i \geq p^* \cdot \omega^i$$

Proof. 1. Suppose, on the contrary, that $p^* \cdot x'_j \leq p^* \cdot \omega^j$, then x'_j is feasible. Since $u^j(x'_j) > u^j(x^*_j)$, x^*_j can not be the solution to MP^j, contradicted.

2. Suppose, on the contrary, that $p^* \cdot x'_i < p^* \cdot \omega^i$, then there exists $\varepsilon > 0$ such that
$$p^* \cdot (x'_i + \varepsilon) < p^* \cdot \omega^i$$
By monotonicity,
$$u^i(x'_i + \varepsilon) > u^i(x'_i) \geq u^i(x^*_i)$$
contradicted.

□

Then we prove the theorem by showing that the total cost of any feasible allocation that Pareto dominates the competitive equilibrium must be greater than the total wealth, which is a contradiction.

Proof of The 1st Fundamental Welfare Theorem. Assume that x^* is not Pareto efficient, then there exists another $x' \in F(\mathcal{E})$ such that
$$\begin{aligned} u^j(x'_j) &> u^j(x^*_j) \quad \exists j \in \mathcal{N} \\ u^i(x'_i) &\geq u^i(x^*_i) \quad \forall i \in \mathcal{N} \end{aligned}$$
By the two observations shown above, we have
$$\begin{aligned} &p^* \cdot x'_j > p^* \cdot \omega^j \\ &p^* \cdot x'_i \geq p^* \cdot \omega^i \quad \forall i \in \mathcal{N} \end{aligned}$$
Then we have
$$p^* \cdot \sum_{i\in\mathcal{N}} x'_i = \sum_{i\in\mathcal{N}} p^* \cdot x'_i > p^* \cdot \sum_{i\in\mathcal{N}} \omega^i$$

However, since $x' \in F(\mathcal{E})$,we have

$$\sum_{i\in\mathcal{N}} x'_i \leq \sum_{i\in\mathcal{N}} \omega^i \Rightarrow p^* \cdot \sum_{i\in\mathcal{N}} x_i \leq p^* \cdot \sum_{i\in\mathcal{N}} \omega^i$$

contradicted. □

Homework 3. Show the following Theorem.

Theorem (The 1st Fundamental Welfare Theorem). *Consider a private ownership economy*

$$\mathcal{E}^{PO} = \langle \mathcal{N}; (X_i)_{i\in\mathcal{N}}; (u_i)_{i\in\mathcal{N}}; (\theta_i)_{i\in\mathcal{N}}; (Y_j)_{j=1,\dots,J} \rangle$$

Let $((x^*, y^*), p^*)$ be a Walrasian competitive equilibrium for $\mathcal{E}^{PO}$. Then, the allocation (x^*, y^*) is Pareto efficient for $\mathcal{E}^{PO}$.

Proof. Suppose, on the contrary, that (x^*, y^*) is not Pareto efficient, then there exists a feasible allocation (x, y) such that $u_i(x_i) \geq u_i(x_i^*)$ for all i and $u_j(x_j) > u_j(x_j^*)$ for some j. Then we have $p^* \cdot x_i \geq p^* \cdot x_i^*$ for all i and $p^* \cdot x_j > p^* \cdot x_j^*$ for some j. Therefore,

$$p^* \cdot \sum_{i\in\mathcal{N}} x_i = \sum_{i\in\mathcal{N}} p^* \cdot x_i > \sum_{i\in\mathcal{N}} p^* \cdot x_i^* = p^* \cdot (\Omega + \sum_j y_j^*)$$

As (x, y) is feasible, we have $\sum_{i\in\mathcal{N}} x_i \leqq \Omega + \sum_j y_j$, and therefore,

$$p^* \cdot \sum_{i\in\mathcal{N}} x_i \leqq p^* \cdot (\Omega + \sum_j y_j)$$

then $p \cdot \sum_j y_j > p^* \cdot \sum_j y_j^*$ contradicted. □

Homework 4. Let there be m consumption goods and n primary factors. For each commodity $j = 1, \dots, m$, there is a production function $f_j : \mathbb{R}_+^n \to \mathbb{R}_+$ such that f_j is (i) increasing; (ii) homogeneous of degree one; (iii) strictly quasi-concave; (iv) differentiable on $\mathbb{R}_+^n$; and it permits free disposal. Let

$$\mathcal{A}_j \equiv \{a_j \in \mathbb{R}_+^n \mid f_j(a_j) \geqq 1\}$$

and $w \in \mathbb{R}_+^n$ be the vector of factor prices. Then define the cost function

$$c_j(w) = \min_{a_j\in\mathcal{A}_j} \quad w \cdot a_j$$

and

$$a_j(w) = \operatorname*{argmin}_{a_j\in\mathcal{A}_j} \quad w \cdot a_j$$

for each commodity j. Show the following properties about the cost function $c_j(w)$.

1) c_j is differentiable with respect to $w > 0$;
2) c_j is homogeneous of degree one;
3) c_j is concave on w;
4) c_j is monotonic increasing.

Proof. 1) The cost minimization program is

$$\begin{aligned} \min_{a_j} \quad & c_j = w \cdot a_j \\ \text{s.t.} \quad & f(a_j) = 1 \end{aligned}$$

The first-order condition is

$$w = \lambda \nabla f(a_j) \text{ or } w_i = \lambda \frac{\partial f}{\partial a_{ij}}, \quad \forall i \tag{11}$$

Take $w^* > 0$ and consider the change of the factor prices from w^* to $(w_i^* + \Delta_i, w_{-i}^*)$. Since $f(a_j(w_i^* + \Delta_i, w_{-i}^*) = 1$ we have

$$\sum_{k=1}^{n} \frac{\partial f}{\partial a_{kj}} \lim_{\Delta_i \to 0} \frac{a_{kj}(w_i^* + \Delta_i, w_{-i}^*) - a_{kj}(w^*)}{\Delta_i} = 0$$

Then

$$\begin{aligned} & \lim_{\Delta_i \to 0} \frac{c_j(w_i^* + \Delta_i, w_{-i}^*) - c_j(w^*)}{\Delta_i} \\ =& a_{ij}(w^*) + \sum_{k=1}^{n} w_k \lim_{\Delta_i \to 0} \frac{a_{kj}(w_i^* + \Delta_i, w_{-i}^*) - a_{kj}(w^*)}{\Delta_i} \\ =& a_{ij}(w^*) + \lambda \sum_{k=1}^{n} \frac{\partial f}{\partial a_{kj}} \lim_{\Delta_i \to 0} \frac{a_{kj}(w_i^* + \Delta_i, w_{-i}^*) - a_{kj}(w^*)}{\Delta_i} \\ =& a_{ij}(w^*) \end{aligned}$$

Hence, $\frac{\partial c_j}{\partial w_i} = a_{ij}$ exists..

2) For any $\alpha > 0$, we have

$$c_j(\alpha w) = \min_{a_j \in \mathcal{A}_j} \quad \alpha w \cdot a_j = \alpha \min_{a_j \in \mathcal{A}_j} \quad w \cdot a_j = \alpha c_j(w)$$

3) For any $w, w' \in \mathbb{R}_+^n$ and $\lambda \in [0,1]$, let $w'' = \lambda w + (1-\lambda) w'$. Then by definition,

$$\begin{aligned} c_j(w'') = w'' \cdot a_j(w'') &= \lambda w \cdot a_j(w'') + (1-\lambda) w' \cdot a_j(w'') \\ &\geq \lambda w \cdot a_j(w) + (1-\lambda) w' \cdot a_j(w') \\ &= \lambda c_j(w) + (1-\lambda) c_j(w') \end{aligned}$$

4) For any $w, w' \in \mathbb{R}_+^n$ such that $w \leqq w'$, we have

$$c_j(w') = w' \cdot a_j(w') \geq w \cdot a_j(w') \geq w \cdot a_j(w) = c_j(w)$$

□