MATH BOOT CAMP 2017

Instructor: Tanisa Tawichsri

1 Introduction

This course will include some of the frequently used mathematical concepts in Economics.

1.1 Outline

- Set thoery

  Naive Set Theory by Halmos

- Logic Concepts: Propositions, Negations, Quantifiers, Necessary and Sufficient Conditions

  Sundaram Appendix A.2-A.4

- Functions: Continuity, Monotonicity, Concavity, Differentiability

  Sundaram 1.4; SB 2, 13, 20.1and 21

- Weierstrass Theorem

  Sundaram 3; SB 30.1

- Open, Closed, Compact, Convex sets

  Sundaram 1.2; SB 12.3-12.5 and 29.2

- Unconstrained Optimization: Necessary and Sufficient Conditions

  Sundaram 4; SB 17

- Constrained Optimization: Lagrange and Kuhn-Tucker Methods, Constraint Qualifications

  Sundaram 5 and 6; SB 18, 19.1, 19.3 and 19.5

- Comparative Statics: Inverse Function, Implicit Function and Envelope Theorems

  Sundaram 1.6.3; SB 19.2

- Dynamic Optimization: special case of standard growth model with closed form solution

  Sundaram 11.1 and 11.2; SB

1.2 Important Examples

Here is the general form of an optimization problem:

$$\min_{x} \quad f(x)$$
$$\text{subject to } x \in D$$

An optimization problem in parametric form:

$$\min_{x} \quad f(x,\theta)$$
$$\text{subject to } x \in D(\theta)$$

1. Utility maximization subject to a budget constraint

$$\max_{x} \quad U(x)$$
$$\text{subject to}$$
$$x \in \mathcal{B}(p, I) \subset \mathbb{R}^k_+$$
$$p \in \mathbb{R}^k_+,\ I \in \mathbb{R}_+$$

2. Profit maximization

$$\max_{\{L,K\}} \quad pF(L,K) - wL - rK$$
$$\text{subject to}$$
$$(L,K) \geq 0$$

3. Cost minimization

$$\min_{L,K} \quad wL + rK$$

subject to

$$F(L,K) \geq \bar{Y}$$
$$(L,K) \geq 0$$

4. Least Squares

$$\min_{\beta} \sum_i (y_i - \beta x_i)^2$$

5. Pareto optimization in an endowment economy

$$\max_{\{c_i\}} \quad \sum_i \lambda_i U(c_i)$$

subject to

$$\sum_i c_i = \bar{X}$$
$$c_i \geq 0 \qquad \forall i, i = 1,..,n$$

6. Dynamic optimization in production economy (consumption and leisure choice)

$$\max_{\{c_t,h_t,x_t\}} \quad \sum_{t=0}^{\infty} \beta^t U(c_t, 1-h_t)$$

subject to

$$c_t + x_t = F(h_t, k_t)$$
$$k_{t+1} = (1-\delta)k_t + x_t$$
$$k_0 \text{ given }, 0 \leq h_t \leq 1, c_t \geq 0, k_t \geq 0 \qquad \forall t$$

2 Logic Concepts

2.1 Propositions: Contrapositives and Converses

Definition: Given two propositions P and Q, and a statement "if P, then Q". The contrapositive of this statement is that "if Q is not true, then P must not be true".

If a statement is true, then its contrapositive is also true. That is, a statement and its contrapositive are logically equivalent. Thus, the following are equivalent

$$\begin{aligned} P \Rightarrow Q, \qquad & x > 0 \Rightarrow x^3 > 0 \\ \sim Q \Rightarrow \sim P, \qquad & x^3 \ngtr 0 \Rightarrow x \ngtr 0 \end{aligned}$$

Sometimes it is easier to show the contrapositive than to show the statement, thus the contrapositive is important.

Definition: The converse of the statement is "if Q then P". However, if the statement is true, the converse need not be true. There is no logic relationship between a statement and its converse. Although the converse holds when the statement is true in the above example , it does not hold in the following example,

$$If \;\; x > 0 \Rightarrow x^2 > 0$$

Definition: In the case that both a statement and its converse hold, we say P holds if and only if Q holds.

$$P \Leftrightarrow Q$$

2.2 Quantifiers and Negation

There are two types of logical quantifiers, the "**existence**" and the "**for all**" quantifiers.

Definition: The "for all" quantifiers denote that a property holds for every element in some set A.

An example of a "for all" quantifier:

"All rich men are happy".

Definition: The "existence" quantifiers denote that the property holds for at least one element in the set A.

An example of an existence quantifier is

"There exists a rich man who is unhappy".

The second statement is a negation of the first one. In fact, to negate an existence quantifier, we use the "for all" quantifier, and viceversa. Think of the statement

"There exists a rich man"

If this were not true, the statement should be

"No man is rich"

which can be stated as

"Every man is poor (=not rich)"

Following this logic, we can construct the **negation** of any statement in mathematics (and logic).

Example 1: Suppose the statement is

$$\text{For all } x \in \mathbb{R}, x^2 > 0$$

The negation of the statement would be

$$\text{There exists at least one } x \in \mathbb{R} \text{ such that}, x^2 \not> 0$$

■

Example 2: If a statement involves more than one quantifier, the problem of negation is more involved. The statement

$$\forall x \in X,\ \exists y \in Y,\ s.t.\ (x, y) \in \Pi(X, Y)$$

The negation is

$$\exists x \in X,\ \forall y \in Y\ s.t.\ (x, y) \notin \Pi(X, Y)$$

And the following is wrong negation,

$$\forall y \in Y,\ \exists x \in X,\ s.t.\ (x, y) \notin \Pi(X, Y)$$

■

From example 2, we see that **the order of the quantifiers is very important. Changing the order changes the statement.** For example, the statement

$$For \text{ every } x > 0,\ \text{there exist } y > 0,\ \text{such that } y^2 = x.$$

if we change the order, then

$$\text{There exist } y > 0,\ \text{for all } x > 0,\ \text{such that } y^2 = x.$$

The first one is true, and the second one is false. Thus, more attention should be paid to the order of quantifiers when forming the negation.

Example 3: Consider the following definition of an open set. The set $X \subset \mathbb{R}$ is open if and only if

$$\forall x \in X, \exists \epsilon > 0 : \forall y \in B(X, \epsilon),\ y \in X$$

The negation of the statement is:

$$\exists x \in X : \forall \epsilon > 0, \exists y \in B(x, \epsilon),\ y \notin X$$

That is if we could find one x such that, no matter the value of ϵ, there would be no way to construct an open ball with radius ϵ contained in the set X. In other words, all open balls around x contain at least one point y which is outside of the set. ■

Example 4: The set $(0, 1)$ is open, because no matter what number the interval we pick, we can always find a number to the right and a number to the left of it that still belong to the interval. On the other hand, if the set is $[0, 1]$, there is no number to the right of 1 that belongs to the set. Thus, the set is not open. ■

Example 5: In $\mathbb{R}^2$, set $\mathcal{D} = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}$ is closed. Take the point $(1, 0) \in \mathcal{D}$, and there is no $\epsilon > 0$ such that $B((1, 0), \epsilon) \subset \mathcal{D}$. ■

2.3 Necessary and Sufficient Conditions

Definition: Suppose the following holds: $P \Rightarrow Q$. By contrapositive, we know that if not Q, then not P, so we say Q is a **necessary** condition for P. However, it might be the case that Q is true, but P is not. In this case, Q is not sufficient.

For example, assume the statement is, when it rains, the floor gets wet. We know that if it rains, then the floor will get wet. But if the floor is wet, it is not necessary that it rains. It could be from other reason. Therefore, the floor gets wet is a necessary condition for rain, but not sufficient.

Definition: If $P \Rightarrow Q$, P is the **sufficient** condition for Q.

If it rains, the floor will get wet. Rain is, therefore, a sufficient condition for the floor to get wet.

Example: The first order conditions are a set of necessary conditions for a local maximum. But these conditions also hold in minima points, so they are not sufficient. The second order conditions together with first order conditions provide sufficient conditions.

The objective function is strictly concave at the point where the first order conditions are met is also a sufficient condition for a local maximum. (When the objective function is strictly concave, its Hessian matrix is negative).

Definition: A condition Q is said to be both necessary and sufficient for P if

$$P \Leftrightarrow Q$$

If the constraint set $\mathcal{D}$ is convex and open, and the objective function is concave and differentiable, the first order conditions are necessary and sufficient condition for a maximum.

Example: If $f(x) \in \mathbb{R}$ and is twice continuously differentiable (C^2), $f''(x) \leq 0$ is the necessary and sufficient condition for f is concave.

2.4 Proofs

Prove the statement: $P \Rightarrow Q$.There are three methods to prove it,

1). A direct proof, assuming P holds, then prove Q also holds.

2). Contrapositive: $\sim Q \Rightarrow \sim P$

3). Contradiction: show that if P and $\sim Q$, then there is a contradiction .

Note $P \Rightarrow Q \equiv \sim P \vee Q$

Then $\sim (P \Rightarrow Q) \equiv P \wedge \sim Q$

3 Sets

3.1 Set Theory

We will take an axiomatic approach to the definition of *set*. Sets have elements or members.

Example

- A pack of wolves is *a set* of wolves
- A line is *a set* of points
- A plane, which is a set of all lines, is *a set of sets* (of points).

If x belongs to A (i.e., x is an element of A or contained in A) then we write

$$x \in A$$

Axiom of Extension: Two sets are equal *if and only if* they have the same elements. If every element of A is an element of B, we say that A is a subset of B, or B includes A

$$A \subset B \textit{ or } B \supset A$$

If $A \subset B$ and $A \neq B$, then A is a *proper subset* of B.

Properties

- *reflexive* $A \subset A$
- *transitive* If $A \subset B$ and $B \subset C$, then $A \subset C$

Axiom of Specification: To every set A and to every condition $S(x)$, there corresponds a set B whose elements are exactly those elements x of A for which $S(x)$ holds

$$B = \{x \in A : S(A)\}$$

The usual logical operators	
belonging	$x \in A$
equality	$A = B$
and	$\wedge$
or (either —or —, or both)	$\vee$
not	$\sim$
if— then — (or implies)	$P \to Q$
if and only if (iff)	$P \leftrightarrow Q$
for some (or there exists)	$\exists x$
for all (or for every)	$\forall x$

Note, immediately from the axiom of extension, and axiom of specification that such B is unique.

Exercise Construct an empty set from a condition $S(x)$ as $x \neq x$, and that there is only one empty set.

Let empty set e $\varnothing$.

Axiom of Pairing: For any two sets, there exists a set that they both belong to.

$$\{x \in A : x = a \textit{ or } x = b\}$$

By axiom of extension, the set a, b is unique. a, b is a pair *(unordered pair).*, while $a, a = a$ is a singleton.

Axiom of Union For every collection of sets, there exists a set that contains all the elements that belong to at least one set of the given collection

$$\{x \in U : x \in X \text{for some } X \text{ in } \mathcal{C}\} \textit{ or } U = \{x : x \in X \text{ for some} X \text{ in } \mathcal{C}\}$$

The general definition implies the union of pairs of sets A, B as

$$A \cup B = \{x : x \in A \textit{ in } x \in B\}$$

Properties

1. $A \cup \varnothing = A$

2. $A \cup B = B \cup A$ *(commutativity)*

3. $A \cup (B \cup C) = (A \cup B) \cup C$ *(associativity)*

4. $A \cup A = A$ *(idempotence)*

5. $A \subset B$ *if and only if* $A \cup B = B$

Proof: *exercise*

Intersection: If A and B are sets, the *intersection* of A and B is the set $A \cap B$ defined by

$$A \cap B = \{x \in A : x = a \textit{ and } x = b\}$$

Exercise: Derive analogous properties of intersection to the properties of union.

Distributive Laws

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C),$$
$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

Proof: *Exercise*

The difference between A and B, also known as *relative complement* is the set $A - B$ defined by

$$A - B = \{x \in A : x \notin B\}$$

The *absolute complement* of A, A', is the relative complement of A with respect to the universe E, which contains all the sets in consideration.

Properties

1. $(A')' = A$

2. $\varnothing' = E$, $E' = \varnothing$

3. $A \cap A' = \varnothing$, $A \cup A' = E$

4. $A \subset B$ if and only if $B' \subset A'$

De Morgan Laws

$$(A \cup B)' = A' \cap B'$$

$$(A \cap B)' = A' \cup B'$$

Proof:

For the second statement, first consider the RHS. x is either in the set A' or B', then x is either not in A, or not in B, i.e. x is not in A and B (or $x \notin A \cap B$). Therefore x is in $(A \cap B)'$. From the LHS. Now, suppose x is in $(A \cap B)'$, then by definition, x is not in $(A \cap B)$. In another word, x is not in either A or B, i.e. x is in either A' or B'.

3.2 Ordered Pairs

Now, if we think about arranging elements of a set A in some order. The *ordered pair* of a and b is the set (a, b) defined by

$$(a, b) = \{\{a\}, \{a, b\}\}.$$

The main properties of the order pair is that if (a, b) and (x, y) are ordered pairs and if $(a, b) = (x, y)$, then $a = x$ and $b = y$.

Define a *Cartesian product* of A and B

$$A \times B = \{x : x = (a, b), \textit{for some a in A and for some b in B}\}.$$

3.3 Relations

Using order pairs (x, y) we can formulate the mathematical theory of relations in set-theoretic language. Take an example of a relation, such as marriage, we can consider ordered pairs (x, y) of which x is a man, y is a woman, and x is married to y.

Relation: a set R is a relation if each element of R is an ordered pair. If $z \in R$, then there exist x and y so that $z = (x, y)$. If $(x, y) \in R$, we can write

$$xRy$$

, saying that x stands in the relation R to y.

Example
Let X be any set, and let R be the set of all those pairs (x, y) in $X \times X$ for which $x = y$. The relation R here is the relation of equality between elements of X; if x and y are in X, then xRy means the same as $x = y$.

We can define the associated sets of ordered pairs as the domain and the range of R (abbreviated $domR$ and $ranR$), defined by

$$\begin{aligned} domR &= \{x : \text{ for some } \ y \ (xRy)\} \\ ranR &= \{y : \text{ for some } \ x \ (xRy)\}. \end{aligned}$$

If R is a relation included in a Cartesian product $X \times Y$ (so that $domR \subset X$ and $ranR \subset Y$), we may say that R is a relation *from* X *to* Y. Instead of a relation from X to X, we may say that R is a relation *in* X.

Definition

- A relation R in X is *reflexive* if xRx for every $x \in X$
- A relation R in X is *symmetric* if xRy implies that yRx.

- A relation R in X is *transitive* if xRy and yRz imply that xRz.

3.4 Functions

If X and Y are sets, a *function* from (or on) X to (or into) Y si a relation f such that $dom\ f = X$ and such that for eachh x in X there is a unique element y in Y with $(x, y) \in f$.

The uniqueness condition can be written formally: if $(x, y) \in f$ and $(x, z) \in f$, then $y = z$. For each x in X, the unique y in Y such that $(x, y) \in f$ is denoted by $f(x)$.

For function, usually we will often see $f(x) = y$ rather than $(x, y) \in f$ or xfy. f is a function from X to Y is often denoted

$$f : X \to Y$$

4 Important set and function concepts

Let $x \in \mathbb{R}^n$. An open ball $B(x, r)$ with center x and radius r is a set

$$B(x, r) = \{y \in \mathbb{R}^n : d(x, y) < r\}$$

where d is some measure of distance, for example with $n = 2$, $d(x, y) = \sqrt{\sum_i (x_i - y_i)^2}$, $i = 1, 2$.

4.1 Open and closed sets

Open sets: A set $\mathcal{D} \subset \mathbb{R}^n$ is open if for all $x \in \mathcal{D}$, there exist $r > 0$ such that $B(x, r) \subset \mathcal{D}$.

Example 1: In $\mathbb{R}$, the set $(0, 1)$ is open. Take any $x \in (0, 1)$, let $r = \frac{\min\{x, 1-x\}}{2}$ and $d(x, y) = |x - y|$.

If $x < 1 - x$, then

$$B(x, r) = \left(\frac{1}{2}x, \frac{3}{2}x\right)$$

Note that $x > 0$, then $\frac{x}{2} > 0$, and since $1 - x > x$, $\frac{3}{2}x < 2x < 1$. Thus, $B(x, r) \subset (0, 1)$.

If, $x > 1 - x$, then

$$B(x,r) = \left(\frac{3}{2}x - \frac{1}{2}, \frac{1}{2}x + \frac{1}{2}\right)$$

Notice that $x > 1 - x \Rightarrow x > \frac{1}{2} \Rightarrow \frac{3}{2}x - \frac{1}{2} > 0$, and $x < 1 \Rightarrow \frac{3}{2}x - \frac{1}{2} < \frac{1}{2}x + \frac{1}{2} < 1$.

Finally, if $x = 1 - x \Rightarrow x = \frac{1}{2}$, so take

$$B(\frac{1}{2}, r) = \left(\frac{1}{4}, \frac{3}{4}\right)$$

■

Example 2 In $\mathbb{R}^2$, $\mathcal{D} = \{(x,y) \in \mathbb{R}^2 : 0 < x < 1, 0 < y < 1\}$ is open. Take any point $(x,y) \in \mathcal{D}$. WLOG, let $x \geq y$. Let $r = \frac{1}{2}\min\{1-x, y\}$.

If $1 - x > y$, then $r = \frac{y}{2}$ and $B((x,y),r) = \{(a,b) \in \mathbb{R}^2 : x - \frac{y}{2} < a < x + \frac{y}{2}, y - \frac{y}{2} < b < y + \frac{y}{2}\}$.

$$x + \frac{y}{2} \leq x + \frac{1-x}{2} < x + 1 - x \leq 1$$

$$x - \frac{y}{2} \geq x - y \geq 0$$

$$y + \frac{y}{2} \leq y + \frac{1-x}{2} < y + \frac{1-y}{2} < y + 1 - y = 1$$

$$y - \frac{y}{2} \geq y - y \geq 0$$

Thus, $B((x,y),r) \subset \mathcal{D}$. Similarly, you can show $B((x,y),r) \subset \mathcal{D}$ if $1 - x < y$ and $1 - x = y$. Therefore, $\mathcal{D}$ is open. ■

Closed set: A set $\mathcal{D} \subset \mathbb{R}^n$ is closed when its complement is open.

The complement of $\mathcal{D}$ is $\mathcal{D}^c = \{x \in \mathbb{R}^n : x \notin \mathcal{D}\}$. **Example** The set $[0,1] \subset \mathbb{R}$ is closed, since its complement is $(-\infty, 0) \cup (1, \infty)$, an open set.

Intuitively, a set is closed when the set has a well defined point after which there are no more elements in the set. A set is open when, no matter which point of the set you look at, there is always a point to the right and to the left of it.

4.2 Bounded Sets and Compact Sets

Bounded set: A set $\mathcal{D} \subset \mathbb{R}^n$ is bounded if there exists $r > 0$ such that $\mathcal{D} \subset B(0, r)$. That is, $\mathcal{D}$ is bounded if there exists an open ball that can completely contain $\mathcal{D}$.

For example, the set $(0, 1) \subset \mathbb{R}$ is bounded, but the set of integers $\{1, 2, 3, ...\}$ is not.

Compact set: A set $\mathcal{D} \subset \mathbb{R}^n$ is compact if it is both closed and bounded.

For example, the set $(0, 1)$ is bounded, but not closed, and therefore not compact. The set $[0, 1]$ is compact.

4.3 Convex Sets

Take an arbitrary collection of points $\{x_1, \ldots, x_m\} \in \mathbb{R}^n$. A point $z \in \mathbb{R}^n$ is a *convex combination* of the points $\{x_1, \ldots, x_m\}$ if $z = \sum_i^m \lambda_i x_i$, with $\lambda_i \geq 0$ for all $i = 1, 2, \ldots, m$ and $\sum_i^m \lambda_i = 1$.

Convex set: A set is convex if any convex combination of any two points in the set is also in the set. That is, take any two points $(x, y) \in \mathcal{D}$, for all $\alpha \in [0, 1]$, $\alpha x + (1 - \alpha)y \in \mathcal{D}$. Intuitively, the set is convex if a straight line joining any two points in the set is inside the set.

Example 1: In $\mathbb{R}$, The sets $(0, 1)$ and $[0, 1]$ are both examples of convex sets. The set $[0, 1] \cup [2, 3]$ is not a convex set. ■

Example 2: In $\mathbb{R}^2$, the open ball with radius 1 is a convex set.

$$\mathcal{D} = \left\{(x, y) \in \mathbb{R}^2 : x^2 + y^2 < 1\right\}$$

The set

$$\mathcal{D} = \left\{(x, y) \in \mathbb{R}^2 : y \leq x^2\right\}$$

is not convex. ■

Example 3: Prove a set is a convex set by definition.

Consider the set $\mathcal{D} = \{(x, y) \in \mathbb{R}^2 : 0 < x < 1, 0 < y < 1\}$. Take any two points $(x_1, y_1), (x_2, y_2) \in \mathcal{D}$, and any $\lambda_1 \in \mathbb{R}_+, \lambda_2 \in \mathbb{R}_+$ such that $\lambda_1 + \lambda_2 = 1$. The linear combination of the two points is $\lambda_1(x_1, y_1) + \lambda_2(x_2, y_2)$,

$$\lambda_1(x_1, y_1) + \lambda_2(x_2, y_2) < \lambda_1(1, 1) + \lambda_2(1, 1) = (\lambda_1 + \lambda_2, \lambda_1 + \lambda_2) = (1, 1)$$
$$\lambda_1(x_1, y_1) + \lambda_2(x_2, y_2) > \lambda_1(0, 0) + \lambda_2(0, 0) = (0, 0)$$

Thus, the point $\lambda_1(x_1, y_1) + \lambda_2(x_2, y_2) \in \mathcal{D}$. ■

5 Functions

Let S, T be subsets of $\mathbb{R}^n$ and $\mathbb{R}^l$. A $function$ f is a mapping from S to T, denoted as $f : S \to T$. The set S is the domain of f, and T is its range.

5.1 Continuous Functions

Continuous Function: A function $f : \mathcal{D} \to \mathbb{R}$ is continuous at $x \in \mathcal{D}$ if, for all sequences $\{x_k\} \in \mathcal{D}$ and $x_k \to x$, $f(x_k) \to f(x)$. f is continuous if it is continuous at every $x \in \mathcal{D}$.

Intuitively, a continuous function is a function that does not have "breaks". That is, a function is continuous function you could trace it entirely without ever lifting your pencil.

The function $f(x) = x$ is continuous, but the function

$$f(x) = \begin{cases} 0, & x \leq 0 \\ 1, & x > 0 \end{cases}$$

is not continuous at $x = 0$.

Theorem: If $f(x)$ and $g(x) : \mathcal{D} \to \mathbb{R}^n$, are continuous functions at $x \in \mathcal{D}$, then $f(x) + g(x)$, $f(x) - g(x)$, $f(x) * g(x)$ are continuous at $x \in \mathcal{D}$.

5.2 Differentiability

Differentiability: A function $f : \mathcal{D} \to \mathbb{R}$ is differentiable at a point $x \in \mathcal{D}$ if

$$\lim_{y \to x} \frac{f(y) - f(x)}{y - x} = A$$

A is called the derivative of f at x. Intuitively, the derivative is the slope of f between points x and y when y goes very close to x.

In general, if f is differentiable at x, there cannot be a break, kink, or cusp at x

Let $f(x) = ax^2, x \in \mathbb{R}$. We denote the derivative of f with respect to x as $f'(x) = 2ax$ or $\frac{df(x)}{dx} = 2ax$, and the second derivative is $f''(x) = \frac{d^2 f(x)}{dx^2} = 2a$.

Example 1: $f(x) = |x|$ is not differentiable at $x = 0$. ■

Let $f : \mathbb{R}^2 \to \mathbb{R}$. The matrix of first derivatives is called the *Jacobian* and the matrix of second derivatives is called the *Hessian*.

$$Jf(x,y) = Df(x,y) = \begin{pmatrix} \dfrac{\partial f(x,y)}{\partial x} & \dfrac{\partial f(x,y)}{\partial y} \end{pmatrix}$$

$$Hf(x,y) = D^2 f(x,y) = \begin{pmatrix} \dfrac{\partial^2 f(x,y)}{\partial x^2} & \dfrac{\partial^2 f(x,y)}{\partial x \partial y} \\ \dfrac{\partial^2 f(x,y)}{\partial y \partial x} & \dfrac{\partial^2 f(x,y)}{\partial y^2} \end{pmatrix}$$

The Hessian is a symmetric matrix, by Young's theorem:

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}$$

The chain rule: $D(f(g(x))) = Df(g) * Dg(x)$.

Notice: If a function is differentiable at x, then it must also be continuous at x. However, the converse is not true.

Continuously Differentiable: A function is continuously differentiable, denoted $f \in \mathbb{C}^1$, if the derivative $f'(x)$ exists, and is itself a continuous function.

5.3 Monotonous Functions

Monotonous Functions:

A function $f : \mathcal{D} \to \mathbb{R}$ is *monotonically increasing* in $x \in \mathcal{D}$ if, for any pair $x_1, x_2 \in \mathcal{D}$,

$$\begin{aligned} x_1 &\leq x_2 \Rightarrow f(x_1) \leq f(x_2). \\ x_1 &\ll x_2 \Rightarrow f(x_1) < f(x_2). \end{aligned}$$

The function is *monotonically decreasing* if

$$\begin{aligned} x_1 &\leq x_2 \Rightarrow f(x_1) \geq f(x_2). \\ x_1 &\ll x_2 \Rightarrow f(x_1) > f(x_2). \end{aligned}$$

Strongly Monotonous Functions:

A function $f : \mathcal{D} \to \mathbb{R}$ is *strongly monotone increasing* in $x \in \mathcal{D}$ if, for any pair $x_1, x_2 \in \mathcal{D}$,

$$x_1 \leq x_2 \;\&\; x_1 \neq x_2 \Rightarrow f(x_1) < f(x_2).$$

The function is *strongly monotone decreasing* if

$$x_1 \leq x_2 \;\&\; x_1 \neq x_2 \Rightarrow f(x_1) > f(x_2).$$

If the function is continuously differentiable (C^1), the following hold:

1. If f is monotonically increasing, then $f' \geq 0$.
2. If f is monotonically decreasing, then $f' \leq 0$.

5.4 Concave and Convex Functions

Definition: Let $\mathcal{D} \subset \mathbb{R}^n$. A function $f : \mathcal{D} \to \mathbb{R}$ is **concave** if, for all $x, y \in \mathcal{D}$ and all $\alpha \in [0, 1]$,

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

If the inequalities are strict, the function is strictly concave.

f is a **convex** function, for all $x, y \in \mathcal{D}$ and all $\alpha \in [0, 1]$,

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

If the inequalities are strict, the function is strictly convex.

In the case $f(x) \in \mathbb{R}$ and twice continuously differentiable (C^2), the following hold:

1. If f is convex, then $f'' \geq 0$
2. If f is strictly convex, then $f'' > 0$
3. If f is concave, then $f'' \leq 0$
4. If f is strictly concave, then $f'' < 0$

5.5 Quasi-concave and Quasi-Convex Functions

Quasi-concave Functions: A function $f : \mathcal{D} \to \mathbb{R}$ is quasi-concave on $\mathcal{D}$ if and only if for all $x, y \in \mathcal{D}$ and for all $\lambda \in (0, 1)$,it is the case that

$$f[\lambda x + (1 - \lambda)y] \geq \min\{f(x), f(y)\}.$$

Quasi-convex Functions: A function $f : \mathcal{D} \to \mathbb{R}$ is quasi-convex on $\mathcal{D}$ if and only if for all $x, y \in \mathcal{D}$ and for all $\lambda \in (0,1)$,it is the case that

$$f[\lambda x + (1-\lambda)y] \leq \max\{f(x), f(y)\}.$$

5.6 Inverse Functions

Definition: If function $f : \mathcal{D} \to \mathbb{U}$ is one-to-one and onto, there is a unique function $g : U \to \mathcal{D}$ such that $f(g(b)) = b$ for all $b \in U$. Function g is called the inverse function of f.

6 Vector Spaces

6.1 Vector Space

Let $\mathbf{u}, \mathbf{v}, \mathbf{w}, \mathbf{z}$ be vectors in set V, and a, b are scalars. The operator "addition" is defined on V,i.e., for any $\mathbf{u}, \mathbf{v} \in V$, $\mathbf{u} + \mathbf{v} \in V$; the operator "multiplication" is defined as $a * \mathbf{v} \in V$ if $\mathbf{v} \in V$. V is a vector space if the following properties are satisfied:

1. $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$,
2. $\mathbf{u} + (\mathbf{v} + \mathbf{z}) = (\mathbf{u} + \mathbf{v}) + \mathbf{z}$
3. There is an element 0 in V such that, for all $\mathbf{v} \in V$, $\mathbf{v} + 0 = \mathbf{v}$
4. For every $\mathbf{v} \in V$,there is an element $\mathbf{w} \in V$, such that $\mathbf{v} + w = 0$
5. $a * (\mathbf{u} + \mathbf{v}) = a * \mathbf{v} + a * \mathbf{u}$
6. $(a + b) * \mathbf{u} = a * \mathbf{u} + b * \mathbf{u}$

7. $(ab) * \mathbf{v} = a * (b\mathbf{v})$

8. $1 * \mathbf{u} = \mathbf{u}$

6.2 Normed Vector Space

Vectors are measured by specifying *norm*, measuring the lengths of the vectors, and inner product , measuring angles between vectors.

$\|.\|$ is a norm if it satisfies

1. Only the zero vector has zero length. Every other vector has a positive length

$$\|x\| > 0 \text{ if } x \neq 0$$

2. Multiplying a vector by a scalar changes only its length, not its direction

$$\|\alpha x\| = |\alpha| \|x\| \text{ for any scalar } \alpha$$

3. The *triangle inequality* holds:

$$\|x + y\| \leq \|x\| + \|y\| \text{ for any vectors } x \text{ and } y.$$

Definition
A *normed vector space* is a pair $(V, \|.\|)$ where V is a vector space and $\|.\|$ a norm on V.

6.2.1 Banach Space

Banach space or a complete norm vector space. Informally, it is a normed vector space that is complete, in a sense that a caught sequence of vectors always converges to a well defined limit in the space.

Definition

A sequence $x_1, x_2, x_3, ...$ of real numbers is called a *Cauchy sequence*, if for every $\varepsilon > 0$, there exist a positive integer N such that for all natural numbers $m, n > N$,

$$\|x_m - x_n\| < \varepsilon$$

.

Definition

A Banach space is a vector space X over the filed R of real numbers, or over the field C of complex numbers, which is equipped with a *norm* and which is *complete* with respect to that norm. That is to say, for every Cauchy sequence $\{x_n\}_{n=1}^{\infty}$ in X, there exists an element x in X such that

$$\lim_{n\to\infty} x_n = x, \text{ i.e., } \lim_{n\to\infty} \|x_n - x\| = 0.$$

7 Optimization

7.1 The Basic Problem

Let $f : \mathcal{D} \to \mathbb{R}$ be a continuous function and $\mathcal{D} \subset \mathbb{R}^n$. The maximization problem is

$$\max_{x} \ f(x) \text{ subject to } x \in \mathcal{D}$$

The solution is a point $x^* \in \mathcal{D}$ such that

$$f(x^*) \geq f(y) \qquad \forall y \in \mathcal{D}$$

If such a point exists, we call it a global maximum. The Weierstrass theorem is used to guarantee the existence of a globe solution.

7.2 Weierstrass Theorem

The Weierstrass theorem provides sufficient conditions to guarantee the existence of optima. It says that if an objective function is continuous over a compact set, there exists at least one maximum and one minimum. A set is compact when it is closed and bounded in $\mathbb{R}^n$. It is important to highlight that these are **sufficient** conditions. That is, if these conditions can not be satisfied, there can still be optima. The statement of the theorem is as follows:

Weierstrass Theorem: *Let $\mathcal{D} \subset \mathbb{R}^n$ be compact, and let $f : \mathcal{D} \to \mathbb{R}$ be a continuous function on $\mathcal{D}$. Then f attains a maximum and a minimum on $\mathcal{D}$, i.e., there exist points z_1 and z_2 such that*

$$f(z_1) \geq f(x) \geq f(z_2) \qquad \forall x \in \mathcal{D}$$

The compactness of $\mathcal{D}$ tells us that there are well defined minimum and maximum points in $\mathcal{D}$.

The continuity of f guarantees that the function is well defined and does not go to infinity as long as we apply it on a bounded set. Let's look at some examples.

Is there a maximum in the following cases? Why or why not?

1. $f(x) = x$, $x \in [0, 1]$.
2. $f(x) = x$, $x \in (0, 1)$.
3. $f(x) = x$, $x \in \mathbb{R}$.
4. $f(x) = \begin{cases} \frac{1}{x}, x \in (0, 1] \\ 0, x = 0 \end{cases}$.
5. $f(x) = \begin{cases} 1, x \in (0, 1] \\ 0, x = 0 \end{cases}$.

The conditions in Weierstrass Theorem is sufficient not necessary. There are optimization problems that do not satisfy Weierstrass Theorem but have optimal solutions.

The theorem is very useful in applications, to check whether a given problem has a solution.

Example 1: Consider the problem of maximizing utility subject to a budget constraint.

$$\begin{aligned} &\max_x u(x) \\ s.t.\ x &\in \mathcal{B}(p, I) = \{x \in \mathbb{R}^n_+ : p \cdot x \leq I\} \end{aligned}$$

The budget set is compact as long as the vector $p \gg 0$. To see this, take the element i of vector x. Notice that $0 \leq x_i \leq \frac{I}{p_i}$. If $u(x)$ is continuous, then we know there is a solution.

Example 2:Sundaram, Exercise 13, chapter 3, page 98. The solution is

1. Suppose there is an $x^* > 0$ and $p(x^*) = 0$. This means that the monopolist will always choose $x \in [0, x^*]$. If he sets $x = x^*$, the revenue is zero, cost is nonnegative, so the profit

is negative. Since the demand curve is downward sloping, the revenue for any $x > x^*$ is 0 with some positive cost. Thus, he has negative profit when choosing $x \geq x^*$. We can compactify his choice set by setting $0 \leq x \leq x^*$ and apply Weierstrass.

2. Suppose there is some $x' > 0 : c(x) \geq xp(x), \forall x \geq x'$. Again, the monopolist will always choose some $x \in [0, x']$, which is compact.

3. Assume $p(x) = \bar{p}$ and $c(x) \to \infty$ as $x \to \infty$. In this case, we cannot use Weierstrass Theorem. Even more, there might not have a solution. To see why, assume $c(x) = \ln(x)$. The revenues grow at a constant rate $\bar{p}$ as x increases. The costs grow at a rate $1/x$. The monopolist would want to sell an infinite amount of its product. However, if $c(x) = x^2$, then $x \in [0, \bar{p}]$, a compact set, and there is solution.

7.3 Maximum Theorems

7.3.1 Correspondences

Correspondence: Let Θ and S be subsets of R^l and R^n, respectively. A *correspondence* Φ from Θ to S is a map that associates with each element $\theta \in \Theta$ a (nonempty) subset $\Theta(\theta) \subset S$.

To distinguish a correspondence notationally from a function, we will denote a correspondence Φ from Θ to S by $\Phi : \Theta \to P(S)$, where $P(S)$ denotes the *power set* of S, i.e., the set of all nonempty subsets of S.

Any function f from Θ to S may also be viewed as a single-valued correspondence from Θ to S.

7.3.2 Upper- and Lower-Semicontinuous Correspondences

Upper-semicontinuous correspondence: A correspondence $\Phi : \Theta \to P(S)$ is said to be *upper-semicontinuous* or *usc* at a point $\theta \in \Theta$ if for all open sets V such that $\Phi(\theta) \subset V$, there exists an open set U containing θ, such that $\theta' \in U \cap \Theta$ implies $\Phi(\theta') \subset V$. We say that Φ is *usc* on Θ if Φ is usc at each $\theta \in \Theta$.

Lower-semicontinuous correspondence: A correspondence $\Phi : \Theta \to P(S)$ is said to be *lower-semicontinuous* or *lsc* at a point $\theta \in \Theta$ if for all open sets V such that $V \cap \Phi(\theta) \neq \varnothing$, there exists an open set U containing θ, such that $\theta' \in U \cap \Theta$ implies $V \cap \Phi(\theta') \neq \varnothing$. The correspondence is said to be lsc on Θ if it is lsc at each $\theta \in \Theta$.

See example 9.1 in Sundaram on page 226.

Continuous correspondence: A correspondence $\Phi : \Theta \to P(S)$ is said to be continuous at $\theta \in \Theta$ if Φ is both usc and lsc at θ. The correspondence is *continuous on* Θ if Φ is continuous at each $\theta \in \Theta$.

7.3.3 Additional Definitions

Let $\Theta \subset R^n$ and $S \subset R^l$. A correspondence $\Phi : \Theta \to P(S)$ is said to be

1. *closed-valued* at $\theta \in \Theta$ if $\Phi(\theta)$ is a closed set;
2. *compact-valued* at $\theta \in \Theta$ if $\Phi(\theta)$ is a compact set; and
3. *convex-valued* at $\theta \in \Theta$ if $\Phi(\theta)$ is a convex set;

7.3.4 The Maximum Theorems

Theorem: The Maximum Theorem *Let $f : X \times \Theta \to \mathbb{R}$ to be a continuous function, and $D : \Theta \to P(S)$ be a compact-valued, continuous correspondence. Let $f^* : \Theta \to R$ and $D^* : \Theta \to P(S)$ be defined by*

$f^*(\theta) = \max\{f)x, \theta)|x \in D(\theta)\}$

$D^*(\theta) = \arg\max\{f(x, \theta)|x \in \mathcal{D}(\theta)\} = \{x \in D(\theta)|f(x, \theta) = f^*(\theta)\}.$

Then $f*$ is a continuous function on Θ, and $D*$ is a compact-valued, upper-semicontinuous correspondence on Θ.

Proof, Sundaram page 235.

Theorem: The Maximum Theorem under Convexity: *Suppose f is a continuous function on $S \times \Theta$ and D is a compact-valued continuous correspondence on Θ. Let*
$f^*(\theta) = \max\{f(x, \theta)|x \in D(\theta)\}$
$D^*(\theta) = \arg\max f(x, \theta)|x \in D(\theta) = x \in D(\theta)|f(x, \theta) = f^*(\theta)\}.$

Then:

1. *f^* is a continuous function on Θ and D^* is a usc correspondence on Θ.*
2. *If $f(., \theta)$ is a concave in x for each θ, and D is convex-valued (i.e., $D(\theta)$ is a convex set for each θ), then D^* is a convex-valued correspondence. When "concave" is replaced by "strictly concave," then D^* is a single-valued usc correspondence, hence a continuous function.*
3. *If f is concave on $S \times \Theta$, and D has a convex graph, then f^* is a concave function, and D^* is a convex-valued usc correspondence. If "concave" is replaced by "strictly concave," then f^* is also strictly concave, and D^* is single-valued everywhere, and therefore, a continuous function.*

Proof, Sundaram page 238.

8 Unconstrained Optimization

In this section we consider problems where the objective function is differentiable. The fact that these problems are unconstrained does not mean there are no constraints: rather, they mean that the constraints are not binding, and therefore we can act as if they are not there.

That is, if the constraint is defined by some set $\mathcal{D}$, the solution will always be in a point inside the interior of the set $\mathcal{D}$. We define the interior of $\mathcal{D}$ as

$$int\mathcal{D} = \{x \in \mathcal{D} : \exists r > 0, s.t. \ B(x, r) \subset \mathcal{D}\}$$

An optimization problem which maximizes f over D achieves a maximum at x, and if $x \in int\mathcal{D}$,this problem is called an unconstrained optimum.

We only deal with local maxima: a point $x \in \mathcal{D}$ is a local maximum of f on $\mathcal{D}$ if there exists $r > 0$ such that $f(x) \geq f(y), \forall y \in B(x, r) \cap \mathcal{D}$. It is an unconstrained local maximum if $B(x, r) \subset \mathcal{D}$,which implies $x \in int\ \mathcal{D}$.

Example: 0 is an unconstrained local maximum of $f(x) = -x^2$ for $x \in R$.

8.1 First Order Conditions

Proposition: Suppose $x^* \in \mathcal{D}$ is an unconstrained local maximum of $f(x)$. If f is differentiable at x^*, then $Df(x^*) = 0$. *The same is true if x^* is a minimum.*

We call the points that make the derivative zero critical points. Be aware that these points can be local maxima, local minima, or a saddle point. The proposition merely says that, if x^* is a maximum, then $Df(x^*)$ is zero. $Df(x^*) = 0$ provides a $necessary$ condition for a local optimal.

Consider the problem

$$\max_x -x^2$$

The first order conditions are

$$2x = 0 \Rightarrow x^* = 0$$

Why should the first order conditions hold for a maximum? If $f'(x) < 0$, then it is possible to increase the value of $f(x)$ by decreasing x. The opposite can be said if $f'(x) > 0$.

The first order conditions are only necessary conditions. It does not say anything about its converse.

Consider the problem

$$\max_x x^2$$

The first order condition is $2x = 0$, which is satisfied at $x = 0$, but this is a minimum.

On the other hand, if the function is x^3, the first order conditions say $x = 0$, but this is neither a maximum nor a minimum.

The second order conditions together with FOC provide the sufficient conditions for a maximum and a minimum. If the second derivative is negative (positive), the point where the first order condition is met is a maximum (minimum). If the domain of the function is a set $D \subset \mathbb{R}^n$, then if $Df(x^*) = 0$ and $D^2 f(x)$ is negative (positive) definite at x^*, x^* is a local maximum (minimum).

8.2 Second Order Conditions

The second order conditions together with FOCs provide sufficient conditions to guarantee an optimum.

Let $F : \mathcal{D} \to \mathbb{R}$ be C^2, where $\mathcal{D} \subseteq \mathbb{R}^n$ is open, and the first order conditions hold at some $x^* \in \mathcal{D}$. Then

1. If x^* is a local maximum, then $D^2 f(x^*)$ is a negative semi definite matrix

2. If x^* is a local minimum, then $D^2 f(x^*)$ is a positive semi definite matrix

3. If $D^2 f(x^*)$ is a negative definite matrix, then x^* is a strict local maximum

4. If $D^2 f(x^*)$ is a positive definite matrix, then x^* is a strict local minimum

Example 1

Find the critical points of

$$f(x) = 2x^3 - 3x^2$$

The first order conditions identify two points at which $f'(x) = 0 : x = 0$ and $x = 1$. The SOC say $f''(0) = -6$ and $f''(1) = 6$. Thus, the function evaluated at $x = 0$ is a local maximum and at $x = 1$ is a local minimum. But neither of these are global minimum or maximum, since $\lim_{x\to\infty} f(x) = \infty$ and $\lim_{x\to-\infty} = -\infty$

Example 2

Find the critical points of

$$f(x, y) = 2x^3 + xy^2 + 5x^2 + y^2$$

.

The FOCs are:

$$\frac{\partial f(x,y)}{\partial x} = 6x^2 + y^2 + 10x = 0$$
$$\frac{\partial f(x,y)}{\partial y} = 2xy + 2y = 0$$

The critical points are $(0, 0), (-1, -2), (-1, 2), (-\frac{5}{3}, 0)$. Calculate the Hessian.

$$H(x,y) = \begin{pmatrix} \frac{\partial^2 f(x,y)}{\partial x^2} & \frac{\partial^2 f(x,y)}{\partial x \partial y} \\ \frac{\partial^2 f(x,y)}{\partial y \partial x} & \frac{\partial^2 f(x,y)}{\partial y^2} \end{pmatrix} = \begin{pmatrix} 12x + 10 & 2y \\ 2y & 2x + 2 \end{pmatrix}$$

$$
\begin{aligned}
H(0,0) &= \begin{pmatrix} 10 & 0 \\ 0 & 2 \end{pmatrix},\ H(-1,-2) = \begin{pmatrix} -2 & -4 \\ -4 & 0 \end{pmatrix}, \\
H(-1,2) &= \begin{pmatrix} -2 & 4 \\ 4 & 0 \end{pmatrix},\ H(-\frac{5}{3},0) = \begin{pmatrix} -10 & 0 \\ 0 & -\frac{4}{3} \end{pmatrix}
\end{aligned}
$$

Check if $H(x,y)$ is psd or nsd at each point.

$H_1(0,0) = 10 > 0,\quad H_2(0,0) = 20 > 0 \Rightarrow$ psd $\Rightarrow (0,0)$ is a local minimum

$H_1(-1,-2) = -2 < 0,\quad H_2(-1,-2) = -16 < 0 \Rightarrow$ cannot say whether max, min, or none

$H_1(-1,2) = -2 < 0,\quad H_2(-1,2) = -16 < 0 \Rightarrow$ cannot say whether max, min, or none

$H_1(-\frac{5}{3},0) = -10 < 0,\quad H_2(-\frac{5}{3},0) = \frac{40}{3} > 0 \Rightarrow$ nsd $\Longrightarrow (-\frac{5}{3},0)$ is local maximum

The point $(0,0)$ is not a global minimum. $f(0,0) = 0 > f(-3,0) = -9$. The point $(-\frac{5}{3},0)$ is not a global maximum. $f(-\frac{5}{3},0) = \frac{40}{3} < f(1,2) = 15$.

Even when this procedure does not identify global maxima, it can help in doing so. Consider the problem example 1, and suppose now x is constrained to the interval $[-1,2]$. We know that the global maximum, must either be at an interior point, where the first derivatives are zero, or at a corner: at $x = -1$ or at $x = 2$. This procedure tells us that there are four options, and we should evaluate the function at each option to find the global maximum. We can easily check that $f(-1) = -5, f(0) = 0, f(1) = -1, f(2) = 4$. Thus, 4 is the global maximum.

Unconstrained optima are usually easier to identify than constrained problems. Thus, it sometimes pays to convert a constrained problem into an unconstrained one. We do this by incorporating the restrictions into the objective function directly.

Example 3

Consider the utility maximization with quasilinear preferences:

$$\max \; u(x_1) + x_2$$
$$\text{subject to}$$
$$p_1 x_1 + x_2 = I$$
$$x_1 \geq 0, x_2 \geq 0$$

We can rewrite the problem as

$$\max \; u(x_1) + I - p_1 x_1$$
$$\text{subject to}$$
$$x_1 \geq 0, I - p_1 x_1 \geq 0$$

Finding the solution to this problem is easy: if $I \geq p_1 x_1^*$, $u'(x_1^*) = p_1, x_2^* = I - p_1 x_1^*$, and if $I < p_1 x_1^*$, then $x_1 = \frac{I}{p_1}$.

Example 4

Another example is the problem of the social planner dividing k goods between two individuals. Let $u : \mathbb{R}^k \to \mathbb{R}$. Assume $u_i(x) > 0, i = 1, ..., k$:

$$\max \; \lambda u(x_1) + (1 - \lambda) u(x_2)$$
$$\text{subject to } x_1 \geq 0, x_2 \geq 0, x_1 + x_2 \leq \omega$$

Since $u_i > 0$, we know that the third constraint will bind, so we can set $x_2 = \omega - x_1$. The problem becomes:

$$\max \; \lambda u(x_1) + (1 - \lambda) u(\omega - x_1)$$

The solution is $x*$ that satisfies

$$\lambda u'(x_1^*) = (1-\lambda)u'(\omega - x_1^*)$$

9 Comparative Statics

We are often interested in the relation between different variables in equilibrium. This relation would tell us how the optimal choice would change given a change in a parameter. For example, in the utility maximization problem, we might want to know how one's quantity demaned changes when there is an income increase.

For this purpose, the implicit function theorem tells us the relation of different variables at optima. The envelope theorem tells us how a small change in parameter value affects the objective function value at optima without the need of explicitly recalculating the solution.

9.1 Implicit Function Theorem

Theorem 8.1: Let $G(x, y)$ be a C^1 function on a ball about (x^*, y^*) in R^2. Suppose that $G(x^*, y^*) = c$ and consider the expression $G(x, y) = c$. If $(\partial G/\partial y)(x^*, y^*) \neq 0$,then there exists a C^1 function $y = y(x)$ defined on an interval U about the point x^* such that:

1. $G(x, y(x)) = c$ for all $x \in U$,
2. $y(x^*) = y^*$, and
3. $y'(x^*) = -\frac{\frac{\partial G}{\partial x}(x^*,y^*)}{\frac{\partial G}{\partial y}(x^*,y^*)}$

Theorem 8.1:: Let $G(x_1, ..., x_k, y)$ be a C^1 function around the point $(x_1^*, ..., x_k^*, y^*)$ Suppose that $(x_1^*, ..., x_k^*, y^*)$ satisfies

$$G(x_1^*, ..., x_k^*, y^*) = c$$

and that

$$\frac{\partial G}{\partial y}(x_1^*, ..., x_k^*, y^*) \neq 0$$

Then there is a C^1 function $y = y(x_1, ..., x_k)$ defined on an open ball B about $(x_1^*, ..., x_k^*)$ so that:

1. $G(x_1, ..., x_k, y(x_1, ..., x_k)) = c$ for all $x \in U$,
2. $y(x_1^*, ..., x_k^*) = y^*$, and
3. for each index i,

$$\frac{\partial y}{\partial x_i}(x_1^*, ..., x_k^*) = -\frac{\frac{\partial G}{\partial x_i}(x_1^*, ..., x_k^*, y^*)}{\frac{\partial G}{\partial y}(x_1^*, ..., x_k^*, y^*)}$$

The implicit function theorem provides sufficient conditions to guarantee the existence of a relation between these variables in an optimum. Moreover, even when it might be impossible to explicitly determine this relation in closed form, the theorem allows us to explore how one variable reacts to a change in a parameter at an optimum.

It is important to make sure that the theorem provides sufficient conditions: if these hold, then we can guarantee the existence of the relation. If these do not hold, then we just can't say whether there is a relation or not.

Example

Consider the problem of a firm minimizing costs subject to producing a given amount of output $\bar{Q}$. The problem is

$$\min_{K,L} \; wL + rK$$

Subject to

$$F(K, L) = \bar{Q}$$

The efficient level of K will depend on L and Q. At the optimal K^*, L^*, how one variable reacts to the changes of the other variable in order to keep optimal?

F is C^1 and $F_K(K^*, L^*) \neq 0$. Applying the implicit function theorem to the constraint, we obtain the relation between capital and labor in equilibrium.

$$\frac{\partial K}{\partial L} = -\frac{F_L(K^*, L^*)}{F_K(K^*, L^*)}$$

Notice that at the optimum, the first order conditions determine:

$$w = \lambda F_L(K, L)$$
$$r = \lambda F_K(K, L)$$

We call $\frac{F_L(K^*,L^*)}{F_K(K^*,L^*)}$ the marginal rate of transformation, and $\frac{w}{r}$ is the marginal rate of substitution, so, in equilibrium, these rates must be the same.

9.2 Inverse Function Theorem

Theorem: Let $f : R^n \to R^n$ be a C^1 function with $f(x^*) = y^*$. If $\partial f(x^*)/\partial x$ is nonsingular, then there exists an open ball $B_r(x^*)$ about x^* and an open set V about y^* such that f is a one-to-one and onto map from $B_r(x^*)$ to V. The natural inverse $f^{-1} : V \to B_r(x^*)$ is also C^1 and

$$\partial f^{-1}(y^*)/\partial y = \frac{1}{\partial f(x^*)/\partial x}.$$

9.3 Envelope Theorem without constraints

The envelope theorem tells how a parameter change would impact our objective function at optima.

Theorem 8.3: Let X be the choice set and $\theta \in \Theta$ a relevant parameter. Let $f : X * \Theta \to \mathbb{R}$ denote the parametrized objective function. The value V and the optimal choice function X^* are given by:

1. $V(\theta) = \sup_{x \in X} f(x, \theta)$

2. $x^*(\theta) = \{x \in X : f(x, \theta) = V(\theta)\}$

   The envelope theorem states that if $\theta \in int\Theta$ and V is differentiable at θ, then $V'(\theta) = f_\theta(x^*, \theta)$.

   Proof (Milgrom and Segal 2002): for any $\theta \in int\Theta$, $\forall \theta' \in \Theta$

$$\begin{aligned} \frac{f(x^*(\theta), \theta') - f(x^*(\theta), \theta)}{\theta' - \theta} &\leq \frac{V(\theta') - V(\theta)}{\theta' - \theta} \\ \frac{f(x^*(\theta'), \theta') - f(x^*(\theta'), \theta)}{\theta' - \theta} &\geq \frac{V(\theta') - V(\theta)}{\theta' - \theta} \end{aligned}$$

   Taking the limit as $\theta' \to \theta$, we get

$$f_\theta(x^*, \theta) \leq V'(\theta)$$

$$f_\theta(x^*, \theta) \geq V'(\theta)$$

   Since V is differentiable at θ, it must be that

$$f_\theta(x^*, \theta) = V'(\theta)$$

Example

The firm maximizing profits. The value function is

$$\pi(w, r) = \max_{L,K} \; F(L, K) - wL - rK$$

We are interested in how profits would change given a change in wages. The envelope theorem says $\pi_w(w, r) = -L^*$. If we didn't have the envelope theorem, we would calculate this as follows:

$$\begin{aligned} \pi(w, r) &\equiv F(L^*, K^*) - wL^* - rK^* \\ \frac{\partial \pi(w, r)}{\partial w} &\equiv (F_L(L^*, K^*) - w)\frac{\partial L^*}{\partial w} + (F_K(L^*, K^*) - r)\frac{\partial K^*}{\partial w} - L^* \end{aligned}$$

The first order conditions say that the terms in brackets are zero.

10 Equality Constraints: Lagrange

The most common problems in economics deal with equality constraints. For example the utility maximization problem has an equality constraint. The problem is

$$\max_x f(x)$$
$$\text{subject to } g(x) = 0$$

To solve this, we use the Lagrange theorem.

Lagrange theorem: *Let* $f : \mathbb{U} \to \mathbb{R}$ *and* $g_i : \mathbb{U} \to \mathbb{R}$ *be* C^1 *functions,* $i = 1, 2, \ldots, k$. *Suppose* x^* *is a local maximum or minimum of f on the set*

$$\mathcal{D} = U \cap \{x \in \mathbb{R}^n : g_i(x) = 0,\ i = 1, \ldots, k\}$$

where $U \subset \mathbb{R}^n$ *is an open set. Suppose also that* $rank(Dg(x^*)) = k$ *or* $Dg(x^*)$ *has* full rank. *Then there exists a vector* $\lambda^* = \lambda_1, \ldots, \lambda_k$ *such that*

$$\frac{\partial f(x^*)}{\partial x_j} - \sum_{i=1}^{k} \lambda_i \frac{\partial g_i(x^*)}{\partial x_j} = 0,\ j = 1, ..., n$$

The theorem provides **necessary** conditions for a constrained optimum. As before, the SOCs together with this FOCs provide sufficient conditions. It is important that $rank(Dg(x^*)) = k$. This is the constraint qualification. If this does not met, the theory of Lagrange does not apply.

The first step to solve the problem use Lagrangean is to build a Lagrangean:

$$\mathcal{L}(x; \lambda) = f(x) - \sum_{i=1}^{k} \lambda_i g_i(x)$$

The first order conditions from maximizing the Lagrangean are

$$\begin{aligned}
\frac{\partial \mathcal{L}(x^*;\lambda)}{\partial x_j} &= \frac{\partial f(x^*)}{\partial x_j} - \sum_{i=1}^{k} \lambda_i \frac{\partial g_i(x^*)}{\partial x_j} = 0,\ j = 1, ..., n \\
\frac{\partial \mathcal{L}(x^*;\lambda)}{\partial \lambda_i} &= g_i(x) = 0,\ i = 1, ..., k
\end{aligned}$$

The second order conditions are the Hessian,

$$H = \begin{pmatrix} \frac{\partial^2 \mathcal{L}(x^*;\lambda)}{\partial \lambda \partial \lambda'} & \frac{\partial^2 \mathcal{L}(x^*;\lambda)}{\partial \lambda \partial x'} \\ \frac{\partial^2 \mathcal{L}(x^*;\lambda)}{\partial x \partial \lambda'} & \frac{\partial \mathcal{L}^2(x^*;\lambda)}{\partial x \partial x'} \end{pmatrix}$$

The second order conditions define a quadratic equation. As before, if this equation is positive definite, the solution is a minimum, and if it is negative definite, the solution is a maximum. We need check for negative/positive definiteness.

The Hessian:

$$\begin{aligned}
H &= \begin{pmatrix} \frac{\partial^2 \mathcal{L}(x^*;\lambda)}{\partial \lambda \partial \lambda'} & \frac{\partial^2 \mathcal{L}(x^*;\lambda)}{\partial \lambda \partial x'} \\ \frac{\partial^2 \mathcal{L}(x^*;\lambda)}{\partial x \partial \lambda'} & \frac{\partial \mathcal{L}^2(x^*;\lambda)}{\partial x \partial x'} \end{pmatrix} \\
&= \begin{pmatrix}
0 & \dots & 0 & | & -\frac{\partial g_1(x^*)}{\partial x_1} & \dots & -\frac{\partial g_1(x^*)}{\partial x_n} \\
0 & \dots & 0 & | & \vdots & \ddots & \vdots \\
0 & \dots & 0 & | & -\frac{\partial g_k(x^*)}{\partial x_1} & \dots & -\frac{\partial g_k(x^*)}{\partial x_n} \\
- & - & - & | & - & - & - \\
-\frac{\partial g_1(x^*)}{\partial x_1} & \dots & -\frac{\partial g_k(x^*)}{\partial x_1} & | & \frac{\partial \mathcal{L}(x^*)}{\partial x_1^2} & \dots & \frac{\partial \mathcal{L}(x^*)}{\partial x_1 \partial x_n} \\
\vdots & \ddots & \vdots & | & \vdots & \ddots & \vdots \\
-\frac{\partial g_1(x^*)}{\partial x_n} & \dots & -\frac{\partial g_k(x^*)}{\partial x_n} & | & \frac{\partial \mathcal{L}(x^*)}{\partial x_1 \partial x_n} & \dots & \frac{\partial L(x^*)}{\partial x_n^2}
\end{pmatrix}
\end{aligned}$$

If the last $(n-k)$ leading principal minors of H alternate in sign, with the sign of the determinant of H equal to $(-1)^n$, then the solution is a local maximum. For a minimum, the last $(n-k)$ leading principal minors all have the same sign as $(-1)^k$.

Example 1: Let f and g be C^2 functions on $\mathbb{R}^2$. Consider the problem of maximizing f on the constraint $g(x,y) = c$.

Form the lagrangean,

$$\mathcal{L}(x, y, \lambda) = f(x, y) - \lambda(g(x, y) - c)$$

Suppose that (x^*, y^*, λ) satisfies $\frac{\partial \mathcal{L}(x^*,y^*,\lambda^*)}{\partial x} = 0$, $\frac{\partial \mathcal{L}(x^*,y^*,\lambda^*)}{\partial y} = 0$.

The solution is a local maximum if

$$\begin{vmatrix} 0 & \frac{\partial g(x^*,y^*,\lambda^*)}{\partial x} & \frac{\partial g(x^*,y^*,\lambda^*)}{\partial y} \\ \frac{\partial g(x^*,y^*,\lambda^*)}{\partial x} & \frac{\partial^2 \mathcal{L}(x^*,y^*,\lambda^*)}{\partial x^2} & \frac{\partial^2 \mathcal{L}(x^*,y^*,\lambda^*)}{\partial x \partial y} \\ \frac{\partial g(x^*,y^*,\lambda^*)}{\partial y} & \frac{\partial^2 \mathcal{L}(x^*,y^*,\lambda^*)}{\partial y \partial x} & \frac{\partial^2 \mathcal{L}(x^*,y^*,\lambda^*)}{\partial y^2} \end{vmatrix} > 0$$

If the sign is negative, the solution is a minimum.

10.1 Constraint Qualification

This condition says that the rank of $Dg(x^*)$ be equal to the number of constraints. $Dg(x^*)$ is the Jacobian of the constraint $g(x)$ *at* x^*. If this does not met, then the theorem of Lagrange does not apply. Thus, after finding first order conditions you always need to verify if there are any points at which the constraint qualification is violated. The maximum might be at this point and the Lagrange method will not identify it.

For example, consider the problem

$$\max -y$$
$$s.t.\ y^3 - x^2 = 0$$

Since $x^2 \geq 0, \forall x$, we have $y \geq 0$. Thus, the function is maximized at $y = 0$, where $x = 0$. At this point, the constraint qualification is not met:

$$Dg(x^*, y^*) = (-2x^*, 3y^{*2}) = (0, 0)$$

so $rank(Dg(x^*, y^*)) = 0$. Thus, the Lagrange method can not identify this optimal points, and there is no λ such that $Df(0, 0) + \lambda Dg(0, 0) = 0$.

10.2 A Cookbook Procedure: The Lagrangean Method

Consider the problem

$$\max_x f(x)$$
$$s.t.\ g(x) = 0$$

Step 1**:**

Identify the points $x \in \mathcal{D}$ in which the constraint qualification does not hold. Evaluate the function f at these points.

Step 2:

Build the Lagrangean:

$$\mathcal{L}(x; \lambda) = f(x) - \sum_{i=1}^{k} \lambda_i g_i(x)$$

Step 3:

Find all critical points.

$$\frac{\partial \mathcal{L}}{\partial x_j}(x; \lambda) = 0 \qquad \forall j = 1, \ldots, n$$
$$\frac{\partial \mathcal{L}}{\partial \lambda_i}(x; \lambda) = 0 \qquad \forall i = 1, \ldots, k$$

Evaluate f at each of these points, and also at the points where the constraint qualification fails.

Step 4:

Compare the function f at each of the points evaluated.

Example 1

Consider the problem:

$$\begin{aligned} \max\ x \\ st. x^3 + y^2 &= 0 \end{aligned}$$

1. What is the optimal solution for this problem?

2. Does the Lagrange method work in this problem? Why?

Solution:

1. Since $x^3 + y^2 = 0$, $x^3 = -y^2 \leq 0$ or $x \leq 0$. Thus, the maximum value of this problem is 0, which is reached at point $(x, y) = (0, 0)$.

2. The Jacobian of the constraint is:

$$(3x^2, 2y)$$

The constraint qualification is violated at $(x, y) = (0, 0)$, where the rank of the Jacobian is 0.

3. The Lagrangian of the problem is

$$L(x, y) = x - \lambda(x^3 + y^2)$$

and the FOCs are:

$$\begin{aligned} \frac{\partial L}{\partial x} &= 1 - 3\lambda x^2 = 0 \\ \frac{\partial L}{\partial y} &= -2\lambda y = 0 \\ \frac{\partial L}{\partial \lambda} &= x^3 + y^2 = 0 \end{aligned}$$

The second FOC will be satisfied by either $\lambda = 0$ or $y = 0$.

a. if $\lambda = 0$, the first FOC $\frac{\partial L}{\partial x} = 1 \neq 0$..Thus, $\lambda = 0$ can not be a solution.

b. if y=0, the third FOC $\Longrightarrow x = 0$, then the first FOC $\frac{\partial L}{\partial x} = 1 \neq 0$.

Thus, $y = 0$ can not be a solution.

Therefore, no solution for the FOCs of Lagrangian.

Example 2

Consider the problem:

$$\max \ xy$$

Subject to

$$x^2 + y^2 = 2a^2, \ a > 0$$

Step 1: Analyze the constraint qualification:

$$Dg(x, y) = \begin{pmatrix} 2x, & 2y \end{pmatrix}$$

The constraint qualification fails at $(0, 0)$. $f(0, 0) = 0$.

Step 2: Set up the Lagrangean

$$\mathcal{L}(x, y) = xy - \lambda(x^2 + y^2 - 2a^2)$$

Step 3: Find the critical points:

$$\begin{aligned} x: & \quad y - 2\lambda x = 0 \\ y: & \quad x - 2\lambda y = 0 \\ \lambda: & \; 2a^2 - x^2 - y^2 = 0 \end{aligned}$$

The first FOC $\implies y = 2\lambda x$, insert in second FOC to get $x - 4\lambda^2 x = 0$.This equation provides two alternatives: either $x = 0$ or $x \neq 0$.

Case 1: $x = 0$

From the first FOC, this implies $y = 0$, but this cannot be a solution since it violates the constraint.

Case 2: $x \neq 0$

$x - 4\lambda^2 x = 0$ implies $\lambda = \pm 1/2$.

Case 2.1: $\lambda = 1/2$

From first FOC, $y = x$. From the third FOC, $x^2 = a^2$. Thus, we identify two critical points:

$$
\begin{aligned}
(x_1, y_1, \lambda_1) &= (a, a, 1/2) \\
(x_2, y_2, \lambda_2) &= (-a, -a, 1/2)
\end{aligned}
$$

Case 2.2: $\lambda = -1/2$

From first FOC, $y = -x$. This identifies two new critical points:

$$
\begin{aligned}
(x_3, y_3, \lambda_3) &= (-a, a, -1/2) \\
(x_4, y_4, \lambda_4) &= (a, -a, -1/2)
\end{aligned}
$$

Evaluate the objective function at each point:

$$
\begin{aligned}
f(x_1, y_1) &= f(x_2, y_2) = a^2 \\
f(x_3, y_3) &= f(x_4, y_4) = -a^2
\end{aligned}
$$

Step 4:

Compare the 5 points. The maximum points are at (x_1, y_1, λ_1) and (x_2, y_2, λ_2) and the minimum points at (x_3, y_3, λ_3) and (x_4, y_4, λ_4).

Work out the problem of utility maximization subject to the budget constraint in Chapter 5, page 128 of Sundaram.

10.3 Envelope Theorem with Constraint Set

Let $X \subset \mathbb{R}^n$ for some integer n, Θ be an open real interval, and $f : X \times \Theta \to \mathbb{R}$ and $g : \mathbb{R}^n \to \mathbb{R}^m$ for some integer m. Consider the problem

$$
\begin{aligned}
V(\theta) &= \max_{x \in \mathbb{R}^n} f(x, \theta) \\
s.t.\ g(x, \theta) &= 0
\end{aligned}
$$

The lagrangean is $\mathcal{L}(x, \lambda, \theta) = f(x, \theta) - \lambda \cdot g(x, \theta)$. Assume that

1. The solution $x^*(\theta)$ to the problem exists and is unique for every $\theta \in \Theta$.

2. The functions f, g, and x^* are continuously differentiable on their domains.

3. For a point $\theta_0 \in \Theta$, the FOCs hold at $(x^*(\theta_0), \theta_0)$ with associated multipliers $\lambda^* = (\lambda_1^*, ..., \lambda_m^*) \geq 0$.

Then V is differentiable at θ_0 and

$$V'(\theta_0) = \mathcal{L}_\theta(x^*(\theta_0), \lambda^*, \theta_0).$$

Proof: Differentiate $V(\theta) = f(x^*(\theta), \theta) - \lambda^*(\theta) \cdot g(x^*(\theta), \theta)$ with respect to θ, evaluate it at $\theta = \theta_0$, and use the first order conditions.

In words: under the three assumptions, the value function is differentiable and we can calculate it simply by differentiating the Langrangean $\mathcal{L}(x, \lambda, \theta)$ *before optimizing*, then evaluate the derivative at the point (x, λ) that solves this problem.

11 Inequality Constraints

11.1 Inequality Constraints

Most problems in economics deal with inequality constraints. There is a natural extension of the Lagrange theorem to deal with inequality constraints, the Kuhn-Tucker theorem.

$$\max_x \quad f(x)$$
$$\text{subject to}$$
$$x \in \mathcal{D} = \{x \in \mathbb{R}^n : g(x) \leq 0\}$$

The Kuhn Tucker theorem:

Let $f : \mathbb{U} \to \mathbb{R}$ and $g_i : \mathbb{U} \to \mathbb{R}$ be C^1 functions, $i = 1, 2, \ldots, k$. Suppose x^* is a local maximum or minimum of f on the set

$$\mathcal{D} = U \cap \{x \in \mathbb{R}^n : g_i(x) \leq 0, i = 1, \ldots, k\}$$

where $U \subset \mathbb{R}^n$ is an open set. Suppose also that $rank(Dg(x)) = k$ or $Dg(x^*)$ has full rank*. Then there exists a vector $\lambda^* = \lambda_1, \ldots, \lambda_k$ such that*

$$\frac{\partial \mathcal{L}(x^*; \lambda)}{\partial x_j} = \frac{\partial f(x^*)}{\partial x_j} - \sum_{i=1}^{k} \lambda_i \frac{\partial g_i(x^*)}{\partial x_j} = 0, \; j = 1, ..., n$$
$$\lambda_i^* \geq 0 \text{ and } \lambda_i^* g_i(x^*) = 0, \; i = 1, ..., k.$$

The condition $\lambda_i^* g_i(x^*) = 0$ is called the complementary slackness condition. Note that if the constraint does not bind, and $g_i(x^*) < 0$, then it must be that $\lambda_i^* = 0$. Recall that in the Lagrange problem the multipliers were the value of loosening the constraint. If the constraint does not bind, then loosening them has no value, which implies that the corresponding multiplier is zero.

As with Lagrange, the theorem provides **necessary** conditions for a constrained optimum. As before, the second order conditions together with the FOCs provide sufficient conditions. It is important that $rank(Dg(x^*)) = k$. This is the constraint qualification.

Procedure: To solve a problem with inequality constraints we proceed as with equality constraints.

First,find the points that the constraint qualification fail. i.e., discuss all of the possible combinations of binding constraints, and for each combination, check the jacobian matrix of the binding constraint only, and then get the points where the constraint qualification fail.

Second, build a Lagrangean:

$$\mathcal{L}(x;\lambda) = f(x) - \sum_{i=1}^{k} \lambda_i g_i(x)$$

Third, find first order conditions:

$$\frac{\partial \mathcal{L}(x^*;\lambda)}{\partial x_j} = \frac{\partial f(x^*)}{\partial x_j} - \sum_{i=1}^{k} \lambda_i \frac{\partial g_i(x^*)}{\partial x_j} = 0,\ j = 1, ..., n$$

$$\lambda_i^* \geq 0 \text{ and } \lambda_i^* g_i(x^*) = 0,\ i = 1, ..., k.$$

identify all the feasible solutions (x^*, λ^*) that solve the system.

Fourth, evaluate the objective function at each point that are found in the second and third step, and compare the solutions.

Example 1:

$$\max\ x^2 - y$$

$$s.t. 1 - x^2 - y^2 \geq 0$$

Since the constraint set is compact and the objective function is continuous, a maximum exists.

The Jacobian of the constraint is $Dg(x,y) = (2x, 2y)$, the constraint qualification fails at $x = y = 0$, and $f(x,y) = 0$.

Set up the Lagrangean:

$$\mathcal{L} = x^2 - y + \lambda(1 - x^2 - y^2)$$

The FOCs are:

$$\begin{aligned}
&2x - 2\lambda x = 0 \\
&-1 - 2\lambda y = 0 \\
&\lambda \geq 0,\ 1 - x^2 - y^2 \geq 0,\ \lambda(1 - x^2 - y^2) = 0
\end{aligned}$$

The first condition says that either $x = 0$ or $\lambda = 1$. If $\lambda = 1$, then the critical point is

$$(x, y, \lambda) = \left(\pm\frac{\sqrt{3}}{2}, -\frac{1}{2}, 1\right), \text{ and } f(x,y) = \frac{5}{4}$$

If $x = 0$, then if $\lambda \neq 0 \Longrightarrow 1 - x^2 - y^2 = 0 \Longrightarrow y = \pm 1$, we have

$$(x, y, \lambda) = \left(0, -1, \frac{1}{2}\right), \text{ and } f(x,y) = 1 < \frac{5}{4}$$

if $\lambda = 0$, contradicted with the second FOC.

So the solution is $(x,y) = \left(\pm\frac{\sqrt{3}}{2}, -\frac{1}{2}\right)$.

11.2 Mixed Constraints

The optimization problem is:

$$\begin{aligned}
&\max_x \ f(x) \\
&\qquad \text{subject to} \\
&\qquad x \in \mathcal{D} = \{x \in \mathbb{R}^n : g(x) = 0, h(x) \leq 0\}
\end{aligned}$$

where $g : R^n \to R^k$ and $h : R^n \to R^l$.

First find all of the points that constraint qualification fails.

Second, build a Lagrangean:

$$\mathcal{L}(x; \lambda) = f(x) - \sum_{i=1}^{k} \lambda_i g_i(x) - \sum_{j=1}^{l} \lambda_j h_j(x)$$

Identify first order conditions:

$$\begin{aligned}
\frac{\partial \mathcal{L}(x^*; \lambda)}{\partial x_j} &= \frac{\partial f(x^*)}{\partial x_j} - \sum_{i=1}^{k} \lambda_i^* \frac{\partial g_i(x^*)}{\partial x_j} - \sum_{j=1}^{l} \lambda_j^* \frac{\partial h_i(x^*)}{\partial x_j} = 0 \\
g_i(x) &= 0, \text{ for } i = 1, ...k. \\
\lambda_j^* &\geq 0 \text{ and } \lambda_j^* h_j(x^*) = 0 \text{ for } j = 1, ..., l
\end{aligned}$$

Finally, identify all the feasible solutions (x^*, λ^*) that solve the system, evaluate the value function at each point find in the first and second step, and compare.

12 One Last Example: Dynamic Optimization

The following is an example of the standard growth model. There are two kinds of solution macroeconomists are mostly interested in. The first one is the social planner's solution. In these cases, we assume that there is a well informed, benevolent agent that makes the decision for everyone. The solution to problems involving the social planner are always efficient, in the sense that, given the characteristics of the problem, there is no way to make some agent better off without making anyone else worse off. The second concept is the equilibrium concept, where each agent makes decisions to make themselves better off, without caring about anyone else.

The following is a special case of the standard growth model, and the solution is the social planner's solution. In general, these problems do not admit closed form solutions because they involve solving second degree differential equations. The next example is an exception.

The standard growth model is

$$\max_{\{c_t,h_t,x_t\}} \quad \sum_{t=0}^{\infty} \beta^t U(c_t, h_t)$$

subject to

$$\begin{aligned} & c_t + x_t = F(h_t, k_t) \\ & k_{t+1} = (1-\delta)k_t + x_t \\ & k_0 \text{ given }, 0 \leq h_t \leq 1, c_t \geq 0, k_t \geq 0 \qquad \forall t \end{aligned}$$

For this problem to have a solution we need to assume $\beta < 1$. Intuitively, this means that the representative agent values the present more than the future. This is consistent with positive real interest rates.

To allow for closed form solutions, assume

$$U(c,h) = \ln(c)$$
$$F(H,K) = H^{1-\alpha}K^{\alpha}$$
$$\delta = 1$$

Thus, the problem becomes

$$\max_{\{c_t,h_t,k_t\}} \quad \sum_{t=0}^{\infty} \beta^t \ln c_t$$

subject to

$$c_t + k_{t+1} = h_t^{1-\alpha}k_t^{\alpha}$$
$$k_0 \text{ given }, 0 \leq h_t \leq 1, c_t \geq 0, k_t \geq 0 \qquad \forall t$$

The restriction on $0 < \alpha < 1$ is necessary to have a well behaved solution (provides strict concavity, and therefore a unique solution).

It is straightforward to verify that at the optimum $h_t = 1$ for all t (the solution is a corner). Instead of setting up a Lagrangean, I will solve for c_t in the restriction and replace in the objective function. Thus, the maximization is now

$$\max_{\{k_t\}} \quad \sum_{t=0}^{\infty} \beta^t \ln\left(k_t^{\alpha} - k_{t+1}\right)$$

subject to

$$k_t^{\alpha} - k_{t+1} \geq 0, k_t \geq 0 \qquad \forall t$$

To solve, let's ignore the constraint and then check whether the solution satisfies it. If it does, we are done. The first order condition with respect to any particular k_{t+1} is

$$-\frac{\beta^t}{k_t^{\alpha} - k_{t+1}} + \beta^{t+1}\alpha\frac{k_{t+1}^{\alpha-1}}{k_{t+1}^{\alpha} - k_{t+2}} = 0$$

This is a second degree difference equation. To solve it, we need two fixed points. The first point is k_0. The second point is the steady state. The steady state capital stock is the level of capital stock $k_s > 0$ at which the following holds: if $k_t = k_s \Rightarrow k_{t+1} = k_s$. It is easy to find the steady state in this problem. Just set $k_t = k_{t+1} = k_{t+2} = k_s$ to find

$$1 = \beta\alpha k_s^{\alpha-1}$$

or $k_s = (\beta\alpha)^{1-\alpha}$.

Thus, k_s and k_0 determine the initial and final conditions to solve the difference equation. But we will not follow this procedure. The functional forms that we are using, plus the full depreciation, make the problem easier to solve, and the solution can be obtained in closed form. This problem can be solved by "guessing" a solution and verifying whether it is so. Since the solution is unique, this must be the global maximum.

Suggest the following solution: $k_{t+1} = Ak_t^{\alpha}$. Insert this solution in the first order condition. This becomes

$$\frac{\beta\alpha k_{t+1}^{\alpha-1}}{(1-A)k_{t+1}^{\alpha}} = \frac{1}{(1-A)k_t^{\alpha}} \Rightarrow$$
$$k_{t+1} = \beta\alpha k_t^{\alpha}$$

Given the value of k_0, the solution form identifies the entire sequence $\{k_t\}_{t=0}^{\infty}$. If $k_0 > 0 \Rightarrow \{k_t\}_{t=0}^{\infty} > 0$, and $k_t^{\alpha} - k_{t+1} = k_t^{\alpha}(1-\alpha\beta) > 0$, so the restrictions we originally ignored are satisfied, and the solution we identified is a global maximum.

13 Appendix

13.1 Matrix

$A_{n\times m}$ is an $n \times m$ matrix, which is a rectangular array of numbers, denote

$$A_{n\times m} = \underset{(n\times m)}{\begin{pmatrix} a_{11} & \dots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nm} \end{pmatrix}}$$

13.1.1 Transpose

A^T is a transpose of A:

$$A^T = \underset{(m\times n)}{\begin{pmatrix} a_{11} & \dots & a_{n1} \\ \vdots & \ddots & \vdots \\ a_{1m} & \dots & a_{nm} \end{pmatrix}}$$

13.1.2 Addition and Multiplication

Scalar Multiplication: For $\alpha \in \mathbb{R}$

$$\alpha A = \begin{pmatrix} \alpha a_{11} & \dots & \alpha a_{1m} \\ \vdots & \ddots & \vdots \\ \alpha a_{n1} & \dots & \alpha a_{nm} \end{pmatrix}$$

$$(n \times m)$$

Let A and B be $(n \times m)$ matrices. Then $A + B$ is a $(n \times m)$ matrix with the (i, j) element $a_{ij} + b_{ij}$

$$A + B = \begin{pmatrix} a_{11} + b_{11} & \dots & a_{1m} + b_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} + b_{n1} & \dots & a_{nm} + b_{nm} \end{pmatrix}$$

AB^T is a $(n \times n)$ matrix, the (i, j) element is the inner product of the i-th row of A and j-th column of B^T,

$$AB^T = \begin{pmatrix} \sum_{i=1}^m a_{1i} b_{i1} & \dots & \sum_{i=1}^m a_{1i} b_{in} \\ \vdots & \ddots & \vdots \\ \sum_{i=1}^m a_{ni} b_{i1} & \dots & \sum_{i=1}^m a_{ni} b_{in} \end{pmatrix}$$

Properties:

1) $A + B = B + A$

note: AB is not necessarily equal BA

2) $A + (B + C) = (A + B) + C$

3) $A(B + C) = AB + AC$

$A(BC) = (AB)C$

4) $(A + B)^T = A^T + B^T$

$(AB)^T = B^T A^T$

$(\alpha A)^T = \alpha A^T, \alpha \in \mathbb{R}$

13.1.3 Rank of a Matrix

Definition: The vectors $x_1, \ldots, x_k$ are linearly dependent if $\exists \alpha_i \in \mathbb{R}, i = 1, \ldots, k, \alpha_i \neq 0$ for some i such that

$$\sum_{i=1}^{k} \alpha_i x_i = 0$$

If there are no such real numbers, the vectors are linearly independent.

Example 1: The vectors

$$x_1 = \begin{pmatrix} 1 \\ 7 \\ 3 \end{pmatrix}, x_2 = \begin{pmatrix} 9 \\ 2 \\ 1 \end{pmatrix}, x_3 = \begin{pmatrix} 10 \\ 9 \\ 4 \end{pmatrix}$$

are linearly dependent. Take $(\alpha_1, \alpha_2, \alpha_3) = (1, 1, -1)$, then $\sum_{i=1}^{3} \alpha_i x_i = 0$. ■

Example 2: The vectors

$$x_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, x_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, x_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

are linearly independent. Take any $(\alpha_1, \alpha_2, \alpha_3) \neq (0, 0, 0)$,

$$\alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3 = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

■

Definition: The row rank of a matrix A is defined as the maximum number of linearly independent row vectors in A.

The column rank is defined similarly. The row rank of A is always equal to the column rank of A. We say A is of full rank if all the rows or columns are linearly independent.

Theorem Let A be a given $n \times m$ matrix. If B is an nxm matrix obtained from A

1. by interchanging any two rows of A, or

2. by multiplying each entry in a give row by a nonzero constant α, or

3. by replacing a given row, say the i-th by itself plus a scalar multiple α of some other row, say the j-th,

the rank $\rho(B)$ of B is the same as the rank $\rho(A)$ of A. The same result is true if the word "row" in each of the operations above is replaced by "column".

Properties: $rank(A) = rank(A^T)$

$rank(AB) \leq \min\{rank(A), rank(B)\}$

13.1.4 Determinants

The determinant is a transformation that assigns every square matrix A a real number, denoted $|A|$. If matrix A does not have full rank, its determinant is zero. The determinants of full rank matrices are different from zero.

Define $A(ij)$ as the submatrix that results from deleting row i and column j from A.

Define the $(i,j)-$th cofactor of A as $C_{ij}(A) = (-1)^{i+j}|A(ij)|$.

The determinant of A is:

$$|A| = \sum_{j=1}^{n} a_{ij} C_{ij}(A), \text{ for any } i$$

In the case of 2×2 matrices, $C_{11}(A) = a_{22}$ and $C_{12}(A) = -a_{21}$, so the determinant is

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

In the (3×3) case,

$$|A| = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

Example 1.

$$A = \begin{pmatrix} 1 & 3 \\ 4 & 1 \end{pmatrix}$$

■

Example 2.

$$A = \begin{pmatrix} 1 & 3 & 1 \\ 4 & 1 & 2 \\ 2 & 1 & 1 \end{pmatrix}$$

■

13.1.5 Inverse

Let A be a $(n \times n)$ matrix. The inverse of A, denoted as A^{-1} has the property that $AA^{-1} = I$, where I is the identity matrix. A matrix has an inverse if and only if its determinant is nonzero. Define the matrix $C(A)$ as:

$$C(A) = \begin{pmatrix} C_{11}(A) & \dots & C_{1n}(A) \\ \vdots & \ddots & \vdots \\ C_{n1}(A) & \dots & C_{nn}(A) \end{pmatrix}$$

where $C_{ij}(A) = (-1)^{i+j}|A(ij)|$.

The adjoint of A is the transpose of $C(A)$:

$$adj(A) = \begin{pmatrix} C_{11}(A) & \dots & C_{n1}(A) \\ \vdots & \ddots & \vdots \\ C_{1n}(A) & \dots & C_{nn}(A) \end{pmatrix}$$

The inverse can be obtained by dividing every element in $adj(A)$ by the determinant $|A|$.

$$A^{-1} = \frac{adj(A)}{|A|}$$

For example, the inverse of a 2X2 matrix is

$$A^{-1} = \frac{1}{a_{11}a_{22} - a_{21}a_{12}} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}$$

Example 1:

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

■

Properties:

(1) $(A^{-1})^{-1} = A$

(2) $(AB)^{-1} = B^{-1}A^{-1}$

(3) $|A^{-1}| = \frac{1}{|A|}$

(4) $(A^T)^{-1} = (A^{-1})^T$

13.1.6 Definiteness

Theorem 1: Let A be any symmetric $(n \times n)$ matrix, A is

(1) Positive definite if $x'Ax > 0$ for all $x \in \mathbb{R}^n$ and $x \neq 0$.

(2) Positive semidefinite if $x'Ax \geq 0$ for all $x \in \mathbb{R}^n$ and $x \neq 0$.

(3) Negative definite if $x'Ax < 0$ for all $x \in \mathbb{R}^n$ and $x \neq 0$.

(4) Negative semidefinite if $x'Ax \leq 0$ for all $x \in \mathbb{R}^n$ and $x \neq 0$.

A way to check for definiteness is as follows. Let A_k be the matrix that results from deleting the rows and columns from $(k+1)$ to n in A. A_k is called the k-th principal minor of matrix A. Then A is

(1) Negative definite if and only if $(-1)^k|A_k| > 0$ for all $k \in \{1, \ldots, n\}$.

(2) Positive definite if and only if $|A_k| > 0$ for all $k \in \{1, \ldots, n\}$. ■

We cannot change the strict inequality to a weak one to get semidefiniteness. To see why, consider the matrix

$$\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

Notice $|A_1| = |A_2| = 0$, thus, A would be considered negative semidefinite. However, for any $x = (x_1, x_2) \neq (0, 0)$, $x'Ax = x_2^2 \geq 0$, and therefore the conclusion would be wrong.

13.1.7 Matrix Algebra

This sections specifies a way to find a solution to a square linear system of equations $Ax = B$, where A is a $(n \times n)$ matrix, $x \in \mathbb{R}^n$ is the variable we want to solve for and B is an $(n \times 1)$ matrix. There are several ways of attacking this problem. One of the way is to use Cramer's rule.

Cramer's Rule

Notice that if A^{-1} exists, the solution is simply $x = A^{-1}B$,

$$x = A^{-1}B = \frac{Adj(A)}{|A|}B = \frac{1}{|A|}\begin{pmatrix} C_{11}(A) & \dots & C_{n1}(A) \\ \vdots & \ddots & \vdots \\ C_{1n}(A) & \dots & C_{nn}(A) \end{pmatrix} B = \frac{1}{|A|}\begin{pmatrix} \sum_{i=1}^{n} C_{i1}b_i \\ \sum_{i=1}^{n} C_{i2}b_i \\ \vdots \\ \sum_{i=1}^{n} C_{in}b_i \end{pmatrix} \Rightarrow$$

$$x_j^* = \frac{1}{|A|}\sum_{i=1}^{n} C_{ij}b_i$$

Now, define A^i as the matrix A with the column i replaced by the vector B. In the case with $n = 3$

$$A^1 = \begin{pmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{pmatrix}$$

Notice that $|A^1| = \sum_{i=1}^{3} C_{i1}b_i$. Generally, $|A^j| = \sum_{i=1}^{n} C_{ij}b_i$ for all of the n. According to Cramer's rule,

$$x_j^* = \frac{|A^j|}{|A|}$$

Example 1: Use Cramer's rule to solve

$$6x_1 + 3x_2 + x_3 = 22$$
$$x_1 + 4x_2 - 2x_3 = 12$$
$$4x_1 - x_2 + 5x_3 = 10$$

Solution

$$|A| = \begin{vmatrix} 6 & 3 & 1 \\ 1 & 4 & -2 \\ 4 & -1 & 5 \end{vmatrix} = 52 \qquad |A^1| = \begin{vmatrix} 22 & 3 & 1 \\ 12 & 4 & -2 \\ 10 & -1 & 5 \end{vmatrix} = 104$$

$$|A^2| = \begin{vmatrix} 6 & 22 & 1 \\ 1 & 12 & -2 \\ 4 & 10 & 5 \end{vmatrix} = 156 \qquad |A^3| = \begin{vmatrix} 6 & 3 & 22 \\ 1 & 4 & 12 \\ 4 & -1 & 10 \end{vmatrix} = 52$$

$$x_1^* = \frac{|A^1|}{|A|} = 2,\ x_2^* = \frac{|A^2|}{|A|} = 3,\ x_3^* = \frac{|A3|}{|A|} = 1$$

Note: If the it A^{-1} doesn't exists, the problem is likely to have multiple solutions or none at all.

■