Mathematics for Economics

Gianluca Damiani[1]

[1]e-mail: gdamiani@ad.unc.edu. Ph.D. Student in Economics, UNC at Chapel Hill. These notes are mainly my transcription of the lectures I attended of ECON 700 and ECON 701 at UNC, Fall 2023

Contents

Chapter 1

Sets

A set is a collection of elements characterized by having the same property. Defining, for each element, a propositional function $P(x)$, namely a function that can assume only two values (true or false), then we can formally define a set as:

$$C = \{x \in X : P(x)\}$$

Where X is the universal set, i.e., the set of all elements.

Further, a set can be **empty** ($\emptyset$) if it does not contain any element, **non-empty**, or **singleton** if it contains only one element.

An important definition is that of the subset.

Definition 1.0.1 (subset)**.** Given two sets, A, B, we can write that $A \subseteq B$ if for all $x \in A$, then $x \in B$. A proper subset is when $A \subset B$ and $B \not\subset A$. Further, if $A \subseteq B$ and $B \subseteq A$, then $A = B$.

A first result is the following.

Theorem 1.0.1. *Let A, B, C be sets. Then, if $A \subseteq B$ and $B \subseteq C$, then $A \subseteq C$.*

Further, $A \subset C$ it either $A \subset B$ or $B \subset C$ or both.

Proof. To see the first point, by definition, $A \subseteq B$ means that for all $x \in A$, then $x \in B$. $B \subseteq C$ means that for all $x \in B$, then $x \in C$. So $x \in A$ is also in C.

To see the second point, $A \subset C$ means that $C \not\subset A$. By contradiction, assume it is. So we can write $B \subseteq C \subseteq A$ and therefore $B \subseteq A$. But this contradicts $A \subset B$. □

Definition 1.0.2. (equal sets) Given two sets, A, B, they are equal ($A = B$) if and only if they contain the same elements, namely $x \in A \iff x \in B$. If two sets are not equal, they are different.

Definition 1.0.3. (complementary sets) Let X be a set and $A \subseteq X$. Then we can write A^C as the set of elements of X that are not in A.

Assuming X as the set of all elements, or universal sets, we can write:

- $X^C = \emptyset$
- $\emptyset^C = X$
- $(A^C)^C = A$

Definition 1.0.4. The power set of A is the set of all possible subsets of A. A set of subsets is called *family*. The power set is denoted as $\mathcal{P}(A)$.

Notice that $\mathcal{P}(A)$ always has 2^n elements (where n is the number of elements of A).

1.1 Sets: operations

Definition 1.1.1. (Union of sets) Let $A, B \subseteq X$. Then we define $A \cup B$ as the set of all elements that are either in A *or* in B.

Definition 1.1.2. (Intersection of sets) Let $A, B \subseteq X$. Then we define $A \cap B$ as the set of all elements that either in A *and* in B.

If $A \cap B = \emptyset$, then A and B are two **disjoint sets**.

Theorem 1.1.1. *Let A, B be sets. Then:*
1) $A \cup B = A \iff B \subseteq A$.
2) $A \cap B = B \iff B \subseteq A$.
And $\emptyset \subseteq A$ for all $A \subset X$.

Proof. 1) $A \cup B = A \Rightarrow B \subseteq A$. $A \cup B = A$ is the hypothesis. We want to show that, given that, for all $x \in B$, then $x \in A$. Since the equality implies $A \cup B \subseteq A$, then for all $x \in B$, $x \in A$.

$B \subseteq A \Rightarrow A \cup B = A$. $B \subseteq A$ means that for all $x \in B$, then $x \in A$. $A \cup B$ means that either $x \in A$ or $x \in B$. Since for all $x \in B$, then $x \in A$, therefore $x \in A$.

2) $A \cap B = B \Rightarrow B \subseteq A$. $A \cap B$ means that $x \in A$ and $x \in B$. For being equal to B, then, the elements of B must be contained in the elements of A. So we can write $A \cap B \subseteq A$.

$B \subseteq A \Rightarrow A \cap B = B$. B subset of A means that for all $x \in B$, then $x \in A$. Since, $A \cap B$ means $x \in A$ and $x \in B$, then, we have that if $x \in B$, then $x \in A$, implies that the set of x such that x belongs to both A and B is equal to B. □

Theorem 1.1.2. *Let A, B, C be sets. Then the following properties hold:*
1) $A \cap A^C =$. $A \cup A^C = X$
2)$A \cap B = B \cap A$ and $A \cup B = B \cup A$ (commutativity)
3)$A \cap (B \cap C) = (A \cap B) \cap C = A \cap B \cap C$ and $A \cup (B \cup C) = (A \cup B) \cup C = A \cup B \cup C$ (associativity)
4)$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ and $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ (Distributive law)
5) $(A \cap B)^C = A^C \cup B^C$ and $(A \cup B)^C = A^C \cap B^C$ (De Morgan's Laws)

Proof. We prove only 4a) and 5a).

Let's start with 4a). We want to show that $A \cup (B \cap C) \iff (A \cup B) \cap (A \cup C)$. See $A \cup (B \cap C) \Rightarrow (A \cup B) \cap (A \cup C)$. We can have either $x \in A$ or $x \in (B \cap C)$, which means $x \in B$ and $x \in C$. In both cases, we can have $x \in A$ or $x \in B$, or $x \in A$ or $x \in C$. Then, we can write: $A \cup (B \cap C) \subseteq (A \cup B) \cap (A \cup C)$.

Let's see now: $(A \cup B) \cap (A \cup C) \Rightarrow A \cup (B \cap C)$. We have either $x \in A$ or $x \in B$, and $x \in A$ or $x \in C$. If $x \in A$, then we have $x \in A$ also in $A \cup (B \cap C)$. If $x \notin A$, then $x \in B$ and $x \in C$.

Let's see now 5a), $(A \cap B)^C \iff A^C \cup B^C$. $(A \cap B)^C \Rightarrow A^C \cup B^C$. An x that belongs neither to A nor B, then it belongs either to A^C or B^C.

$A^C \cup B^C \Rightarrow (A \cap B)^C$. An x that belongs either to A^C or B^C does not belong to A or B. Therefore, it does not belong to their intersection, and it belongs to the complement of their intersection instead. □

Notice that the properties above can be extended for an infinite number of sets. In that case, we can write $\bigcup_{i=1}^{n} A_i$ and $\bigcap_{i=1}^{n} A_i$. In similar notation, we can write De Morgan's Laws as:

- $\left[\bigcup_{i=1}^{n} A_i\right]^C = \bigcap_{i=1}^{n} A_i^C$
- $\left[\bigcap_{i=1}^{n} A_i\right]^C = \bigcup_{i=1}^{n} A_i^C$

Definition 1.1.3. (Set Difference) Let $A, B \subseteq X$. Then we define $A \setminus B$ as the set difference of A and B, i.e., the set of elements of A that are not in B. This can also be written as $A \cap B^C$.

Notice that $A \setminus B$ makes sense even if $A \not\subset B$.

Theorem 1.1.3. *Let A, B be sets. Then:*
1) $(A \setminus B)^C = A^C \cup B$
2) $B \setminus (B \setminus A) = A \iff A \subseteq B$
3) $B \setminus A = B \iff B \cap A = \emptyset$

Proof. □

Definition 1.1.4. (Ordered Pair) If x, y are any objects, we define (x, y) to be a new object, consisting of x as its first component and y as its second. If $x = x'$ and $y = y'$, then (x, y) and (x', y') are equal

Definition 1.1.5. (Cartesian Product) If A, B are sets, the Cartesian Product $A \times B$ is the collection of ordered pairs such that the first elements belong to A and the second belong to B. We can write:

$$A \times B = \Big\{(x, y) : x \in A \text{ and } y \in B\Big\}$$

1.2 Binary Relations

Given two sets, A and B, a relation between them is any subset R of $A \times B$. In other words, $(a, b) \in R$, and write aRb. If $A = B$, then we have $R \subseteq A^2$.

Definition 1.2.1 (Binary relation)**.** A binary relation R on A can have the following properties:

- Completeness: $\forall a, b \in A$, aRb or bRa (an example of non complete relation is $=$);
- Reflexivity: $\forall a \in A$, then aRa (an example is $\geq$. Instead $>$ is not);
- Symmetry: $\forall a, b \in A$, $aRb \Rightarrow bRa$;
- Transitivity: $\forall a, b, c \in A$, $aRb, bRc \Rightarrow aRc$
- Irreflexivity: $\forall a \in A$, aRa is never true.
- Asymmetry: $\forall a, b \in A$, if aRb, then bRa is not true.
- Antisymmetry: $\forall a, b \in A$, if aRb and bRa, then $a = b$
- Negative transitivity: $\forall a, b, c \in A$, $\neg aRb, \neg bRc \Rightarrow \neg aRc$.

Proposition 1. *1. An asymmetric relation is irreflexive*

2. A transitive and irreflexive relation is asymmetric

3. An asymmetric relation is antisymmetric

4. A antisymmetric and irreflexive relation is asymmetric

Proof. 1. Assume an asymmetric R is reflexive. Then we have aRa, and aRa is not true. A contradiction.

2. Suppose aRb, with R transitive and irreflexivee. If antisymmetry does not hold, we have bRa, and then, by transitivity $aRbRa$, which is a contradiction.

3. Notice that asymmetry implies antisymmetry, since the first holds for all $a, b \in A$, comprised, but not only, the case where $b \neq a$

4. Suppose R is antisymmetric and irreflexive. Take aRb. If $a \neq b$, then bRa is untrue. If $a = b$, then bRa is untrue (by irreflexivity).

□

Definition 1.2.2. A relation R on S is an equivalence relation if it is reflexive, symmetric, and transitive. Then, let $R \subseteq A^2$ be an equivalence relation on A, and x be an element of A. The equivalence class of x with respect to A is defined as $E_x = \{y \in A : yRx\}$

Definition 1.2.3. Let A be a subset of X. A partition of A is a collection $\mathcal{A}$ of non-empty sunsets of A such that:

- each $x \in A$ belongs to some subsets of A
- for all $S, T \in \mathcal{A}$, if $S \neq T$, then $S \cap T = \emptyset$

Definition 1.2.4. A binary relation is called an **order** if it is reflexive, transitive, and antisymmetric. It is called a **strict order** if it is irreflexive, transitive, and antisymmetric. An order that is also complete is called **complete order**.

An example of order is $\geq$. Instead, $>$ is a strict order (but not an order: indeed, it is not reflexive).

Some orders may not be complete. Then, they are defined **partial orders**.

Definition 1.2.5. An ordered set is a set S on which an order is defined.

Chapter 2

Numbers

2.1 Natural Numbers

Let's start with the **natural numbers**. The set of natural numbers is:

$$\mathbb{N} = \{1, 2, \dots\}$$

Definition 2.1.1. A prime number is a natural number greater than 1 with no positive divisor other than 1 and itself.

2.2 Integers and Rational Numbers

Definition 2.2.1. The set of **integers** is defined as:

$$\mathbb{Z} = \{x = a - b : \text{for some} \quad a, b \in \mathbb{N}\}$$

It is obvious that we can write $\mathbb{Z}_+ = \mathbb{N}$, and therefore $\mathbb{Z} = \mathbb{Z}_+ \cup \mathbb{Z}_- \cup \emptyset$.

Definition 2.2.2. Rational Numbers The set of rational numbers is defined as:

$$\mathbb{Q} = \left\{ \frac{m}{n} : m, n \in \mathbb{Z} \quad \text{and} \quad b \neq 0 \right\}$$

Obviously, we have the following:

$$\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q}$$

However, the set of rational numbers has some gaps. For instance, we know, by Pythagorean Theorem, that the length of the hypotenuse of a right triangle with both sides equal to 1 is $\sqrt{2}$. But this is not a rational number.

Theorem 2.2.1. *$\sqrt{2}$ is not a rational number.*

Proof. . If $\sqrt{2}$ is a rational number, then it can be written as $\frac{m}{n}$, where $m, n \in mathbbZ$ and they are not both even. Therefore, we can write $\sqrt{2} = \frac{m}{n}$. Squaring both sides, we have $2 = (\frac{m}{n})^2$, i.e. $2 = \frac{m^2}{n^2}$. Then we can write:

$$m^2 = 2n^2$$

Therefore m^2 is even, and m is even. Since m is even, we can write it as $2r$, therefore, substituting above, we have:

$$(2r)^2 = 2n^2$$

Dividing both sides by 2, we have:

$$2r^2 = n^2$$

And, as above, n must be even. But this contradicts the assumption of m, n not being both even. □

2.3 Real Numbers

Since rational numbers have gaps (not only $\sqrt{2}$, we need to construct a system that is richer enough such that any subset is not empty. To do this, we start with a series of axioms, defining a **Field**.

Definition 2.3.1. A field is a set F endowed with two operations, addition, and multiplication, which satisfy the following axioms.

For addition:

1. $\forall x, y \in F, x + y \in F$
2. $\forall x, y \in F, x + y = y + x$
3. $\forall x, y, z \in F, x + (y + z) = (x + y) + z$
4. It exists a $0 \in F$ such that $x + 0 = x, \forall x \in F$
5. $\forall x \in F$ it exists a $(-x)$ such that $(-x) + x = 0$

For multiplication:

1. $\forall x, y \in F, x \cdot y \in F$
2. $\forall x, y \in F, x \cdot y = y \cdot x$
3. $\forall x, y, z \in F, (x \cdot y) \cdot z = x \cdot (y \cdot z)$
4. It exists an $1 \in F$ such that $x \cdot 1 = x, \forall x \in F$

5. It exists an $x \in F, x \neq 0, \forall x \in F$ such that $\frac{1}{x} \cdot x = 1$

Besides, for both, it holds:

- $\forall x, y \in F, (x + y) \cdot z = z \cdot x + z \cdot y$

It is obvious that $\mathbb{Z}$ is not a field.

Definition 2.3.2. An **Ordered Field** is a field F on which we can define an order $\geq$ such that:

- $\forall x, y, z \in F$, if $x \leq y$, then $x + z \leq y + z$.
- $\forall x, y, z \in F$, if $x \leq y$, and $z \geq 0$,, then $x \cdot y \leq y \cdot z$

A further definition is required before defining the set of real numbers.

Definition 2.3.3. Suppose that $A, B \subseteq \mathcal{R}$ satisfy:

$$a \leq b, \forall a \in A, b \in B$$

Then, it exists a $c \in \mathcal{R}$ such that:

$$a \leq c \leq b$$

For all $a \in A, b \in B$

This is called the **Completeness axiom**.
To see that $\mathbb{Q}$ does not satisfy it, let's take two subsets of it:

$$A = \left\{ x \in \mathbb{Q} : x < \sqrt{2} \right\}$$

and

$$B = \left\{ x \in \mathbb{Q} : x > \sqrt{2} \right\}$$

But since $\sqrt{2} \notin \mathbb{Q}$, then $\mathbb{Q}$ is not complete.
Therefore we can define $\mathbb{R}$.

Definition 2.3.4. The set $\mathbb{R}$ is an ordered field that satisfies the axiom of Completeness.

If a set is ordered, then we can define some special elements in it, or outside it. These are the **Upper Bound, Lower Bound, Greatest Lower Bound (Infimum), Lower Upper Bound (Supremum)**.

Definition 2.3.5. Let S be ordered and $E \subseteq S$. Then:

- If there exists $\beta \in S$ such that $\forall x \in E, x \leq \beta$, this is called **Upper Bound**.
- If there exists $\alpha \in S$ such that $\forall x \in E, x \geq \alpha$, this is called **Lower Bound**

- If E has both an upper bound and a lower bound, then it said to be **bounded**
- The smallest upper bound is called **Supremum**: $\beta = \sup E$
- The greatest lower bound is called **Infimum**: $\alpha = \inf E$

Notice that it is not required that the supremum and the infimum are elements of E. But if they exist, they are unique (recall, we are in an ordered set, and an order is, by definition, antisymmetric).

Definition 2.3.6. An ordered set S it is said to have the **Supremum Property** if:

- Every bounded above non-empty subset has a Supremum
- Every bounded below non-empty subset has an Infimum

Theorem 2.3.1. *The Supremum Property and Completeness are equivalent. That is: Supremum Property* $\iff$ *Completeness*

Proof. Let's start with: Completeness $\Rightarrow$ Supremum Property.

Take $E \subseteq \mathbb{R}$, E is bounded and non-empty. We can denote U as the set of all upper bounds. Since E is bounded, $U \neq \emptyset$. Then, we know that $\forall x \in E$, then we have at least one $u \in U$ such that $x \leq u$. By completeness, we have a c such that, $x \leq c \leq u$, for all $x \in E, u \in U$. Then c is a supremum of E. The same case for lower bounds.

Let's see now: Supremum property $\Rightarrow$ Completeness. Take $A, B \subseteq \mathbb{R}$ such that $x \leq y, \forall x \in A, y \in B$. Then, A is bounded above, and B is bounded below. By the supremum property, $\alpha = \sup A$ and $\beta = \inf B$ exist. Since a least upper bound for A exists, we can write:

$$x \leq \alpha \leq y, \forall x \in A, y \in B$$

So completeness holds. □

Definition 2.3.7. If the least upper bound belongs to the set $E \subseteq \mathbb{R}$, then it is called **maximum**. Similarly, if the greatest lower bound belongs to $E \subseteq \mathbb{R}$, it is called **minimum**.

Again, by antisymmetry, if a **maximum(minimum)** exists, it isn unique.

Theorem 2.3.2 (Archimedean Property)**.** *The set* $\mathbb{N}$ *is not bounded above in* $\mathbb{R}$*. Equivalently, we can say:*

1. *For each* $z \in \mathbb{R}$*, it exists an* $n \in \mathbb{N}$ *such that* $n > z$
2. *For each* $x \in \mathbb{R}_{++}, y \in \mathbb{R}$*, there exists an* $n \in \mathbb{N}$ *such that* $ny > z$
3. *For each* $x \in \mathbb{R}_{++}$*, there exists an* $n \in \mathbb{N}$ *such that* $0 < \frac{1}{n} < x$*.*

Proof. Let's prove only the main proposition. By contradiction, suppose $\mathbb{N}$ is bounded above. Then, by supremum property ($\mathbb{N}$ is not empty), we have an $\alpha = \sup\mathbb{N}$. Since α is a supremum, then $\alpha - 1$ is not, and then we can write:

$$\alpha - 1 < n_0$$

For at least one $n_0 \in \mathbb{N}$. Then, we can write:

$$\alpha < n_0 + 1$$

Since $n_0 + 1 \in \mathbb{N}$. But this contradicts α being a supremum. □

Notationally, we can introduce upper and lower bounds for sets that are not bounded. These are given by $-\infty, +\infty$. Still, these are **not real numbers**.

However, the following conventions are customary:

- $x + \infty = \infty$
- $\frac{x}{\pm\infty} = 0$
- $x \cdot \infty = \infty$, $x \cdot -\infty = -\infty$
- $-x(\infty) = -\infty$, $-x \cdot (-\infty) = \infty$

We can also define the subsets of $\mathbb{R}$ (intervals) as follows:

- $[a, b) = \{x \in \mathbb{R} : a \leq x < b\}$ This interval is **half-closed**
- $[a, b] = \{x \in \mathbb{R} : a \leq x \leq b\}$ This interval is **closed**
- $(a, b) = \{x \in \mathbb{R} : a < x < b\}$ This interval is **open**
- $(a, b] = \{x \in \mathbb{R} : a < x \leq b\}$ This interval is **half-closed**

If the interval is unbounded, then we can write:

- $[a, \infty) = \{x \in \mathbb{R} : a \leq x\}$
- $(a, \infty) = \{x \in \mathbb{R} : a < x\}$
- $(-\infty, b) = \{x \in \mathbb{R} : b > x\}$
- $(-\infty, b] = \{x \in \mathbb{R} : b \geq x\}$

Finally, an important concept is that of **Absolute Value**.

Definition 2.3.8. If $x \in \mathbb{R}$, then the absolute value of x, denoted as $|x|$ is:

$$|x| = \begin{cases} x & \text{if } x \geq 0 \\ -x & \text{otherwise} \end{cases}$$

Theorem 2.3.3. *Let $x, y \in \mathbb{R}$ and $a \geq 0$. Then:*

1. $|x| \geq 0$
2. $|x| \leq a$ *if and only if* $-a \leq x \leq a$
3. $|x\dot{y}| = |x| \cdot |y|$
4. $|x + y| = |x| + |y|$
5. $||x| - |y|| \leq |x - y|$
6. $|x - y| < c \Rightarrow |x| \leq |y| + c$

Proof. □

Chapter 3

Functions

Definition 3.0.1. Let A and B be sets. A function between A and B is a relation $f \subseteq A \times B$, non-empty, such that it $(a, b) \in f$ and $(a, b') \in f$, then $a = a'$.

We can also define the **Domain** as the set of all first elements of f, and the **Range** as the set of all the second elements of f. The set B, which contains all the second elements, is called **Codomain**

A function can be written as:

$$f : A \to B$$

If we allow for f be such that $(a, b) \in f$ and $(a, b') \in f$ and $b \neq b'$, this is called **Correspondence**, and can be written as:

$$f : A \rightrightarrows B$$

Notice that for each correspondence, one can always define a function $g : A \to 2^B$.

Definition 3.0.2. The **graph** of a function is defined as:

$$G = \Big\{(x, y) \in A \times B : y \in f(x)\Big\}$$

Notice that this is the graph of a correspondence. For a function (that is, a special case of correspondence when the set of values is a singleton), we write the graph as:

$$G = \Big\{(x, y) \in A \times B : y = f(x)\Big\}$$

Definition 3.0.3. Consider a function. Then say that x^* is a **Fixed Point** if:

$$x^* = f(x^*)$$

A function can be **surjective**, **injective** or **bijective**.

Definition 3.0.4. A function $f : A \to B$ is **surjective** (onto) if every element of B is mapped by at least one element if A. In other words, if B is equal to the **range** of f.

A function $f : A \to B$ is **injective** (one-to-one) if for all elements $a, a' \in A$, $f(a) = f(a')$ implies that $a = a'$. In other words, if each element in the codomain is mapped by at most one element in the domain

Finally, a function $f : A \to B$ is **biijective** if it is both surjective and injective.

An example of a function that is not surjective is $y = x^2$ when defined on $\mathbb{R}^2$ because no matter what values x takes, y can never be negative. If defined only on $\mathbb{R}_+$, then it becomes surjective and injective, and therefore bijective.

A function can also act on sets. In this case we talk of **Image** and **Pre-image**

Definition 3.0.5. Suppose that $f : A \to B$, and $C \subseteq A$, we define the **image** of C in B as:

$$f(C) = \Big\{ f(x) \in B : x \in C \Big\} \subseteq B$$

And, for $D \subseteq B$ the **pre-image** of D on A as:

$$f^{-1}(D) = \Big\{ x \in A : f(x) \in D \Big\} \subseteq A$$

Take for example $y = x^2$, defined on $\mathbb{R}$. Then, if we take a subset of $\mathbb{R}$, as the interval $[0, 2)$, we can write the image as follows:

$$f([0, 2)) = \Big\{ f(x) : x \in [0, 2) \Big\} = [0, 4)$$

Taking a subset $\{1\}$, we can write the preimage as:

$$f^{-1}(\{1\}) = \Big\{ x \in A : x \in \{-1\} \Big\} = \emptyset$$

Theorem 3.0.1. *Suppose that $f : A \to B$, and $C \subseteq A$. Then:*

$$C \subseteq f^{-1}[f(C)]$$

Proof. $forall x' \in C$, we have $f(x') \in f(C)$ and $f^{-1}(f(C)) = \{x \in A : f(x) \in f(C)\}$. Then $x' \in f^{-1}(f(C))$ □

The opposite is true if f is one-to-one.

Theorem 3.0.2. *Suppose $f : A \to B$, and $D \subseteq B$ then:*

$$f[f^{-1}(D)] \subset D$$

Proof. For all $y \in f(f^{-1}(D))$, it exists an $x \in f^{-1}(D)$ such that $y = f(x)$ and $f^{-1}(D) = \{x \in A : f(x) \in D\}$. So $y \in D$. □

If f is one-to-one, also the opposite direction is true. So, we have $f[f^{-1}(D)] = D$

Theorem 3.0.3. *Suppose that $f : A \to B$, and $C_1, C_2 \subseteq A$, then:*

$$f(C_1 \cup C_2) \subseteq f(C_1) \cup f(C_2)$$

Proof. Suppose $y \in f(C_1 \cup C_2)$. Then, it exists $x \in C_1 \cup C_2$ such that $y = f(x)$, so $x \in C_1$ and $x \in C_2$, and $y \in f(C_1)$ and $y \in f(C_2)$. □

If f is one-to-one, also the opposite direction is true. So, $f(C_1 \cup C_2) = f(C_1) \cup f(C_2)$

Given two functions $f : A \to B$ and $g : B \to C$, then, for any $a \in A$, we have $f(a) \in B$. B is the domain of g, and then $g(f(a)) \in C$. This function is called **composition**, and denoted as:

$$(g \circ f)(a) = g[f(a)]$$

Theorem 3.0.4. *Let $f : A \to B$, $g : B \to C$ two biijective functions. Then also, the composition is bijective*

Proof. Since g is bijective, then it is surjective. This means that for every $c \in C$, there exists at most one $f(a) \in B$ such that $g[f(a)] \in C$. Since f is surjective too, then it exists at most one a such that $f(a) \in B$. Suppose it is not. That is, it exists $a \neq a'$ such that $(g \circ f)(a) = (g \circ f)(a')$. Then $f(a) = b \neq f(a') = b'$ and $g(b) \neq g(b')$. This is a contradiction. □

Given a bijection $f : A \to B$, then to each y corresponds one and only one $x \in A$. Then we can define a function B into A, called the **inverse**.

Definition 3.0.6. Let $f : A \to B$ be bijective. The inverse function $f^{-1} : B \to A$ is given by:

$$f^{-1} = \Big\{(y, x) \in B \times A : (x, y) \in f\Big\}$$

Theorem 3.0.5. *Let $f : A \to B, g : B \to C$ be two bijective functions. Then $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$*

Proof. Since $(g \circ f)^{-1}$ is the composition of two invertible functions, it is invertible too. We can write:

$$(g \circ f) = \Big\{(a, c) : \text{it exists } \ b \in B \ \text{ s.t. } (a, b) \in f^{-1}, (b, c) \in g^{-1}\Big\}$$

Then:

$$(g \circ f)^{-1} = \Big\{(c, a) : \text{it exists } \ b \in B \ \text{ s.t. } (b, a) \in f^{-1}, (c, b) \in g^{-1}\Big\} =$$
$$= f^{-1} \circ g^{-1}$$

□

Chapter 4

Cardinality

To compare the relative sizes of sets, the concept of cardinality has been defined.

Two sets can be of the same size if they are **equinumerous**.

Definition 4.0.1. Two sets S and T are equinumerous, $S \sim T$ if it exists a bijective function $f : S \to T$

For instance, the set $\{1, 2, 3, \ldots, 50\}$ is equinumerous to the set $\{1, 4, 9, \ldots, 2500\}$ for the function $f(x) = x^2$.

Definition 4.0.2. A set S is called **finite** if $S \neq \emptyset$, or, there is a bijection $f : \{1, 2, \ldots, n\} \to S$ and an $n \in \mathbb{N}$.

A set that is not finite is **infinite**.

If there is a bijection $f : \mathbb{N} \to S$, the set is denumerable. We can write $N \sim \mathbb{N}$.

If a set is finite or innumerable, it is called **countable**. If not, it is called **uncountable**.

Theorem 4.0.1. *The set of even numbers and the set of odd numbers are denumerable*

Proof. For the even numbers, we can find the function $f(n) = 2n$. For the odd numbers, the function $f(n) = 2n + 1$. □

Theorem 4.0.2. $\mathbb{Z}$ *is countable.*

Proof. We can find the following function that maps each element of $\mathbb{N}$ into an element of $\mathbb{Z}$:

$$f(n) = \begin{cases} \frac{(n-1)}{2} & \text{if } n \text{ is odd} \\ \frac{-n}{2} & \text{if } n \text{ is even} \end{cases}$$

□

Theorem 4.0.3. *Suppose T and S are both countable. Then $S \cup T$ is also countable.*

Definition 4.0.3. Since T and S are countable, there must be be $f : \mathbb{N} \to T$ and $g : \mathbb{N} \to T$. Then we can have $h : \mathbb{N} \to S \cup T$ as:

$$h(n) = \begin{cases} f(\frac{n+1}{2} \text{if odd} \\ g(\frac{n}{2} \text{if even} \end{cases}$$

Further, if S_i is countable, then $\bigcup_{i=1}^{n} S_i$ is also countable.

Theorem 4.0.4. *Let $S_1, S_2, \ldots, S_n$ be nonempty countable sets. The Cartesian Product $\bigtimes_{i=1}^{n} S_i$ is also countable.*

Proof. □

Theorem 4.0.5. *The denumerable union of denumerable sets is denumerable.*

Proof. □

Theorem 4.0.6. $\mathbb{Q}$ *is countable.*

Proof. Define $S_p = \{\frac{p}{q}, \forall p \in \mathbb{Z}\}$ for every $q \in \mathbb{N}$. This set is countable because it is equinumerous to $\mathbb{Z}$. If we write $\mathbb{Q} = \bigcup_{q \in \mathbb{N}} S_q$, then it is a union of denumerable sets, so it is denumerable and countable. □

Theorem 4.0.7. $\mathbb{R}$ *is uncountable.*

Proof. Let's focus on the interval $(0, 1)$. Suppose it is countable. Then, any element $0 \leq x \leq 1$, can be written as follows:

$$x_1 = 0.a_{11}a_{12}a_{13}\ldots$$
$$x_2 = 0.a_{21}a_{22}a_{23}\ldots$$
$$x_3 = 0.a_{31}a_{32}a_{33}\ldots$$

We can construct a real number $0.b_1b_2b_3\ldots$ such that $b_i \neq a_ii$ for every i. This number belongs to $(0, 1)$, but it is not equal to any x_n because they have different n^{th} decimal digits. □

Chapter 5

Some elements of Linear Algebra

The fundamental objects of Linear Algebra are **vector spaces**, also called **linear spaces**. We can denote them by V. An example is $\mathbb{R}^n$.

The elements of Vector spaces are called **vectors**, and they can be summed or multiplied by an $a \in \mathbb{R}$. Then, V is said to be closed under vector addition and scalar multiplication.

Definition 5.0.1. A vector space is a set V closed under vector addition and scalar multiplication and obeying the following rules:

- $\mathbf{v} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$
- $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$
- $\mathbf{v} + 0 = \mathbf{v}$
- $\mathbf{v} + -\mathbf{v} = 0$
- $\forall a \in \mathbb{R}, a(b \cdot \mathbf{v}) = (a \cdot b)\mathbf{v}$
- $1 \cdot \mathbf{v} = \mathbf{v}$
- $a \cdot (\mathbf{u} + \mathbf{v}) = a \cdot \mathbf{u} + a \cdot \mathbf{v}$
- $(a + b)\mathbf{u} = a \cdot \mathbf{u} + b \cdot \mathbf{v}$

Another important definition is that of **subspaces**.

Definition 5.0.2. Let V be a vector space, and $W \subset V$. W is said to be a **subspace** of V if:

- taking $\mathbf{v}, \mathbf{w} \in W$, then $\mathbf{v} + \mathbf{w} \in W$
- $\mathbf{v} \in W, \alpha \in \mathbb{R}$, then $\alpha \cdot \mathbf{v} \in W$.

It is easy to see that the smallest possible vector subspace is that made up of only vector $\mathbf{0}$. Indeed, $\mathbf{0} + \mathbf{0} = \mathbf{0}$ and any scalar times $\mathbf{0}$ is equal to $\mathbf{0}$.

Definition 5.0.3. Let $v_1, \ldots, v_n$ be n elements in V and $a \in \mathbb{R}$. Then we define the vector:

$$\alpha_1 \cdot \mathbf{v_1} + \cdots + \alpha_n \cdot \mathbf{v}_n$$

the **Linear Combination** of $v_1, \ldots, v_n$.

Proposition 2. *Let W be the set of all linear combinations of $\mathbf{v}_1, \ldots, \mathbf{v}_n \in V$. Then, W is a subspace of V, and it is called the subspace generated by $\mathbf{v}_1, \ldots, \mathbf{v}_n$*

Proof. Take $\mathbf{w}_1$ and $\mathbf{w}_2 \in W$, and write them as:

$$\mathbf{w}_1 = \alpha_1 \cdot \mathbf{v}_1 + \cdots + \alpha_n \cdot \mathbf{v}_n$$

$$\mathbf{w}_2 = \beta_1 \cdot \mathbf{v}_1 + \cdots + \beta_n \cdot \mathbf{v}_n$$

Then we can see that:

$$\mathbf{w}_1 + \mathbf{w}_2 = (\alpha_1 + \beta_1)\mathbf{v_1} + \cdots + (\alpha_n + \beta_n)\mathbf{v_n} \in W$$

$$k \cdot \mathbf{w_1} = k\alpha_1 \mathbf{v}_1 + \cdots + k\alpha_n \mathbf{v}_n$$

□

Definition 5.0.4. $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are **linearly dependent** if it exists $(\alpha_1, \ldots, \alpha_n) \neq (0, \ldots, 0)$ such that:

$$\alpha_1 \cdot \mathbf{v}_1 + \cdots + \alpha_n \cdot \mathbf{v}_n = 0$$

If linearly dependent, we can write a vector $\mathbf{v}_i$ as:

$$\mathbf{v}_i = -\sum_{i \neq j}^{n} \frac{a_i}{a_j} \mathbf{v}_j$$

for some i.

If they are not linearly dependent, then $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are **linearly independent**.

Definition 5.0.5. If $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are linearly independent and the set of all linear combinations of $\mathbf{v}_1, \ldots, \mathbf{v}_n$ is equal to V, (that is, they **generate** V), then $\mathbf{v}_1, \ldots, \mathbf{v}_n$ is a **basis** for V.

The most simple example of basis is the standard basis, namely, in R^2, the set of vectors $(1, 0), (0, 1)$, in R^3 $(1, 0, 0), (0, 1, 0), (0, 0, 1)$ and so on. It is apparent that there are infinitely many bases and that any set of 3 linearly independent vectors is a basis.

Proposition 3. *Let V be a vector space and $\mathbf{v}_1, \ldots, \mathbf{v}_n$ a basis. Then, each vector can be written as a unique linear combination of the basis. Given a set of $(\alpha_1, dots, \alpha_2)$, there is a unique vector that can be written as:*

$$v = \alpha_1 \cdot \mathbf{v}_1 + \cdots + \alpha_n \cdot \mathbf{v}_n$$

Proof. To prove that, at most, one vector can be written as a linear combination of the basis for each $(\alpha_1, dots, \alpha_n)$, suppose it is not. Then we have:

$$v = \alpha_1 \cdot \mathbf{v}_1 + \cdots + \alpha_n \cdot \mathbf{v}_n$$

$$v = \beta_1 \cdot \mathbf{v}_1 + \cdots + \beta_n \cdot \mathbf{v}_n$$

Combining together, we have:

$$(\alpha_1 - \beta_1)\mathbf{v}_1 + \cdots + (\alpha_n - \beta_n)\mathbf{v}_n = 0$$

If $\alpha \neq \beta$, then $\mathbf{v}_1, \ldots, \mathbf{v}_n$ is a set of linearly dependent vectors, thus, the definition of basis is violated. □

For example, in $\mathbb{R}^2$, take the vector $(2, 3)$, it can be written as a linear combination of the basis $((1, 0), (0, 1))$.

$$\begin{bmatrix}2\\3\end{bmatrix} = 2 \cdot \begin{bmatrix}1\\0\end{bmatrix} + 3 \cdot \begin{bmatrix}0\\1\end{bmatrix}$$

Then, we call $\{2, 3\}$ the coordinates of the vector $(2, 3)$.

As seen, a vector space does not have a unique basis. Still, the number of vectors in the basis is unique. This number is called the **dimension** of the vector space.

Proposition 4. *Let $(\mathbf{v}_1, \ldots, \mathbf{v}_n)$ be a basis of V and $(\mathbf{w}_1, \ldots, \mathbf{w}_m)$ a set of vectors, with $m > n$. Then $(\mathbf{w}_1, \ldots, \mathbf{w}_m)$ is linearly independent*

Proof. □

Space $\{\mathbf{0}\}$ has no basis. Therefore, his dimension is zero.

5.1 Matrices

A $(m \times n)$ matrix is an array of numbers:

$$A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix}$$

A vector is a special case of a matrix. Further, we can write a $(m \times n)$ matrix as an array of row vectors $A_i = (a_{i1}, a_{i2}, \ldots, a_{in})$, or column vectors $A^j = (a_{1j}, a_{2j}, \ldots, a_{mj}$

Two matrices $(m \times n)$ can be added: $A + b = C$, elementwise ($(a_{ij} + b_{ij}$ and so on).

Matrix addition is associative and commutative:

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

$$A + B = B + A$$

Further, each matrix can be multiplied by a scalar, $a \in \mathbb{R}$.

Instead, to multiply two matrices, they must be conformable; namely, $(m \times n)$ can be multiplied only by a $(n \times m)$ matrix, and the result is a $(m \times m)$ matrix. The ij-entry of this new matrix is given by the inner product of the A_i column and the A^j row:

$$A_i \cdot A^j = \sum_{k=1}^{n} a_{ik} a_{kj} = a_{i1} a_{1j} + \ldots a_{in} a_{nj}$$

Matrix Multiplication is still associative:

$$(A \cdot B) \cdot C = A \cdot (B \cdot C)$$

But it is not commutative, in general.

Notice, finally, that the product of an $(1 \times n)$ matrix and a $(n \times 1)$ matrix is equivalent to vector scalar multiplication $x^T y$.

Each matrix $(n \times m)A$ has a **transpose**, a matrix $(m \times n)$ denoted A^T. An example:

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}$$

$$A^T = \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \\ a_{13} & a_{32} \end{bmatrix}$$

Notice that $(A \cdot B)^T = B^T \cdot A^T$.

Definition 5.1.1. A matrix is **symmetric** is $A = A^T$

Obviously, a symmetric matrix must be $(n \times n)$, that is a **square matrix**

Definition 5.1.2. A square matrix is **diagonal** if all off the diagonal elements are equal to zero.

Definition 5.1.3. A square matrix is **upper triangolar** if all entries below the main diagonal are equal to zero.

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{bmatrix}$$

A particular diagonal matrix is the **identity** matrix:

$$I = \begin{bmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \\ 0 & 0 & \ldots & 1 \end{bmatrix}$$

Each matrix, or vector, multiplied by I is unchanged.

One of the most important concepts in matrix algebra is that of the **inverse** of a matrix.

Definition 5.1.4. A is invertible (or non-singular) if there exists a B such that:

$$A \cdot B = I$$

and

$$B \cdot A = I$$

Proposition 5. *The inverse, if it exists, is unique.*

Proof. Suppose $B \cdot A = I$ and $C \cdot A = I$. Then we can write:

$$B \cdot A = C \cdot A$$

Then, $B = C$. □

The inverse is denoted as A^{-1}.

Only for the (2×2) case we can find the inverse as follows:

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{a \cdot d - c \cdot b} \cdot \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

if $(a \cdot d - c \cdot b) \neq 0$.

In the case of a diagonal matrix, instead, the inverse is:

$$A = \begin{bmatrix} d_1 & 0 & 0 & \dots & 0 \\ 0 & d_2 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & & \\ 0 & 0 & \dots & 0 & d_n \end{bmatrix} \Rightarrow A^{-1} = \begin{bmatrix} \frac{1}{d_1} & 0 & 0 & \dots & 0 \\ 0 & \frac{1}{d_2} & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & & \\ 0 & 0 & \dots & 0 & \frac{1}{d_n} \end{bmatrix}$$

These are the only cases where the inverse of a matrix is easy to find. Most of the time, it requires some specific algorithm, which can be (relatively) easy or considerably most difficult.

Proposition 6. *If A and B are invertible, then also their product is invertible, and:*

$$(A \cdot B)^{-1} = B^{-1} \cdot A^{-1}$$

Proof. To see this, left-multiply both sides by $(A \cdot B)$. Then, we have:

$$\begin{aligned} &(A \cdot B) \cdot (A \cdot B)^{-1} = (A \cdot B) \cdot (B^{-1} \cdot A^{-1}) \\ &I = A \cdot \underbrace{(B \cdot B^{-1})}_{\text{by associativity}} \cdot A^{-1} \\ &I = A \cdot I \cdot A^{-1} = \\ &I = A \cdot A^{-1} \\ &I = I \end{aligned}$$

□

5.2 Linear Equations

Chapter 6

Sequences

6.1 Metric Spaces

Given 2 elements in a set, a natural operation is that of looking at how close they are. This is done through a function d, called **metric**.

Definition 6.1.1. Let X be a non-empty set. The function $d : X \times X \to \mathbb{R}$ is called **distance** if it satisfies the following conditions:

- $d(x, y) \geq 0$ and $d(x, y) = 0$ if and only if $x = y$
- $d(x, y) = d(y, x)$
- $d(x, z) \leq d(x, y) + d(y, z)$ The pair (X, d) is called **metric space**.

The most used (although not the only one) metric space in economics is the **Euclidean Space**, $\mathbb{R}^n$, where $x = (x_1, x_2, \ldots, x_n)$. In these spaces, the metric is defined by:

$$d_E(x, y) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2} = \|x - y\|$$

The function $\|\| : \mathbb{R}^n \to \mathbb{R}_+$ is called **Euclidean Norm**.

Proposition 7 (Cauchy-Schwarz Inequality)**.** *If $a_1, \ldots, a_n$ and $b_1, \ldots, b_n$ are arbitrary real numbers, then:*

$$\Big(\sum_{k=1}^{n} a_k b_k \Big)^2 \leq \sum_{k=1}^{n} a_k^2 \cdot \sum_{a=1}^{n} b_k^2$$

In vector notation:

$$|a \cdot b| \leq \|a\| \cdot \|b\|$$

Proof. For each $t \in \mathbb{R}$, we can write):

$$\sum_{k=1}^{n}(ta_k + b_k)^2 \geq 0$$

Since it is a sum of squared terms. Solving the square of the binomial, we have:

$$\sum_{k=1}^{n}(t^2a_k^2 + 2ta_kb_k + b_k^2) =$$
$$t^2\underbrace{\sum_{k=1}^{n}a_k^2}_{A} + 2t\underbrace{\sum_{k=1}^{n}a_kb_k}_{B} + \underbrace{\sum_{k=1}^{n}b_k^2}_{C} =$$
$$t^2A + 2tB + C$$

This is a quadratic inequality in t. Since it is non-negative $\forall t$, we must look at the discriminant $b^2 - 4ac$, that is:

$$(2B)^2 - 4AC \geq 0$$
$$B^2 \geq AC$$

Substituting, we have:

$$\Big(\sum_{k=1}^{n}a_kb_k\Big)^2 \leq \sum_{k=1}^{n}a_k^2 \cdot \sum_{k=1}^{n}b_k^2$$

□

Proposition 8. *The Euclidean Norm is a metric for the Euclidean Space*

Proof. To see this, we must check if $\|\cdot\|$ satisfies the three conditions.

- It is obviously greater than zero because it is the square root of the sum of squares. Also, it is equal to zero if and only if $x = y$
- Since it is squared, symmetry is satisfied.
- Let's check for triangle inequality. Take any $x, y \in \mathbb{R}^n$. Then we can write (notice

that we raise square to get rid of the square root:

$$\|a+b\|^2 = \sum_{i=1}^{n}(a_i+b_i)^2 =$$
$$\sum_{i=1}^{n} a_i^2 + \sum_{i=1}^{n} b_i^2 + 2\sum_{i=1}^{n} a_i b_i \leq$$
$$\leq \sum_{i=1}^{n} a_i^2 + \sum_{i=1}^{n} b_i^2 + 2\underbrace{\sqrt{\big(\sum_{i=1}^{n} a_i^2\big)\cdot\big(\sum_{i=1}^{n} b_i^2\big)}}_{\text{by Cauchy-Schwarz Inequality}} =$$
$$\Big(\sqrt{\sum_{i=1}^{n} a_i^2} + \sqrt{\sum_{i=1}^{n} b_i^2}\Big)^2 = (\|a\|+\|b\|)^2$$

But notice that:

$$\|a+b\|^2 \leq (\|a\|+\|b\|)^2$$

Can also be written as:

$$\sqrt{\sum_{i=1}^{n}(a_i+b_i)^2} \leq \sqrt{\sum_{i=1}^{n} a_i^2} + \sqrt{\sum_{i=1}^{n} b_i^2}$$

Assuming $x-y=a, y-z=b$, so $a+b=x-y+y-z=x-z$, we can write:

$$\underbrace{\sqrt{\sum_{i=1}^{n}(x_i-z_i)^2}}_{d_E(x,z)} \leq \underbrace{\sqrt{\sum_{i=1}^{n}(x_i-y_i)^2}}_{d_E(x,y)} + \underbrace{\sqrt{\sum_{i=1}^{n}(y_i-z_i)^2}}_{d_E(y,z)}$$

□

Given a metric space, we can consider a ϵ-**neighborood**.

Definition 6.1.2. Consider a metric space (X,d) and $\epsilon > 0$. Fix an $x \in X$, then we define the ϵ-**neighborood** as the set $N_\epsilon(x)$:

$$N_\epsilon(x) = \big\{y \in X : d(x,y) < \epsilon\big\}$$

6.2 Sequences

Definition 6.2.1. A sequence is a function whose domain is $\mathbb{N}$.

In terms of notation, we can write $\{x_n\}_{i=1}^{\infty}$, or just $\{x_n\}$, where x_n is the $n^{th}-$term of the sequence.

Definition 6.2.2. A sequence $\{x_n\}$ in a metric space (X, d) is said to **converge** to if there exists an $s \in X$ such that, for all $\epsilon > 0$, it exists an $m \in \mathbb{N}$ such that $s_n \in N_\epsilon(s)$ for $n \geq m$

Then, s is the **limit** of the sequence, and we can write $\{s_n\} \to s$.

Another way of writing it (if the sequence is a subset of $\mathbb{R}$, and then the absolute value is a metric on $\mathbb{R}$) is:

$$|s_n - s| < \epsilon$$

If a sequence does not converge, then it **diverges**.

The simplest example of converging sequence in $(\mathbb{R}, |\cdot|)$ is $\{s_n\} = \{\frac{1}{n}\}$. This converges to 0 as $n \to \infty$. Still, notice that if a sequence converges or not, it depends on the space. Then, for instance, $s_n = \frac{1}{n}$ does not converge in $(\mathbb{R}_+ +, |\cdot|)$.

Theorem 6.2.1. *If a sequence $\{x_n\}$ converges, then its limit is unique.*

Proof. Assume $\{x_n\}$ converges to a and b. If $a \neq b$, then we can construct a metric $d(a, b)$. This satisfies the triangle inequality so that:

$$d(a, b) \leq d(a, x_n) + d(x_n, b)$$

Let's assume $\epsilon = \frac{d(a,b)}{2} > 0$. Then, we can write:

$$d(a, x_n) < \frac{d(a, b)}{2} \quad \forall n \geq n_1, n_1 \in \mathbb{N}$$

$$d(x_n, b) < \frac{d(a, b)}{2} \quad \forall n \geq n_2, n_2 \in \mathbb{N}$$

Assuming $n = \max\{n_1, n_2\}$, $\forall m > n$, we can write $d(a, b) < \epsilon$. Combining all the elements, we have:

$$d(a, b) \leq d(a, x_n) + d(x_n, b) < \frac{d(a, b)}{2} + \frac{d(a, b}{2} = d(a, b)$$

This is a contradiction. Then, it must be $a = b$. □

Theorem 6.2.2. *If a sequence $\{x_n\}$ converges, then, for every $\epsilon > 0$, there exists an $M \in \mathbb{M}$ such that $d(x_{n1}, x_{n2}) < \epsilon$, for any $n_1, n_2 > M$.*

Proof. Since $\{x_n\}$ converges, then $\{x_n\} \to x$. By the triangle inequality, we have:

$$d(x_{n1}, x_{n2}) \leq d(x_{n1}, x) + d(x, x_{n2})$$

For any $\epsilon > 0$, there exists an M such that $d(x, x_{n1}), d(x, x_{n2}) < \frac{\epsilon}{2}$ □

Definition 6.2.3. A sequence $\{x_n\}$ in a metric space (X, d) is **bounded** if there exists a $s \in X$ and a $b \geq 0$ such that:

$$d(x, x_n) \leq b$$

for all $n \in \mathbb{N}$.

This means that we can find an interval $(-b, b)$ that contains all the elements of the sequence. An example of bounded sequence can be $\{x_n\} = \{\frac{1}{n}\}$, which is bounded between $(1, 0)$. A not-bounded sequence can be $\{x_n\} = \{2n\}$.

Theorem 6.2.3. *Any convergent sequence is bounded.*

Proof. Since we know that $\{x_n\} \to x$, we can take a ϵ s.t. $d(x_n, x) < \epsilon$ for an $m \in \mathbb{N}, m > n$. If we fix $\epsilon = 1$, we can define b as:

$$b = \max\{1, d(x_n, x)\}$$

Then, $d(x_n, x) \leq 1$. □

However, notice that the opposite may not be true. Not all bounded sequences converge. The classical example is $\{x_n\} = (-1)^n$. It is bounded because it is comprised in the interval $(-1, 1)$ but does not converge.

Theorem 6.2.4. *Suppose $\{x_n\} \to x$ and $x_n \geq 0, \forall n$. Then, $x \geq 0$. This means that the weak inequality is preserved at the limit.*

Proof. Assume $x < 0$. We can find an ϵ such that $d(x_n, x) < \epsilon$. Take $\epsilon = |x|$. Then, we can write:

$$|x_n - x| < |x| \quad \forall n \geq m \quad n, m \in \mathbb{N}$$

Therefore, $x_n < 0$, and we have reached a contradiction. □

However, this result holds only with weak inequality.
With converging sequences, we can make the following operations.

Theorem 6.2.5. *Consider $\{x_n\}, \{y_n\} \subset \mathbb{R}$, convergent sequences with $\{x_n\} \to x$ and $\{y_n\} \to y$. Then the following properties hold:*

1. $x_n + y_n \to x + y$
2. $x_n \cdot y_n \to x \cdot y$
3. $\frac{x_n}{y_n} \to \frac{x}{y}$ *if* $y_n \neq 0$ *and* $x \neq 0$

Proof. Let's see 1). We need to show that:

$$|(x_n + y_n) - (x + y)| < \epsilon$$

For $n \geq m$, and $n, m \in \mathbb{N}$. Rearranging terms and applying the triangle inequality, we have:

$$|(x_n + y_n) - (x + y)| = |(x_n - x) + (y_n - y)| \leq |x_n - x| + |y_n - y|$$

Since we know that $\{x_n\} \to x$, we know that there exists a n_1 such that:

$$|x_n - x| < \epsilon$$

For $n_1 \geq m$. The same for $\{y_n\} \to y$.

$$|y_n - y| < \epsilon$$

For $n_2 \geq m$. Taking $n = \max\{n_1, n_2\}$, and letting $\epsilon = \frac{\epsilon}{2}$ then we have:

$$|(x_n - x) + (y_n - y)| \leq |x_n - x| + |y_n - y| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

for all $n \geq m$.

Let's see 2). Notice that:

$$(x_n y_n - xy) = \underbrace{x_n(y_n - y) + y(x_n - x)}_{\text{by adding and subtracting } yx_n}$$

Then, applying triangle inequality:

$$|x_n y_n - xy| \leq |x_n(y_n - y)| + |y(x_n - x)|$$

But recall that any convergent sequence is bounded, so a supremum exists: let's call it $M_x = sup\{|x_n|\}$. And we can write:

$$|x_n y_n - xy| \leq M_x|y_n - y| + |y||x_n - x|$$

For any $\epsilon > 0$, there are n_1, n_2 such that:

$$|y_n - y| < \frac{\epsilon}{2M_x}$$

$$|x_n - x| < \frac{\epsilon}{2y}$$

So, we can write:

$$\begin{aligned} &|x_n y_n - xy| \leq |x_n(y_n - y)| + |y(x_n - x)| \leq \\ &\leq M_x|y_n - y| + |y||x_n - x| < \\ &M_x \frac{\epsilon}{2M_x} + y\frac{\epsilon}{2y} = \epsilon \end{aligned}$$

Let's see 3) □

Theorem 6.2.6. *Consider a real-vector sequence $\{x_n\} = \{(x_k^1, x_k^2, \ldots, x_k^n)\}$ in the euclidean space $(\mathbb{R}^n, \|\cdot\|)$. $\{x_n\}$ converges if and only if $\{x_k^i\}$ converges for $i = 1, 2, \ldots, n$ and $\{x_k\} \to x$, and $\{x_k^i\} \to x^i$, $\forall i = 1, 2, \ldots, n$, where $s = (s_1, s_2 \ldots, s_n)$.*

Proof. □

A sequence can be:

- **increasing** if $x_{n+1} \leq x_n$ (strictly with $<$)
- **decreasing** if $x_{n+1} \geq x_n$ (strictly with $>$)
- **monotone** is (strictly) increasing or decreasing

Theorem 6.2.7. *A monotone sequence $\{x_n\}$ converges if and only if it is bounded*

Proof. A convergent sequence is bounded. We want to show that an increasing and bounded sequence converges. Take $\{x_n\}$ to be monotone and bounded. If the sequence is increasing and bounded, there is a least upper bound, $x = supS$, where S denotes the non-empty bounded set $\{x_n\}$. Let's prove if $\{x_n\} \to x$. Since x is a least upper bound, then $x - \epsilon$ is not an upper bound. So there is a point x_N in $\{x_n\}$ such that:

$$x - \epsilon < x_N$$

Since $\{x_n\}$ is increasing, then, if $N \leq n$, then $x_N \leq x_n$. Hence:

$$x - \epsilon < x_N \leq x_n \leq x < x + \epsilon$$

And then:

$$|x_n - x| < \epsilon$$

□

We can also define **subsequences**.

Definition 6.2.4. Let $\{x_n\}$ be a sequence and $\{n_k\}$ be any sequence of real numbers such that:

$$n_1 < n_2 < \dots$$

The sequence $\{x_{nk}\}$ is called the **subsequence** of $\{x_n\}$.

For example, a series $\{x_1, x_2, x_3, x_4, x_5 \dots\}$ have a subsequence given by $\{x_2, x_4, \dots\}$.

Theorem 6.2.8. *In a metric space (X, d), a sequence converges to x if and only if any subsequence $\{x_{nk}\}$ converges to x.*

Proof. ($\Rightarrow$) Suppose $\{x_n\}$ convrges to x. Take $\{x_{nk}\}$ as an arbitrary subsequence. For any $\epsilon > 0$ we know that there exists a $n > m$ for which:

$$d(x_n, x) < \epsilon$$

We can similarly find an $n_k > m$ such that:

$$d(x_{nk}, x) < \epsilon$$

($\Leftarrow$) Suppose every subsequence $\{x_{nk}\}$ converges to x. Then $\{x_n\}$ converges to x as well because a sequence is always a subsequence of itself. □

This theorem can be useful to check if a sequence does not converge. Indeed, it is sufficient to find two subsequences with different limits. For example, $\{x_n\} = (-1)^n$ it is not convergent. Indeed, the subsequence $\{x_{2n}\} = (-1)^{2n}$ converges to 1, the subsequence $\{x_{2n+1}\} = (-1)^{2n+1}$ converges to -1.

When a sequence converges to its limit, the terms get closer to each other as n gets larger. This is called **Cauchy Property**.

Definition 6.2.5. A sequence $\{x_n\}$ in a metric space (X, d) is said to be a **Cauchy Sequence** if for all $\epsilon > 0$, there exists a number N such that, for all $n, m \geq N$, we have:

$$d(x_n, x_m) < \epsilon$$

Theorem 6.2.9. *Every convergent sequence is a Cauchy sequence*

Proof. We need to show that, for all ϵ it exists N such that, if $n, m \geq N$, then $d(x_n, x_m) < \epsilon$. We know that $\{x_n\}$ converges to x. By the triangle inequality:

$$d(x_n, x_m) \leq d(x_n, x) + d(x, x_m)$$

We know that $d(x_n, x) < \epsilon$ for all $\epsilon > 0$. Further, if $m > n$, then $d(x_m, x) < \epsilon$. Let's take $\epsilon = \frac{\epsilon}{2}$ Then, we have:

$$d(x_n, x_m) \leq d(x_n, x) + d(x, x_m) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

□

Notice that the opposite is not always true. There can be some Cauchy Sequences that do not converge. It depends on the choice of the metric space. The classical example is: $\{x_n\} = \frac{1}{n}$ on the metric space $\mathbb{R}_+ +$. It is a Cauchy sequence, but it does not converge.

The metric spaces where Cauchy Sequences converge are called **Complete Spances**.

Definition 6.2.6. A metric space (X, d) is called **complete** if every Cauchy Sequence in X converges in X.

We can prove that $\mathbb{R}$ and $\mathbb{R}^n$ are complete spaces.

Theorem 6.2.10. *The metric space* $(\mathbb{R}, |\cdot|)$ *is complete.*

Proof. To show this, we first assume that any Cauchy Sequence is bounded. Then, by Bolzano-Weierstraß theorem (see below), there is convergent subsequence $\{x_{n_k}\} \to x$. To see that any Cauchy Sequence is bounded, we just need to show that there is b such that $|x_n - x_m| \leq b$, for all $n, m \in \mathbb{N}$. Since it is a Cauchy Sequence, we know $|x_n - x_m| < \epsilon$. Taking $\epsilon = 1$, we can define b as $\max\{1, |x_n - x_m|\}$. Then, $|x_n - x_m| \leq b$.

Any bounded sequence has a convergent subsequence (Bolzano-Weierstraß Theorem), so $\{x_{nk}\} \to x$. By triangular inequality:

$$|x_n - x| \leq |x_n - x_{n_k}| + |x_{n_k} - x|$$

Then, $\forall \epsilon > 0$, there are n_1, n_2 such that:

$$|x_n - x_{n_k}| < \frac{\epsilon}{2} \quad \forall n, n_k > n_1$$

$$|x_{n_k} - x| < \frac{\epsilon}{2} \quad \forall n_k > n_2$$

Taking $N = \max\{n_1, n_2\}$, then:

$$|x_n - x| \leq |x_n - x_{n_k}| + |x_{n_k} - x| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

And $\{x_n\} \to x$. □

Theorem 6.2.11. *The metric space $(\mathbb{R}^n, \|\cdot\|)$ is complete*

Proof. □

Theorem 6.2.12. *(Bolzano-Weierstraß Theorem) Every bounded sequence in $(\mathbb{R}, |\cdot|)$ has a convergent subsequence.*

Proof. Take the bounded sequence $\{x_n\} \subseteq [a_1, b_1]$. Since $[a_1, b_1] \subset \mathbb{R}$, there are infinitely many elements in it. We can find a mid-point $\frac{a_1 - b_1}{2}$. Take $[a_1, \frac{a_1 - b_1}{2}]$ to construct a subsequence x_1. We can continue for infinitely many steps up to where b_k and a_k collapse on each other. Indeed $|b_k - a_k| = \frac{1}{2^{k-1}} \cdot |b_1 - a_1| = 0$ for values of k very large. So $a_k = b_k$.

We have therefore built a subsequence, $\{x_{n_k}\}$. We must show that it converges to x. Since $b_n \leq x \leq a_n$, we can write, for all $\epsilon > 0$:

$$|x_{n_k} - x| \leq |x_{n_k} - a_n| + |a_n - x|$$

$|x_{n_k} - a_n|$ is bounded between b_n and a_n, but we also known that $|b_n - a_n| \to 0$. So, we can write:

$$|x_{n_k} - x| \leq |x_{n_k} - a_n| + |a_n - x| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

Then $\{x_{n_k}\}$ converges and $\mathbb{R}$ is complete. □

6.3 Contraction Mapping Theorem

This is a **fixed point theorem**, with some important implications.

Definition 6.3.1. Let (X, d) be a metric space and $T : X \to X$ a function mapping X into itself. T is a **contraction mapping** if it exists a $\beta \in (0, 1)$ (called **modulus**), such that:

$$d(T(x), T(y)) \leq \beta d(x, y) \quad \forall x, y \in X$$

Definition 6.3.2. The fixed points of a mapping T, are elements of X such that:

$$T(x) = x$$

Then, for any arbitrary point $x_0 \in X$, we can construct a sequence of functions $\{f(x_n)\}$ which converges on the fixed point $x^* \in X$.

Then, we have the following theorem.

Theorem 6.3.1 (Contraction Mapping Theorem, or Banach Fixed Point Theorem)**.** *If (X, d) is a complete metric space and $T : X \to X$ is a contraction mapping, with modulus β, then:*

1. *T has exactly one fixed point $x^* \in X$*
2. *for any $x_0 \in X$, $d(T^n(x), x^*) \leq \beta^n d(x_0, x^*)$, where $T^n(x) = T \circ T^{n-1}(x)$*

Proof. In this proof, we need to show the existence and the uniqueness of x^*. We start by constructing a Cauchy Sequence. Take any $x \in X$, and $m, n \in \mathbb{N}$. We can assume $m > n$. Then we have:

$$\begin{aligned}
d(T^m(x), T^n(x) &\leq \underbrace{d(T^m(x), T^{m-1}(x)) + d(T^{m-1}(x), T^n(x))}_{\text{by Triangle inequality}} \\
&\quad d(T^m(x), T^{m-1}(x) + \underbrace{d(T^{m-1}(x), T^{m-2}(x))}_{\text{by triangle inequality on } d(T^{m-1}(x), T^n(x)} \cdots + d(T^{n+1}(x), T^n(x)) \\
&\leq \beta\Big[d(T^{m-1}(x), T^{m-2}(x) + d(T^{m-2}(x), T^{m-1}(x) + \dots d(T^n(x), T^{n-1}(x)\Big] \\
&\leq (\beta^{m-1} + \dots \beta^n) d(T(x), x) \\
&= \frac{\beta^n}{1-\beta} d(T(x), x)
\end{aligned}$$

As $n \to \infty$, then $d(T^m(x), T^n(x)) \to 0$. So $\{T^n(x)\}$ is a Cauchy sequence in a complete space that converges to x^*.

We must prove now the existence of x^*, namely that x^* is a fixed point:

$$T(x^*) = x^*$$

By triangle inequality:

$$\begin{aligned}
d(T(x^*), x^*) &\leq d(T(x^*), T^n(x)) + d(T^n(x), T(x^*)) \\
&\leq \beta d(x^*, T^{n-1}(x)) + d(T^n(x), x^*) \quad \forall n
\end{aligned}$$

Both terms on the right-hand side go to zero as $n \to \infty$. So $d(T(x^*), x^*) = 0$.

Finally, we must show that x^* is the unique fixed point. Suppose it is not, so $x^* \neq y$ are both fixed points. Then we have $T(x^*) = x$ and $T(y) = y$. Thus:

$$d(x^*, y) = d(T(x^*), T(y) \leq \beta d(x^*, y)$$

But $\beta \in (0, 1)$, $d(x^*, y) \leq \beta d(x^*, y)$ is impossible. We have reached a contradiction.

Finally, we can write the convergence path as follows: take any initial point x_0. Since $T(x^*) = x^*$, then:

$$d(T^n(x_0), x^*) = d(T^n(x_0), T(x*)) \leq \beta d(T^{n-1}(x), x^*) \leq \cdots \leq \beta^n(x_0, x^*)$$

□

Chapter 7

Topology: some elements

We focus on the basic topology of Euclidean Spaces.

Definition 7.0.1. A set $S \subseteq \mathbb{R}^n$ is **open**, if, for every $s \in S$, it exists a $\epsilon > 0$ such that $N_\epsilon(s) \subseteq S$.

The simplest example of an open set is an open interval in $\mathbb{R}$.

Definition 7.0.2. A set $S \subseteq \mathbb{R}^n$ is **closed** if its complementary is open.

Again, the simplest example of an open set is a closed interval in $\mathbb{R}$.

Still, notice that a set can be neither closed nor open. For instance, a half-closed interval.

Definition 7.0.3. Let $S \subseteq \mathbb{R}^n$. A point $s \in \mathbb{R}^n$ is called a **limit point** of S if there exists a sequence $\{s_n\} \subseteq S$ and $\{s_n\} \to s$

Theorem 7.0.1. *A set is closed if and only if it contains all its limit points.*

Proof. ($\Rightarrow$) If S is closed , it contains all its accumulation points. Suppose it does not. Then, there is a sequence $\{s_n\} \subseteq S$, such that $\{s_n\} \to s$, $s \notin S$. Since S is closed, then S^C is open. We can find an ϵ' such that $N_\epsilon(s) \subseteq S^C$. However, since $\{s_n\} \to s$, we can find an m such that $\forall \epsilon$, $s_n \in N_\epsilon(s)$, for $n > m$. If $\epsilon' = \epsilon$, then we have $s_n \in S^C$ when $n > m$, but this contradicts the assumption that $\{s_n\} \subseteq S$.

($\Leftarrow$). If S contains all its limit points, then it is closed. Suppose that S is not closed. Then it does not contain all its limit points (contrapositive statement). If S is not closed, then S^C is not open. Then it exists a $x \in S^C$ such that $\forall \epsilon > 0$, $N_\epsilon(x) \cap S \neq \emptyset$. If $\epsilon = \frac{1}{n}$, then we can have $s_n \in N_{\frac{1}{n}}(x) \cap S$. So $\{s_n\} \in S$ and $\{s_n\} \to x$.

But then, it is not true that every converging sequence $\{s_n\} \in S$ implies that the limit also lies in S. □

Theorem 7.0.2. $\mathbb{R}^n$ *and* $\emptyset$ *are both open and closed.*

Proof. $\mathbb{R}^n$ is open. But also closed since it contains all its limit points. $\emptyset$ can be written as $(\mathbb{R}^n)^C$, so it is also open and closed. □

Theorem 7.0.3. *The following properties hold:*

1. *The union of arbitrarily many open sets is open*
2. *The intersection of finitely many open sets is open*
3. *The intersection of arbitrarily many closed sets is closed*
4. *The union of finitely many closed sets is closed.*

Proof. Let's start with 1). This can be written as:

$$\bigcup_{i=1}^{\infty} A_i$$

is open if all A_i are open. Take a $s \in \bigcup_{i=1}^{\infty} A_i$, then $s \in A_i$ (at least one, by the definition of the union of sets). Since A_i is open, then there exists an $\epsilon > 0$ such that $N_\epsilon(s) \subseteq A_i$ and therefore:

$$N_\epsilon(s) \subseteq A_i \bigcup_{i=1}^{\infty} A_i$$

Let's see 2). This can be written as:

$$\bigcap_{i=1}^{n} A_i$$

is open if A_i is open for all $i = 1, \ldots, n$. Take $s \in \bigcap_{i=1}^{n} A_i$, so $s \in A_i$ for every $i = 1, \ldots, n$. For each A_i, then we have an $\epsilon_i > 0$ such that:

$$N_{\epsilon_i}(s) \subseteq A_i.$$

Taking $\epsilon = \max\{\epsilon_i\} > 0$, we have:

$$N_{\epsilon_i}(s) \subseteq \bigcap_{i=1}^{n} A_i$$

And $\bigcap_{i=1}^{n} A_i$ is open.

Let's see 3) This can be written as:

$$\bigcap_{i=1}^{\infty} A_i$$

is closed if all A_i are closed. Take a converging sequence $\{s_n\} \subseteq \bigcap_{i=1}^{\infty} A_i$. Then $\{s_n\}$ must belong to A_i for all i. Since A_i is closed, then its limit is in A_i, too, for all i. Then, $s \in \bigcap_{i=1}^{\infty}$.

Let's see 4). This can be written as:

$$\bigcup_{i=1}^{n} A_i$$

is closed if A_i are closed. We can prove a contrapositive: if $\bigcup_{i=1}^{n} A_i$, then at least one A_i is not closed. Since $\bigcup_{i=1}^{n} A_i$, there is a converging sequence in $\bigcup_{i=1}^{n} A_i$ whose limits is not in $A_i, \forall i = 1, \ldots, n$. There must be at least one A_i which has infinitely many elements in $\{s_n\}$. We can take a subsequence $\{s_{n_k}\}$ that converges at s (because $\{s_n\} \to s$). Since $s \notin A_i$, A_i is not closed. □

Definition 7.0.4. Let $A \subseteq \mathbb{R}^n$. A point $x \in \mathbb{R}$ is a **contact point** of A if, for any $\epsilon > 0$, we have $N_\epsilon(x) \cap A \neq \emptyset$.

Put differently, a contact point is a point such that any open interval containing it also has some element in common with A. Still, a contact point may also not be a point of A if A is open, for instance.

The set of all contact points is called **closure**.

Definition 7.0.5. Let $A \subseteq \mathbb{R}^n$. The set of all contact points:

$$\bar{A} = \left\{ x \in \mathbb{R}^n : N\epsilon(x) \cap A \neq \emptyset, \forall \epsilon \right\}$$

it is called **closure** of A and it is defined as $\bar{A}$ or $cl(A)$.

$A \subseteq \bar{A}$ because any x such that $\{x_n\} \to x$ or it is in A or it is in $\bar{A}$.

Theorem 7.0.4. *$\bar{A}$ is closed.*

Proof. We need to show that any sequence $\{x_n\}$ has its limit points in $\bar{A}$. Assume it is not. Then $x \in A^C$, so x is not a contact point of A. So, we have (by definition of contact point):

$$N_\epsilon(x) \cap A = \emptyset$$

Any open neighborhood is an open set, so for $\{x_n\} \to x$, $x_n \in N_\epsilon(x), \forall \epsilon$. Take a x_n for $n \geq m$, $n, m \in \mathbb{N}$ and $\epsilon' > 0$ such that $\epsilon' < \epsilon$. Therefore, we have $N'_\epsilon(x_n) \subseteq N_\epsilon(x)$. Hence:

$$N_{\epsilon'}(x_n) \cap A = \emptyset$$

and x_n is not a contact point. But this is a contradiction because $x_n \in A$, and therefore in $\bar{A}$. □

An example of closure can be: $A = (0, 1), \bar{A} = [0, 1]$. Another example: $A = \{\frac{1}{n} : n \in \mathbb{N}\}. \bar{A} = A \cup 0$.

$\bar{A}$ is the smallest closed set containing A.

Theorem 7.0.5. *Let $A \subseteq \mathbb{R}^n$. A is closed if and only if $A = \bar{A}$.*

Proof. ($\Rightarrow$)If A is closed, then $\bar{A} = A$. Notice that $A = \bar{A}$, means that $A \subseteq \bar{A}$ and $\bar{A} \subseteq A$. The first is obvious by the definition of closure. Let's see $\bar{A} \subseteq A$. We assume that A is closed, so any convergent sequence in A has its limits in A. We want to show that there exists $x \in \bar{A}$ such that $x \in A$, for all x. $x \in \bar{A}$ means $N_\epsilon(x) \cap A \neq \emptyset$, $\forall \epsilon$. We can take $\{x_n\} \in N_\epsilon(x)$. Since A is closed, its limit x is in A. So, $\bar{A} \subseteq A$.

($\Leftarrow$)If $\bar{A} = A$, then A is closed. This is trivial because we have shown that $\bar{A}$ is closed. □

Definition 7.0.6. Let $A \subseteq \mathbb{R}^n$. The **boundary set** of A, defined by ∂A is:

$$\partial A = \bar{A} \cap \bar{A^C}$$

Theorem 7.0.6. *∂A is closed.*

Proof. This is immediate since $\bar{A}$ and $\bar{A^C}$ are closed sets, and the union of closed sets is closed itself. □

Definition 7.0.7. Let $A \subset \mathbb{R}^n$. $x \in \mathbb{R}^n$ is called an **interior point** of A if there exists $\epsilon > 0$ such that $N_\epsilon(x) \subseteq A$. The set of all interior points is the **interior** of A, $int(A)$.

Theorem 7.0.7. *Let $A \subseteq \mathbb{R}^n$. $int(A) \subseteq A$ and A is open.*

Proof. □

Theorem 7.0.8. *Let $A \subseteq \mathbb{R}^n$. A is open if and only if $A = int(A)$.*

Proof. ($\Rightarrow$) If A is open, then by definition of openness, $\forall x \in A$, it exists $\epsilon > 0$ such that $N_\epsilon(x) \subseteq A$. Then x is in $int(A)$, and $A \subseteq int(A)$. Since $A \subseteq int(A)$, then $A = int(A)$.

($\Leftarrow$) $int(A)$ is open. So if $A = int(A)$, then A is open. □

Theorem 7.0.9. *Let $A \subseteq \mathbb{R}^n$. Then:*

1. $int(A) = A \setminus \partial A$
2. $int(A) \cap \partial A = \emptyset$
3. $\partial A \cup int(A) = \bar{A}$

Proof. □

Definition 7.0.8. Let $A \subseteq \mathbb{R}^n$ is called **(sequentially) compact** if every sequence $\{x_n\} \in A$ has a subsequence $\{x_{n_k}\} \to s \in A$.

Notice that $\{x_n\}$ may not be convergent. And this definition of compactness holds only in metric spaces.

For example, $[a, b] \subseteq \mathbb{R}$ is compact. Indeed, any sequence $\{x_n\}$ in the interval is bounded and therefore has a convergent subsequence to x(Bolzano-Weierstraß theorem). Furthermore, $[a, b]$ is a closed set, so for any subsequence in it, $x \in [a, b]$.

A famous result is the following.

Theorem 7.0.10. *A compact set is closed and bounded.*

Proof. Suppose A is compact. This implies that A is closed. Indeed, taking a converging subsequence $\{x_n\}$, by compactness, there is a converging subsequence $\{s_{n_k}\} \to s \in A$. So $\{x_n\} \to x$, and A is closed.

If A is unbound, then it is not compact. The unboundedness means that $|x_n - x| > b \ \forall n \in \mathbb{N}$. Compactness implies that there is a convergent subsequence in A, but convergence implies boundedness, which is not the case here. □

Theorem 7.0.11. *Suppose that A and B are compact sets. Then:*

1. *$A \cap B$ is compact*
2. *$A \cup B$ is compact*

Proof. Let's see 1). Take a sequence $\{x_n\} \in A \cup B$, so $\{x_n\} \in A$ and B. Since A is compact, there is a subsequence $\{x_{n_k}\}$ whose limit x is in A. Because B is compact, there is a subsequence $\{x_{n_k}\}$ whose limit y is in B. Then $x = y \in A \cap B$.

Let's see 2). Take a sequence $\{x_n\}$ in $A \cup B$. This means that $\{x_n\}$ is in A, in B or both. Since A and B are compact, we can find $\{s_{n_k}\}$ that has a limit in A, or in B, or both. Then $A \cup B$ is compact. □

This can be extended to any finite union or intersection of compact sets.

Theorem 7.0.12. *Suppose A and B are compact. Then $A \times B$ is compact.*

Proof. Take any sequence in $A \times B$, $\{x_n\} = (a_n, b_n)$, wit $a_n \in A$ and $b_n \in B$. Because A is compact, then $\{a_n\}$ has a converging subsequence to $a \in A$. Because B is compact, then $\{b_n\}$ has a converging subsequence to $b \in B$. Then $(a_{n_k}, b_{n_k}) \to (a, b)$. □

This result can be extended to the Cartesian Product of any finite number of sets. Furthermore, any box $\bigtimes_{i=1}^{n} [a_i, b_i] \subseteq \mathbb{R}^n$ is compact.

Theorem 7.0.13. *A closed subset of a compact set is compact.*

Proof. Take $A \subseteq B$, where A is closed, and B is compact (closed and bounded). Since A is closed, it contains all its limit points. Since A is a subset of B, and B is bounded, then A is bounded. So A is compact. □

Theorem 7.0.14 (Heine-Borel Theorem)**.** *Every closed and bounded subset of $\mathbb{R}^n$ is compact.*

Proof. Let A be a closed and bounded subset of $\mathbb{R}^n$. We know that each $A \subseteq \bigtimes_{i=1}^{n} [a_i, b_i] \subseteq \mathbb{R}^n$ which is compact. And a closed subset of a compact set is compact. So A is compact. □

Notice, however, that this theorem could not hold in different metric spaces.

Definition 7.0.9. Two subsets A and B of a metric space are said to be **separated** if both $A \cap \bar{B}$ and $\bar{A} \cap B$ are empty. A subset E of the metric space is said to be **connected** if E is not the union of two nonempty separated sets.

An example of a connected set is $[0, 1]$. Instead $[0, 1) \cup (1, 2]$ is not connected.

Theorem 7.0.15. *A set $E \in \mathbb{R}$ is connected if and only if $\forall x, y \in E$ and $x < z < y, z \in E$. In other words, E must be an interval.*

Proof. Connectedness $\Rightarrow$ interval. Suppose $x < z < y$, and $z \in E$. Define:

$$A_z^- = (-\infty, z) \cap E$$

$$A_z^+ = (z, \infty) \cap E$$

Since $x \in A_z^-$ and $y \in A_z^+$, these are both non-empty. Notice that $A_z^- \subseteq (-\infty, z)$ and $A_z^+ \subseteq (z, \infty)$, so they are separated. E is not connected because it is the union of two non-empty separated sets.

Interval $\Rightarrow$ connectedness. Suppose E is not connected. Then, there are two separate sets A and B such that $A \cup B = E$. We need to show that E is not an interval. Pick $x \in A$ and $y \in B$, $x < y$. Define $z = sup(A \cap [x, y])$ (where $A \cap [x, y] \neq \emptyset$. So $z \in \bar{A}$. Otherwise, it exists an $\epsilon > 0$ such that $z - \epsilon$ is an upper bound of A. Since A, B are separated, $z \notin B$, and thus $x \leq y < z$. If $z \notin A$, $x < z < y$ and $z \in E$. If $z \in A$, then $z \notin \bar{B}$. Then it exists z' such that $z < z' < y$. Thus, $z' \notin A$ (z is sup), $z' \notin B$, ($B \subseteq \bar{B}$), so $z' \notin E$. $\square$

Chapter 8

Continuity

Let (X, d_x) and (Y, d_y) be two metric spaces. A function $f : X \to Y$ maps elements from X to Y. For each sequence $\{x_n\} \subseteq Y$, we have a corresponding sequence $\{y_n\} \subseteq Y$, where $y_n = f(x_n)$.

To preserve the topological properties of X in Y, y_n must converge in Y. This is, in a nutshell, what continuity means.

Definition 8.0.1. A function $f : X \to Y$ is **continuous** if $\forall \epsilon > 0$, $\forall x \in X$, it exists a $\delta > 0$ such that, $x' \in X$:

$$d_x(x, x') < \delta$$

then:

$$d_y(f(x), f(x')) < \epsilon$$

Theorem 8.0.1. *Let $f : X \to Y$. f is continuous if and only if send any convergent sequence $\{x_n\}$ in a convergence sequence $\{y_n\} \in Y$. That is:*

$$\{f(x_n)\} \to \{f(x)\}$$

As long as:

$$\{x_n\} \to x$$

Proof. ($\Rightarrow$) $f(x_n) \in Y$. By continuity of f, $\forall \epsilon > 0, \forall x \in X$ it exists a $\delta > 0$, such that, for $x' \in X$ $d(x, x') < \delta$ and $d(f(x), f(x')) < \epsilon$.

($\Leftarrow$) f preserves convergence, then it is continuous. We prove the contrapositive: f does not preserve convergence, then it is not continuous. That is, $\{x_n\} \to x$, but $\{f(x_n)\} \not\to f(x_n)$. This means that, $\forall \epsilon > 0, \forall x \in X$ it exists a $\delta > 0$, such that, for $x' \in X$ $d(x, x') < \delta$ and $d(f(x_n), f(x')) \geq \epsilon$. Let's take $\delta = \frac{1}{n}$ For each δ_n, $x_n \in N_{\delta_n}(x)$ and $d(f(x_n), f(x')) \geq \epsilon$. □

In other words, if f is continuous, the convergence behavior of $\{x_n\}$ is preserved in Y by $\{f(x_n)\}$.

Theorem 8.0.2. *The composite of continuous functions is continuous.*

Proof. Take $f : X \to Y$ and $g : Y \to Z$. Since f is continuous, then $\{x_n\} \to x$ and $\{f(x_n)\} \to f(x)$. Since g is continuous, then $\{f(x_n)\} \to f(x)$ and $\{g(f(x_n))\} \to g(f(x))$. Then $(g \circ f)(x_n)$ is continuous. □

Theorem 8.0.3. *Let $f : X \to Y$ and $g : X \to Y$, f, g are continuous. Then, the following functions are also continuous:*

1. $h(x) = f(x) + g(x)$
2. $h(x) = f(x)g(x)$
3. $\frac{f(x)}{g(x)}$ *if* $g(x) \neq 0, \forall x \in X$

Proof. This derives directly from the properties of sequences (Theorem 6.5, above). □

Theorem 8.0.4. *Let $f : X \to \mathbb{R}$. The following statements are equivalent:*

1. *f is continuous*
2. *For each closed set $V \subseteq f(X), f^{-1}(V)$ is closed*
3. *For each open set $V \subseteq f(X), f^{-1}(V)$ is open*

Proof. (1) $\Rightarrow$ (2) f is continuous and $V \subseteq f(X)$ is closed. We want to show $f^{-1}(V)$ is closed. Take a convergent sequence $\{x_n\} \subseteq f^{-1}(V)$. We must show that $x \in f^{-1}(V)$. $x_n \in f^{-1}(V)$, so $f(x_n) \in V$. f is continuous, so $\{f(x_n)\} \to f(x)$. V is closed, so $f(x) \in V$. That is $x \in f^{-1}(V)$.

(2) $\Rightarrow$ (3): V is open. We want to show $f^{-1}(V)$ is open. $(f^{-1}(V))^C = \{x \in X : f(x) \notin U\} = f^{-1}(U^C)$. Let $U = V^C$. When V is open, U is closed. $f^{-1}(U)$ is closed. $f^{-1}(V) = f^{-1}(U^C) = (f^{-1}(U))^C$.

(3) $\Rightarrow$ (1): We want to show the ϵ, δ. Pick $x \in X$ and $\epsilon > 0$ be given. $N_\epsilon(f(x))$ is open. $f^{-1}(N_\epsilon(f(x)))$ is open and contains x. Therefore, it exists a $\delta > 0$ such that $N_\delta(x) \subseteq f^{-1}(N_\epsilon(f(x)))$. Each $x' \in N_\delta(x)$ must also belong to $f^{-1}(N_\epsilon(f(x)))$. □

Theorem 8.0.5. *Suppose that $f : A \to B$ is continuous. Then:*

- *If $f(x_0) > 0$, it exists $\delta > 0$ such that $f(x) > 0$, $\forall x \in N_\delta(x_0)$.*
- *If $f(x_0) < 0$, it exists $\delta > 0$ such that $f(x) < 0$, $\forall x \in N_\delta(x_0)$.*

Proof. □

Theorem 8.0.6. *Suppose $f : A \to B$ and $g : f(A) \to C$ are continuous then. $g \circ f$ is continuous.*

Proof. □

Definition 8.0.2. A function $f : X \to \mathbb{R}$ is **bounded** if there is a $B \in \mathbb{R}$ such that $|f(x)| \leq B$, for any $x \in X$.

Theorem 8.0.7. *Suppose that a continuous function $f : X \to \mathbb{R}$, where X is compact. Then $f(X)$ is compact.*

Proof. For every convergent sequence in Y, there is a converging subsequence. Because X is compact, by def. we know that there is a subsequence:

$$\{x_{n_k}\} \to x^* \in X$$

So, $f(x^*) \in f(X)$. Because f is continuous, we have:

$$f(x_{n_k}) \subseteq \{f(x_n)\} \to f(x^*)$$

□

Theorem 8.0.8 (Weierstraß Theorem)**.** *Let $f : X \to \mathbb{R}$, f continuous and X compact, and:*

$$M = \sup f(x)$$

$$m = \inf f(x)$$

Then there exists $p, q \in X$ such that $f(p) = M$ and $f(q) = m$.

Proof. X is compact and f is continuous. Then $f(X)$ is compact and a subset of R. By the least upper bound property, sup and inf exist. Take a sequence $\{x_n\} \subseteq X$ such that $f(x_n) \to M$. Because $f(X)$ is closed, $M \in f(X)$. The same for $m \in F(X)$. □

Definition 8.0.3. A function $f : X \to \mathbb{R}$ is **uniformly continuous** if for every $\epsilon > 0$, there exist $\delta > 0$ such that:

$$|f(x) - f(x')| < \epsilon$$

for all $x, x' \in X$, such that $\|x - x'\| < \delta$.

The definition between this concept and continuity mainly resides in the fact that uniform continuity is a property of a function in a set, whereas continuity is defined at any single point.

Every uniformly continuous function is continuous. The opposite is not true.

Theorem 8.0.9. *Let f be a continuous real-valued function of a compact set X. Then f is uniformly continuous on X.*

Proof. □

Theorem 8.0.10. *Let $f : E \to \mathbb{R}$. Suppose there is $K > 0$ such that:*

$$|f(x) - f(y)| \leq K \cdot |x - y|$$

*f is called **Lipschitz continuous** on E. The inequality is called **Lipschitz condition**.*

Proof. □

Theorem 8.0.11. *If f is Lipschitz continuous, then it is uniformly continuous.*

Proof. This is immediate taking $\epsilon = \frac{\epsilon}{K}$. □

Theorem 8.0.12. *If $f : X \to \mathbb{R}$ is continuous, $E \subseteq X$ is connected, then $f(E)$ is connected.*

Proof. □

Theorem 8.0.13 (Intermediate Value Theorem)**.** *Let $f : [a, b] \to \mathbb{R}$ be continuous. If $f(b) > f(a)$ and $f(a) > c > f(b)$, then there exists a point $x \in (a, b)$ such that $f(x) = c$.*

Proof. $f([a, b])$ is connected, so $c \in f([a, b])$. Thus, there must be $x \in [a, b]$ such that $f(x) = c$. Since $x \neq a$, then $x \in (a, b)$. □

Definition 8.0.4. A function $f : (a, b) \to \mathbb{R}$. Consider any point x such that $a \leq x \leq b$. We define the **right-hand limit** as:

$$f(x^+) = q$$

If $f(t_n) \to q$ as $t_n \to x^+$

We define the **Left-hand limit** as:

$$f(x^-) = q$$

As $f(t_n) \to q$ as $t_n \to x^-$

Then, $lim_{t \to x} f(t)$ exists if and only if $f(x^+) = f(x^-) = lim_{t \to x} f(t)$. And a function is continuous if and only if $f(x^+) = f(x^-) = f(x)$.

Definition 8.0.5. A function $f : (a, b) \to \mathbb{R}$. f is **monotonically increasing** on (a, b) if $a < x < y < b$ implies that $f(x) \leq f(y)$. If $f(x) \geq f(y)$, then the function is **monotonically decreasing**. In either case, the function is **monotone**.

Theorem 8.0.14. *Let f be monotonically increasing on (a, b). Then, $f(x^+)$ and $f(x^-)$ exist at every point of $x \in (a, b)$.*

Proof. □

Theorem 8.0.15. *If $f : [a, b] \to \mathbb{R}$ is monotone function. Then, the set of points at which f is discontinuous is at most countable.*

Proof. □

8.1 Continuity of correspondences

Recall that a correspondence[1] is a set-valued function, namely a function that maps points in the domain to not empty subsets of the co-domain. Then, considering two metric spaces (Θ, d_Θ) and (X, d_X):

$$\psi : \Theta \rightrightarrows X$$

Let's note that:

- A function is a special case of correspondence, namely when the set is a singleton
- We can write any correspondence in functional notation. Denoting 2^B the set of all subsets of B, then we can write:

$$f : A \to 2^X$$

Some concepts pertaining to correspondences are the following.

Definition 8.1.1. Let $\Theta \subseteq \mathbb{R}^n$ and $X \subseteq \mathbb{R}^l$. A correspondence $\psi : \Theta \rightrightarrows X$ is said to be:

- **closed-valued** at $\theta \in \Theta$ if $\psi(\theta)$ is a closed set. If it is closed valued at all $\theta \in \Theta$, then $\psi(\cdot)$ is closed-valued.
- **compact-valued** at $\theta \in \Theta$ if $\psi(\theta)$ is a compact set. If it is compact valued at all $\theta \in \Theta$, then $\psi(\cdot)$ is compact-valued.
- **convex-valued** at $\theta \in \Theta$ if $\psi(\theta)$ is a convex set. If it is convex valued at all $\theta \in \Theta$, then $\psi(\cdot)$ is convex valued.

Definition 8.1.2. The **graph** of a correspondence, denoted $Gr(\psi)$, is defined as:

$$Gr(\psi) = \Big\{(\theta, s) \in \Theta \times S : s \in \psi(\theta)\Big\}$$

Notice that $Gr(\psi) \subseteq \mathbb{R}^n \times \mathbb{R}^l$.

Definition 8.1.3. The correspondence is said to be **closed-graph** if $Gr(\psi) \subseteq \mathbb{R}^n \times \mathbb{R}^l$ is closed (namely that for all converging sequences θ_m in Θ, it exists a converging sequence s_m in $\psi(\theta_m)$ such that $s \in \psi(\theta)$.)

Definition 8.1.4. The correspondence is said to be **convex-graph** if $Gr(\psi) \subseteq \mathbb{R}^n \times \mathbb{R}^l$) is convex.

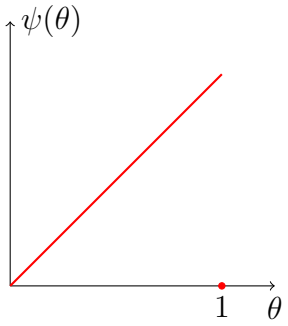
A closed-graph correspondence is a closed value, but the opposite may not be true. The same for convex-graph correspondences. Let's see the following example:

[1] Most of this section and examples are based on Sundaram 1996, 224 et ss.

Example 8.1.1. *Let $\theta = X = [0,1]$, and define $\psi(\theta) =$*

$$\psi(\theta) = \begin{cases} \{\theta\} & 0 \leq \theta < 1 \\ \{0\} & \theta = 1 \end{cases}$$

The graph is represented in the following figure.



The correspondence is compact. But it is not closed graph. Indeed, if we take the sequence $\{x_m, s_m\} = \{1 - \frac{1}{m}, 1 - \frac{1}{m}\}$, this is in the graph. But it converges to $(1,1)$, which is not in the graph.

Definition 8.1.5. Let $\Theta \subseteq \mathbb{R}^n$ and $X \subseteq \mathbb{R}^l$ and a correspondence $\psi : \Theta \rightrightarrows X$. Let W ba any set $\mathbb{R}^l$. We define the **upper inverse** of W under ψ as:

$$\psi_+^{-1}(W) = \Big\{\theta \in \Theta : \psi(\theta) \subseteq W\Big\}$$

and the **lower inverse** of W under ψ:

$$\psi_-^{-1}(W) = \Big\{\theta \in \Theta : \psi(\theta) \cap W \neq \emptyset\Big\}$$

When the correspondence is single-valued, then lower and upper inverse coincide, as well as with the definition of the inverse of a function.

Correspondences are important in economics because sometimes there can be more than one solution to optimization problems. For example, if preferences are not strictly convex, then we have a demand correspondence instead of a demand function.

For correspondences, a stronger notion of continuity is required than for a function. A function $f : X \to S$ is continuous at $x \in X$ if for all open sets V such that $f(x) \in V$, there is an open set U containing x such that, for all $x' \in X \cap U$, then $f(x') \in V$. Another way of seeing is that $f(x) \in V \iff f(x) \cap V = \emptyset$. But this is not true if $f(\cdot)$ is not a singleton any more.

Then, we have the following definitions.

Definition 8.1.6. A correspondence $\psi : \Theta \rightrightarrows X$ is said to be **upper-hemicontinuous** (or u.h.c) at a point $\theta \in \Theta$, if, for all open sets V such that $\psi(\theta) \in V$, there exists an open set U containing θ such that $\theta' \in U \cap \Theta$ implies $\psi(\theta') \subset V$. Then, ψ is upper-hemicontinuous on Θ if it u.h.c. for all $\theta \in \Theta$.

A "intuitive way" of seeing it is that a correspondence which is not u.h.c. at θ "blows up" in any neighborhood of θ, in the sense that part of $\psi(\theta)$ lies outside some small open set containing it.

Example 8.1.2. *Let $\Theta = X = [0, 2]$. Define $\psi : [0, 2] \rightrightarrows [0, 2]$ as:*

$$\psi(\theta) = \begin{cases} \{1\} & \text{if } 0 \leq \theta < 1 \\ [0, 2] & \text{if } 1 \leq \theta \leq 2 \end{cases}$$

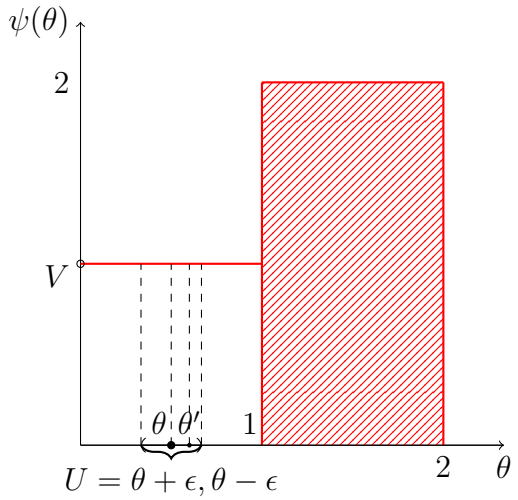


Figure 8.1: A correspondence which is u.h.c. but not l.h.c.

This correspondence is u.h.c. Let's see for $\theta < 1$. Define the interval $(\epsilon + \theta, \theta - \epsilon)$. Then $\psi(\theta) \subset V$. Take $\theta' \in U \cap [0, 2]$, then $\psi(\theta') \in V$ (since, in this case $\psi(\theta) = \psi(\theta')$). A similar reasoning holds for $\theta \geq 1$. Then, taking any open set around $\psi(\theta)$, there exists an open set U containing θ such that $\theta' \in U \cap [0, 2]$ implies $\psi(\theta') \in V$. Then, $\psi(\cdot)$ is upper hemi-continuous (See figure 1).

However, let's see another example (figure 2). Let $\Theta = X = [0, 2]$. Define $\psi : [0, 2] \rightrightarrows [0, 2]$ as:

$$\psi(\theta) = \begin{cases} \{1\} & \text{if } 0 \leq \theta \leq 1 \\ [0, 2] & \text{if } 1 < \theta \leq 2 \end{cases}$$

Following the same argument as above, this correspondence is u.h.c for $\theta \leq 1$, as well as for $\theta > 1$. But for $\theta = 1$, it is not. Indeed, we can find an open interval that contains $\psi(\theta)$ but not $\psi(\theta')$. Let's see $V = \left(\frac{2}{3}, \frac{4}{3}\right)$. This contains $\psi(1)$. But does not contain any $\psi(\theta')$, with $\theta' > 1$.

It is intuitive to see why graphically. In the graph of correspondence in Fig. 1, the rectangular part is closed. Instead, in the graph of Fig. 2, it is not. Then, when $\theta = 1$, $\psi(\theta)$ "blows up".

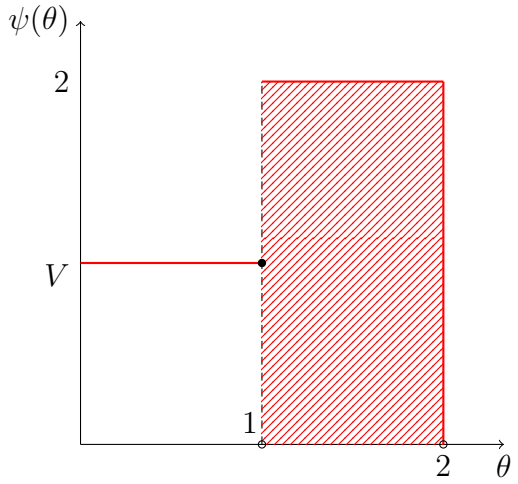


Figure 8.2: A correspondence which is not u.h.c. but l.h.c

We can also define the lower hemi-continuity.

Definition 8.1.7. A correspondence $\psi : \Theta \rightrightarrows X$ is said to be **lower hemi-continuous** (l.h.c) at a point $\theta \in \Theta$, if, for all open sets V such that $V \cap \psi(\theta) \neq \emptyset$, there exists an open set U containing θ such that $\theta' \in U \cap \Theta$ implies $V \cap \psi(\theta') \neq \emptyset$. Then ψ is lower hemi-continuous on Θ if it is l.h.c. for all $\theta \in \Theta$.

As above, geometrically, the idea is that $\psi(\theta)$ does not shrink suddenly as we move slightly away from θ.

Example 8.1.3. *See the example in Figure 1. At $\theta = 1$, consider the interval $V = \left(\frac{3}{2}, \frac{5}{2}\right)$. $V \cap \psi(1) \neq \emptyset$, because $\psi(1) = [0, 2]$. But since any open U containing $\theta = 1$ also contains $\theta < 1$, then there is no open set U containing $\theta = 1$ such that $\psi(\theta') \cap V \neq$ for all $\theta \in U \cap \Theta$. Indeed $\psi(\theta') = \{1\}$ and $\{1\} \cap \left(\frac{3}{2}, \frac{5}{2}\right) = \emptyset$.*

Similarly, in the case of the correspondence represented in Figure 2, at $\theta = 1$, $\psi(\theta) \cap V \neq \emptyset$ if $1 \in V$. Since for any $\theta' \in \Theta$, $1 \in \psi(\theta')$, then ψ is l.h.c. at $\theta = 1$. Furthermore, it is l.h.c. at $\theta \neq 1$, so it is l.h.c. on Θ.

Definition 8.1.8. A Correspondence $\psi : \Theta \rightrightarrows X$ is **continuous** at $\theta \in \Theta$ if it is both l.h.c. and u.h.c. It is continuous on Θ if it is l.h.c. and u.h.c. for all $\theta \in \Theta$.

A result that links the l.h.c. of correspondence with the upper inverse is the following:

Proposition 9. *$\psi : \Theta \rightrightarrows X$ is upper hemi-continuous if and only if $\psi_+^{-1}(G)$ is open in Θ for every G that is open in X.*

Proof. Let's see that u.h.c. implies open upper inverse. By definition of u.h.c., for all open sets V containing $\psi(\theta)$, then there exists an open set U containing θ such that $\theta' \in U \cap \Theta$ implies $\psi(\theta') \subset V$. Instead, the upper inverse is $\psi_+^{-1}(G) = \{\theta \in \Theta : \psi(\theta) \subset$

$G\}$. Assume u.h.c. holds, but not the openness of $\psi_+^{-1}(G)$. The latter implies that, for G open, $\psi_+^{-1}(G)$ is not open. Define $G = V$. If ψ_+^{-1} is not open, then (by definition of open set), $N_\epsilon(\theta) \not\subset \psi_+^{-1}(V)$. Then, for $\theta' \in N_\epsilon(\theta)$, $\psi(\theta')$ is not in V. But this contradicts u.h.c.

Let's see now that the upper inverse implies u.h.c. Let G be an open set that contains $\psi(\theta)$. Then $N_\epsilon(\psi(\theta)) \subseteq G$. Let $G = V$ and $U \cap \Theta = \psi_+^{-1}(G)$, Then $\theta \in \psi_+^{-1}(G)$ implies $\psi(\theta) \subset G$. Since $\psi_+^{-1}(G)$ is open, then $N_\epsilon(\theta) \subset \psi_+^{-1}(G)$, $\forall \epsilon > 0$. Take $\theta' \in N_\epsilon(\theta)$. Then $\theta' \in \psi_+^{-1}$. Therefore $\psi(\theta') \in G$. □

A similar result states that a correspondence is u.h.c. if and only if the lower inverse $\psi_-^{-1}(F)$ is closed in Θ for every F closed in X.

For lower hemi-continuity, we have the following result.

Proposition 10. *$\psi : \Theta \rightrightarrows X$ is lower hemi-continuous if and only if $\psi_+^{-1}(G)$ is closed in Θ for every G that is closed in X.*

Proof. □

Sometimes, it can be useful to characterize u.h.c. and l.h.c. in terms of converging sequences instead of open sets. However, this last characterization is less general because for u.h.c. it requires $\psi(\cdot)$ to be compact-valued.

Proposition 11. *Let $\psi : \Theta \rightrightarrows X$ be a compact-valued correspondence. Then ψ is u.h.c. at $\theta \in \Theta$ if and only if for all sequences $\theta_n \to \theta \in \Theta$ and for all sequences $s_n \in \psi(\theta_n)$, there is a subsequence s_{n_k} of s_n such that s_{n_k} converges to some $s \in \psi(\theta)$.*

Proof. Let's show first that $\psi(\cdot)$ compact-valued and u.h.c. implies convergence. Take $\theta_n \to \theta$ and $s_n \in \psi(\theta)$,$\forall n$, we need to show that:

1. it exists $s_{n_k} \to s$
2. $s \in \psi(\theta)$

$\psi(\cdot)$ is compact-valued, means that $\psi(\theta)$ is a compact set. Take $s_n \in \psi(\theta)$. Since it is a sequence in a compact set, it has a convergent subsequence s_{n_k} (by the sequential definition of compactness). This proves 1). To see 2), let's say $s \notin \psi(\theta)$. Then, there it is also not in a closed set G containing $\psi(\theta)$. But $\psi(\theta_n) \in G$, and $s_{n_k} \in G$. Since $s_{n_k} \to s$, and G is closed, then $s \in G$. Then, we have a contradiction.

Let's show now that convergence and compact-valued imply u.h.c. Suppose $\theta_n \to \theta$, $s_n \in \psi(\theta_n)$ and $s_{n_k} \to s \in \psi(\theta)$. Suppose $\psi(\cdot)$ is not u.h.c. Then, there exists an open set containing $\psi(\theta)$ such that, for all open U containing θ, there exists $\theta' \in U \cap \Theta$ and $\psi(\theta') \not\subset V$. Let U_m be $N_\epsilon(\theta)$, with $\epsilon = \frac{1}{m}$ $m = 1, 2....$, and $\theta_m \in U_m \cap \Theta$ such that $\psi(\theta_n) \not\subset V$. Take $s_m \in \psi(\theta_m)$, $s_m \notin V$. $\theta_m \to \theta$, by construction. Since $s_m \in \psi(\theta_m)$, then $s_{m_k} \to s \in \psi(\theta)$. But $s_m \notin V$, for each m, V is open and therefore $s \notin V$. This contradicts $\psi(\theta) \subset V$. □

Proposition 12. *Let $\psi : \Theta \rightrightarrows X$ be a correspondence. Then $\psi(\cdot)$ is l.h.c. at θ if and only if for any sequence $\theta_n \in \Theta$, such that $\theta_n \to \theta$, and any $s \in \psi(\theta)$, there exists a sequence $s_n \in X$ such that $s_n \in \psi(\theta_n)$, and $s_n \to s$.*

Proof. We show only the first part, l.h.c. implies convergence. Let's take $\psi(\cdot)$ l.h.c at θ, and let $\theta_n \to \theta$ and $s \in \psi(\theta)$. For each $k \in \mathbb{N}$, defines U_k as the neighborhood of θ such that $\psi(\theta) \cap N_\epsilon(s) \neq \emptyset$, $\forall \theta' \in U_k \cap \Theta$. Define a sequence $n_1 < n_2 < \dots n_{k-1} < n_k$ and $\theta_n \in U_k$, for all $n > n_k$. Then, we can define s_n: if $n < n_1$, $s_n \in X$, otherwise $s_n \in \psi(\theta_n) \cap N_\epsilon$. Notice that this is not empty because $\theta_n \in U_k$. Then, $s_n \to s$. □

To conclude, there are several properties of u.h.c. and l.h.c. correspondences.

Proposition 13. *Let $\psi : \Theta \rightrightarrows X$ be a u.h.c. correspondence, where $\Theta \in \mathbb{R}^n$ and $X = \mathbb{R}^l$. Then:*

1. *$\bigcup_{i=1}^n \psi_i$ is u.h.c. if ψ_i are u.h.c.*
2. *if ψ_1 and ψ_2 are u.h.c. and closed valued, then $\psi_1 \cap \psi_2$ is u.h.c. and not empty for all $\theta \in \Theta$*
3. *$\psi \circ \phi$ of two u.h.c. correspondences is u.h.c.*

Proof. □

Proposition 14. *Let $\psi : \Theta \rightrightarrows X$ be a l.h.c. correspondence, where $\Theta \in \mathbb{R}^n$ and $X\mathbb{R}^l$. Then:*

1. *$\bigcup_{i=1}^n \psi_i$ is l.h.c. if ψ_i are l.h.c.*
2. *If ψ_i,$i = 1, 2$ are two l.h.c. convex-valued correspondences such that $\psi_1 \cap \psi_2 \neq \emptyset$, then $\psi_1 \cap \psi_2$ is l.h.c.*
3. *$\psi \circ \phi$ of two l.h.c. correspondences is l.h.c.*

Proof. □

Chapter 9

Differentiation

9.1 Differentiation with One Variable

Definition 9.1.1. Let $f : E \to \mathbb{R}$ where $E \subseteq \mathbb{R}$. For any $x \in E$, if, for any sequence $\{x_n\} \to x$ such that $x_n \neq x, \forall n$, and:

$$\lim_{x_n \to x} \frac{f(x_n) - f(x)}{x_n - x}$$

exists, then f is called **differentiable** at x. This limit is called the **derivative** of $f(x)$, that is, $f'(x)$.

Differentiation is just approximating a function with a linear function. Furthermore, f is differentiable, then $f'(x^+) = f'(x^-)$. In other words, kinks are ruled out.

Theorem 9.1.1. *Let $f : E \to \mathbb{R}$ where $E \subseteq \mathbb{R}$. Both f and g are differentiable. Then the following properties hold:*

1. $[f(x) + g(x)]' = f(x)' + g(x)'$
2. $[f(x) \cdot g(x)]' = f(x)' \cdot g(x) + f(x) \cdot g'(x)$ *(Leibniz Theorem)*
3. $\left(\frac{f(x)}{g(x)}\right)' = \frac{f'(x) \cdot g(x) - f(x) \cdot g'(x)}{g(x)^2}$

Proof. Let's see 1). Let $h = f + g$. So we have:

$$\begin{aligned} &\frac{h(x+t) - h(x)}{t} = \\ &\frac{f(x+t) - f(x) + g(x+t) - g(x)}{t} = \quad \text{(taking the limit, as } t \to 0\text{)} \\ &h'(x) = [f(x) + g(x)]' = f(x)' + g(x)' \end{aligned}$$

Let's see 2). Let $h = f \cdot g$. Then, $x, x + t \in E$, we have:

$$\frac{h(x+t) - h(x)}{t} = \frac{[f(x+t) - f(x)]g(x) + f(x)[g(x+t) - g(x)]}{t}$$

Taking the limit $t \to 0$:

$$h'(x) = f'(x) \cdot g(x) - f(x) \cdot g(x)$$

Let's see 3). Let $h = \frac{f}{g}$. □

Theorem 9.1.2. *Let $f : E \to \mathbb{R}$, where $E \subseteq \mathbb{R}$. If f is differentiable at x, it is continuos at x.*

Proof. □

Theorem 9.1.3. *(Chain Rule) Let $f : E \to \mathbb{R}$ and $g : D \to \mathbb{R}$, where $D, E \subseteq \mathbb{R}$ and $D \subseteq f(E)$. f, g are both differentiable $\forall x$. Let $h(x) = (g \circ f)(x)$, then h is differentiable and its derivative is:*

$$h'(x) = [g'(f(x)] \cdot f'(x)$$

Proof. Fix $x \in E$. Let $y = f(x)$. $t \in E$, and $s = f(t)$. Then we have:

$$f(t) - f(x) = (t - x)[f'(x) + u(t)]$$

$$g(s) - g(y) = (s - y)[g'(y) + v(s)]$$

Where $\lim_{t \to x} u(t) = \lim_{s \to y} v(s) = 0$. Then we have:

$$h(t) - h(x) = g(f(t) - g(f(x)) = [f(t) - f(x)][g'(y) + v(s)]$$
$$(t - x)[f'(x) + u(t)][g'(y = +v(s)]$$

For $t \neq x$.

$$\frac{h(t) - f(x)}{t - x} = [g'(y) + v(s)][f'(x) + u(t)]$$

Since f is continuous, as $t \to x$, $s = f(t) \to y = f(x)$, so:

$$h'(x) = g'(f(x))f'(x).$$

□

Definition 9.1.2. Let f be a real-valued function on some metric space X. f has a **local maximum** at a point $p \in X$ if there exists a $\delta > 0$ such that $f(x) \leq f(p), \forall x \in N_\delta(x)$. f has a **local minimum** at a point $q \in X$ if there exists a $\delta > 0$ such that $f(x) \geq f(q), \forall x \in N_\delta(x)$.

A global maximum can be a local maximum, but the opposite is not true.

Theorem 9.1.4 (Necessary conditions for Interior Max)**.** *Let f be a real-valued, continuous function on $E \subseteq \mathbb{R}$. If f has a local maximum $x_0 \in int(E)$, and if $f'(x_0)$ exists, then $f'(x_0) = 0$.*

Proof. Suppose $x_0 \in int(E)$ is a local max. Then, there is a $\delta > 0$ such that $N_\delta(x_0) \subseteq int(E)$. Take a sequence $\{a_n\} \to 0$, $0 < a_n < \delta$. Then:

$$\frac{f(x + a_n) + f(x)}{a_n} \leq 0$$

Take a sequence $\{b_n\} \to 0$, such that $b_n < 0$ and $-b_n < \delta$. We have:

$$\frac{f(x_0 + b_n) - f(x_0)}{b_n} \geq 0$$

Since f is differentiable, we have:

$$\lim_{n\to\infty} \frac{f(x_0 + a_n) - f(x_0)}{a_n} = \lim_{n\to\infty} \frac{f(x_0 + b_n) - f(x_0)}{b_n} = f'(x_0) = 0$$

□

Theorem 9.1.5. *(Rolle's Theorem) Let $f : [a, b] \to \mathbb{R}$, f differentiable of (a, b). Then, if $f(a) = f(b)$, it exists a $c \in (a, b)$ such that $f'(c) = 0$*

Proof. If f is continuous, $f([a, b])$ is compact, so f has both a max x_M and a min x_m on $[a, b]$. Then $f(x_m) \leq f(a) \leq f(x_M)$. We have three cases:

1. $f(x_m) = f(a) = f(x_M)$, then the function is constant: $f'(x) = 0$
2. $f(x_m) < f(a) = f(b)$, then the minimum $x_m \in (a, b)$ and $f'(x_m) = 0$
3. $f(x_M) > f(a) = f(b)$, then the max $x_M \in (a, b)$ and $f'(x_M) = 0$

In each case, at least one point $c \in (a, b)$ is equal to 0. □

Theorem 9.1.6. *(Mean Value Theorem) Let f be a real-valued differentiable function on $E \subseteq \mathbb{R}$. If there are $a, b \in \mathbb{R}$ such that $a < b$, then it exists $c \in (a, b)$ such that $f'(c)(b - a) = f(b) - f(a)$, and:*

$$f'(c) = \frac{f(b) - f(a)}{b - a}$$

Proof. Define:

$$g(x) = f(x) - \frac{f(b) - f(a)}{b - a} \cdot (x - a)$$

Which is continuous and differentiable in x. Then:

$$g(a) = f(a) - \frac{f(b) - f(a)}{b - a} \cdot (a - a) = f(a)$$

$$g(b) = f(b) - \frac{f(b) - f(a)}{b - a} \cdot (b - a) = f(b) - f(b) + f(a) = f(a)$$

By Rolle's theorem, it exists a $c \in (a,b)$ such that $g'(c) = 0$:

$$g(c) = f(c) - \frac{f(b)-f(a)}{b-a} \cdot (c-a) = 0 \Rightarrow$$

$$g'(c) = f'(c) - \frac{f(b)-f(a)}{b-a} = 0$$

$$f'(c) = \frac{f(b)-f(a)}{b-a}$$

□

Corollary 9.1.6.1. *Suppose f is a real-valued function, and it is differentiable in (a,b). The following properties hold:*

1. *If $f'(x) \geq 0$, for all $x \in (a,b)$, then f is increasing*
2. *If $f'(x) \leq 0$, for all $x \in (a,b)$, then f is decreasing*
3. *If $f'(x) = 0$, for all $x \in (a,b)$, then f is constant*

Proof. Assume $a < x < y < b$, then it exists, by MVT, a $z \in (x,y)$ such that:

$$f(y) - f(x) = f'(z)(y-x)$$

Which is strictly positive (negative) when $f'(z) > (<) 0, \forall z \in (a,b)$. □

Theorem 9.1.7 (Cauchy's mean value theorem)**.** *Let f and g be two real-valued functions, continuous on $[a,b]$ and $g(a) \neq g(b)$. Both are differentiable on (a,b) and $g'(x) \neq 0, \forall x \in [a,b]$. Then, there is a $c \in (a,b)$ such that:*

$$\frac{f(b)-f(a)}{g(b)-g(a)} = \frac{f'(c)}{g'(c)}$$

Proof. Let $h(x) = f(x) - \frac{f(b)-f(a)}{g(b)-g(a)} \cdot g(x)$. Then:

$$\begin{aligned} h(a) =& f(a) - \frac{f(b)-f(a)}{g(b)-g(a)} \cdot g(a) = \\ &= \frac{f(a)g(b) - f(a)g(a) - f(b)g(a) + f(a)g(a)}{g(b)-g(a)} = \\ &= \frac{f(a)g(b) - f(b)g(a)}{g(b)-g(a)} = h(b) \end{aligned}$$

By Mean Value Theorem, there exists a $c \in (a,b)$ such that:

$$\begin{aligned} h'(c)(b-a) =& \Big[f'(c) - \frac{f(b)-f(a)}{g(b)-g(a)} g'(c)\Big](b-a) = 0 \\ & f'(c)(b-a) = g'(c)\frac{f(b)-f(a)}{g(b)-g(a)}(b-a) = \\ & \frac{f'(c)}{g'(c)} = \frac{f(b)-f(a)}{g(b)-g(a)} \end{aligned}$$

□

Theorem 9.1.8 (L'Hopital rule)**.** *Suppose f and g are both continuous, real-valued functions on $[a, b]$ and differentiable on (a, b). Suppose that:*

- $f(x) = g(x) = 0$
- $f'(x) \neq 0, \forall x$

As $x_n \to x \in (a, b)$, the limit of $\frac{f(x_n)}{g(x_n)}$ exists, then:

$$\lim_{x_n \to x} \frac{f(x_n)}{g(x_n)} = \lim_{x_n \to x} \frac{f'(x_n)}{g'(x_n)}$$

Proof. □

$$\lim_{x_n \to x} \frac{f(x_n)}{g(x_n)} = \lim_{x_n \to x} \frac{f(x_n) - 0}{g(x_n) - 0} =$$
$$\lim_{x_n \to x} \frac{f(x_n) - f(x)}{g(x_n) - g(x)} = \lim_{x_n \to x} \frac{\frac{f(x_n) - f(x)}{x_n - x)}}{\frac{g(x_n) - g(x)}{x_n - x}} = \frac{f'(x)}{g'(x)}$$

Definition 9.1.3. If f has a derivative f' on an interval, f' is a real-valued function. If f' is also differentiable, we denote it as:

$$f''(x) = \lim_{x_n \to x} \frac{f'(x_n) - f'(x)}{x_n - x}$$

f'' is called the **second derivative** for f. We can obtain derivatives $f^{(3)}, f^{(4)}, \ldots, f^{(n)}$ whenever they exist. f is r^{th}-order differentiable, if $f^{(r)}$ derivative exists. f is **smooth** if it is **infinitely differentiable**.

Theorem 9.1.9 (Taylor's Formula)**.** *Let $f : \mathbb{R} \to \mathbb{R}$ and assume that $f^(i)$ exists for $i = 1, 2, \ldots, n$. For any $x_0 \in \mathbb{R}$ and $h > 0$, we have:*

$$f(x_0 + h) = f(x_0) + \sum_{k=1}^{n-1} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k + r_n$$

where $r_n = \frac{f^{(n)}(z)}{n!}h^n$ for some $z \in (x_0, x_0 + h)$

Proof. □

Taylor's formula represents the value of $f(x)$ around x_0 by:

- $f(x_0)$, a constant number
- $f'(x_0)(x - x_0)$, a linear function
- $\frac{f''(x_0)}{2}(x - x_0)^2$, a quadratic function

- $r_n = \frac{f^{(n)}(z)}{n!}h^n$, a residual, where $x \in (x_0, x_0 + h)$ becomes smaller as $n \to \infty$.

Then, smooth functions can be approximated by finite polynomials. For example, when $n = 2$:

$$f(x_0 + h) \approx f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)(x - x_0)^2}{2!}$$

9.2 Differentiation with many variables

Definition 9.2.1. Let $f : \mathbb{R}^n \to \mathbb{R}^m$.

$$f(\mathbf{x}) = \begin{bmatrix} f_1(\mathbf{x}) \\ f_2(\mathbf{x}) \\ \vdots \\ f_m(\mathbf{x}) \end{bmatrix} = \begin{bmatrix} f_1(x_1, x_2, x_3, \ldots, x_n) \\ f_2(x_1, x_2, x_3, \ldots, x_n) \\ \vdots \\ f_m(x_1, x_2, x_3, \ldots, x_n) \end{bmatrix}$$

$f_1, f_2, \ldots, f_m$ are called **Component functions** of f. $f_i : \mathbb{R}^n \to \mathbb{R}$

The general idea is still that of approximating a function by an affine linear function. Just, if in the single variable case, we approach x from left and right, now we can do it in many directions.

Definition 9.2.2. The bfpartial derivative of f with respect to x_i, at x is defined as:

$$\frac{\partial f}{\partial x_i} = \lim_{t \to 0} \frac{f(x_1, \ldots, x_{i-1}, x_i + t, x_{i+1}, \ldots, x_n) - f(x_1, \ldots, x_i, \ldots, x_n)}{t}$$

whenever it exists. The matrix of the first partial derivatives of the component functions of f is called the **Jacobian** of f.

$$J(x) = \begin{bmatrix} \frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2}, & \ldots, & \frac{\partial f}{\partial x_n)} \end{bmatrix} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix}$$

Definition 9.2.3. The **directional derivative** of $f : \mathbb{R}^n \to \mathbb{R}^m$ in the direction of u at the point x is defined by:

$$Df(\mathbf{x}; \mathbf{u}) = \lim_{t \to 0} \frac{f(\mathbf{x} + t\mathbf{u}) - f(\mathbf{x})}{t}$$

Where $t \neq 0$ and $\mathbf{u} \neq 0$

If $\mathbf{u} = \mathbf{e}^i$, then $Df(\mathbf{x}; \mathbf{u}) = \frac{\partial f}{\partial x_i}$
In general:

$$D_i f(\mathbf{x}; \mathbf{u}) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} u_i$$

Notice, however, that, differently from the case on one variable, it is not true that differentiability implies continuity. Partial derivatives can exist and still, the function is not continuous. For example:

$$f(x,y) = \begin{cases} \frac{xy}{x^2+y^2} & \text{if } (x,y) \neq (0,0) \\ 0 & \text{if } (x,y) = (0,0) \end{cases}$$

Then the partial derivatives are $(0,0)$ in $(0,0)$ and $(\frac{y}{2x}, \frac{x}{2y}$ otherwise. Approaching $(0,0)$ along $x = y$, then $f_x(x,y) = f_y(x,y)$ if $x = y$ but $f(0,0) = 0$.

A more comprehensive definition of differentiability is the following:

Definition 9.2.4. A function $f : E \to \mathbb{R}^m$ where $E \subseteq \mathbb{R}^n$ is **differentiable** at $x \in E$ if it exists a matrix A_x such that:

$$\frac{\|f(x+h) - f(x) - A_x h\|}{\|h\|} \to 0$$

As $\|h\| \to 0$, where $h \in \mathbb{R}^n$. If f is differentiable at every $x \in E$, then f is differentiable on E and we can define the **total derivative** of f as the function $\mathbf{D}f(x) : E \to \mathbb{R}^{m \times n}$ such that:

$$\mathbf{D}f(x) = A_x, \quad \forall x \in E$$

Theorem 9.2.1. *A function $f : E \to \mathbb{R}^m$, where $E \subseteq \mathbb{R}^m$*

1. *f is differentiable at $x \in E$, then each of its component functions is differentiable at x*
2. *If f is differentiable at x, then the partial derivatives of the component functions exist at x, and the derivative of f at x equals the Jacobian:*

$$\mathbf{D}f(x) = J(x)$$

Proof. □

Theorem 9.2.2 (Young's theorem)**.** *Suppose f is a real-valued function on $E \subseteq \mathbb{R}^n$ and $\mathbf{D}^2 f(x^*)$ exists at $\bar{x} \in E$. Then:*

$$\frac{\partial f(x^*)}{\partial x_i \partial x_j} = \frac{\partial f(x^*)}{\partial x_j \partial x_i}$$

Then, $\mathbf{D}^2 f(x^*)$ is a symmetric matrix.

Proof. □

Theorem 9.2.3. *A function $f : E \to \mathbb{R}^n$. If the partial derivatives of all component functions exist, and they are continuous at x, then f is differentiable at x.*

Proof. □

Definition 9.2.5. Suppose f is a real-valued function on $E \subseteq \mathbb{R}^n$, and it is differentiable. The **gradient vector** of f is defined as follows:

$$\nabla f(x) = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{bmatrix} \in \mathbb{R}^n$$

The second derivative:

$$\mathbf{D}^2 f(x) = \mathbf{D}[\nabla f(x)] = \begin{bmatrix} \frac{\partial^2 f_1}{\partial x_1^2} & \frac{\partial^2 f_1}{\partial x_1 x_2} & \cdots & \frac{\partial^2 f_1}{\partial x_1 x_n} \\ \frac{\partial^2 f_2}{\partial x_2 x_1} & \frac{\partial^2 f_2}{\partial x_2^2} & \cdots & \frac{\partial^2 f_2}{\partial x_2 x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f_n}{\partial x_1 x_n} & \frac{\partial^2 f_n}{\partial x_n x_2} & \cdots & \frac{\partial^2 f_n}{\partial x_n^2} \end{bmatrix}$$

$\mathbf{D}^2 f(x)$ is called **Hessian Matrix** of f.

Theorem 9.2.4. *(Chain Rule) Suppose that $E \subseteq \mathbb{R}^n, B \subseteq \mathbb{R}^p$, $f : E \to \mathbb{R}^p$ is differentiable at $x \in E$ and $g : B \to \mathbb{R}^m$ at $y = f(x) \in B$. Let $h(x) = g \circ f(x)$. then h is differentiable at x and:*

$$\mathbf{D}h(x) = \mathbf{D}g(f(x)\mathbf{D}f(x)$$

Proof. □

Theorem 9.2.5. *Suppose $f : \mathbb{R}^n \to \mathbb{R}$ is differentiable. Then there exists a $c = (1-\lambda)a + \lambda b$, $\lambda \in (0,1)$ such that:*

$$f(b) - f(a) = \mathbf{D}f(c)(b-a)$$

Proof. Let $g(t) = f((1-t)a + tb)$ for $t \in \mathbb{R}$, so g is differentiable. Then:

$$g'(t) = \mathbf{D}f((1-t)a + tb)(b-a)$$

$g(0) = f(a)$ and $g(1) = f(b)$. Thus, it exists a $\lambda \in (0,1)$ such that:

$$g(1) - g(0) = g'(\lambda)(1-0)$$

Then:

$$f(b) - f(z) = \mathbf{D}f(\underbrace{(1-\lambda)a + \lambda b}_{c})(b-a)$$

□

Definition 9.2.6. A differentiable function $f : E \to \mathbb{R}^m$, where $E \subseteq \mathbb{R}^n$ is said to be **continuosly differentiable** if $\mathbf{D}f$ is continuous on E. We denote it as $f \in C^1$.

If $f \in C^\infty$, we say that f is **smooth**.

Theorem 9.2.6 (Taylor's Formula)**.** *Suppose that $f : \mathbb{R}^n \to \mathbb{R}$ is C^2. For $x, x+h \in \mathbb{R}^n$, there exists a $\lambda \in (0,1)$ such that:*

$$f(x-h) = f(x) + \mathbf{D}f(x)h + \frac{1}{2}h^T\mathbf{D}^2f(x+\lambda h)h$$

Proof. □

Suppose that $f(x,y) \in C^2$. Then, by Taylor's formula, the first order approximation at (x^*, y^*) is:

$$f(x^*, y^*) = f_x(x^*, y^*)(x - x^*) + f_y(x^*, y^*)(y - y^*)$$

9.2.1 Homogeneity

Definition 9.2.7. A real valued function $f : \mathbb{R}^n \to \mathbb{R}$ is **homogeneous of degree** $z \in Z$, if, for $t > 0$, we have:

$$f(tx_1, tx_2, \ldots, tx_n) = t^r f(x_1, x_2, \ldots, x_n)$$

When $r = 1$, f exhibits **constant returns to scale**.

Theorem 9.2.7. *Let f be a real-valued function of $\mathbb{R}^n$, and it is homogeneous of degree r. Then, for any $i = 1, 2, \ldots, n$, the partial derivative function $\frac{\partial f(x_1, x_2, \ldots, x_n)}{\partial x_i}$ is homogeneous of degree $r-1$.*

Proof. Let $t > 0$. Then since:

$$f(tx_1, tx_2, \ldots, tx_n) = t^r f(x_1, x_2, \ldots, x_n)$$

We have:

$$f(tx_1, tx_2, \ldots, tx_n) - t^r f(x_1, x_2, \ldots, x_n) = 0$$

Taking the derivative with respect to x_i, we have:

$$\frac{\partial f(tx_1, tx_2, \ldots, tx_n)}{\partial x_i} t - t^r \frac{\partial f(x_1, x_2, \ldots, x_n)}{\partial x_i} = 0$$

So:

$$\frac{\partial f(tx_1, tx_2, \ldots, tx_n)}{\partial x_i} = t^r \frac{\partial f(x_1, x_2, \ldots, x_n)}{\partial x_i} \cdot \frac{1}{t}$$

$$\frac{\partial f(tx_1, tx_2, \ldots, tx_n)}{\partial x_i} = t^{r-1} \frac{\partial f(x_1, x_2, \ldots, x_n)}{\partial x_i}$$

□

Theorem 9.2.8 (Euler's Theorem)**.** *Let f be a real-valued function of $\mathbb{R}^n$, and it is homogeneous of degree r and differentiable. Then at any $\bar{x} = (\bar{x_1}, \bar{x_2}, \ldots, \bar{x_n})$, we have.*

$$\sum_{i=1}^{n} \frac{\partial f(\bar{x_1}, \ldots, \bar{x_n})}{\partial x_i} \cdot \bar{x}_n = r f(\bar{x}_1, \ldots, \bar{x}_n)$$

Proof. Let $t > 0$, we have:

$$f(tx_1, tx_2, \ldots, tx_n) - t^r f(x_1, x_2, \ldots, x_n) = 0$$

Differentiating with respect to t, we have:

$$\sum_{i=1}^{n} \frac{\partial f(t\bar{x}_1, \ldots, t\bar{x}_n)}{\partial x_i} \cdot \bar{x}_i - rt^{r-1} f(x_1, x_2, \ldots, x_n) = 0$$

With $t = 1$, then we have Euler's formula. □

9.2.2 Implicit Function Theorem

An **explicit** function is a function of the form $y = f(x)$. An equation of the form $f(x, y) = c$ is called **implicit function**. However, it is not always true that an implicit function exists. Then, the question is under which conditions we can express y as an explicit function of x (or vice-versa).

A straightforward example in economics is that of indifference curves $u(x, y) = c$, If we can write $y = y(x)$, this implies how the consumption of y changes with the consumption of x to maintain the same level of utility c. If $y(\cdot)$ is differentiable, then we can take the derivative of $u(x, y(x))$ with respect to x and we have:

$$u_x(x, y(x)) + u_y(x, y(x))y'(x) = 0$$

At (x_0, y_0), we have:

$$\frac{dy(\cdot)}{dx} = -\frac{u_x(x_0, y_0)}{u_y(x_0, y_0)}$$

The simplest example of a function that cannot be written in explicit form is the equation of the circle (of radius 1), $x^2 + y^2 = 1$.

Theorem 9.2.9 (Implicit Function Theorem: 2 variables case)**.** *Let $(x_0, y_0) \in \mathbb{R}^2$ be a point such that $f(x_0, y_0) = c$. Then:*

1. *If f is C^1 (continuously differentiable)*
2. *if $f_y(x_0, y_0) \neq 0$*

Then $f(x, y) = c$ define a C^1 implicit function $y = y(x)$ in some neighborood of (x_0, y_0) and:

$$\frac{dy(x)}{dx} = -\frac{f_x(x_0, y_0)}{f_y(x_0, y_0)}.$$

Notice that this result holds only in the neighborhood of (x_0, y_0).

Proof. The first step is that of constructing an implicit function. Assume that $f_y(x_0, y_0)_\alpha > 0$. Since f is continuous, it exists $a, b > 0$ such that $f_y(x, y) > fracalpha2$, $\forall x, y \in B_{a,b} = [x_0 + a, x_0 - a] \times [y_0 + b, y_0 - b]$. Because $f_y(x, y) > 0$, f is strictly increasing in y. Furthermore, since f is continuous, we can construct an open neighborhood around $(x_0, y_0 + b) = D_1$ and one around $(x_0, y_0 - b) = D_1$ such that:

$$f(x, y) > 0, (x, y) \in D_1$$

$$f(x, y) < 0, (x, y) \in D_2$$

And it exists a $c > 0$ such that, $\forall x \in (x_0 - c, x_0 + c)$:

$$f(x, y_0 + b) > 0$$

$$f(x, y_0 - b) < 0$$

Fix x. Because $f(x, y)$ is strictly increasing in y, it exists an unique $y \in (y_0 - b, y_0 + b)$ such that $f(x, y) = 0$. Name this mapping $y(x) = y$. This is the implicit function around x_0.

Let's prove that the implicit function is continuous. Pick $\hat{x} \neq x \in (x_0 - c, x_0 + c)$, and $\hat{y} = y(\hat{x})$ and $y = y(x)$, so:

$$f(x, y) = f(\hat{x}, \hat{y})$$

By the generalized mean value theorem, it exists $\lambda \in (0, 1)$ such that:

$$\begin{aligned} 0 = f(x, y) - f(\hat{x}, \hat{y}) = \\ \mathbf{D}f(x^\lambda, y^\lambda) \begin{bmatrix} x - \hat{x} \\ y - \hat{y} \end{bmatrix} = \\ f_x(x^\lambda, y^\lambda)(x - \hat{x}) + f_y(y - \hat{y})(y - \hat{y}) \end{aligned}$$

Where x^λ, y^λ are convex combinations of $(x, \hat{x})$ and $(y, \hat{y})$. Rearranging, we have:

$$y(x) - y(\hat{x}) = y - \hat{y} = -\frac{f_x(x^\lambda, y^\lambda)}{f_y(x^\lambda, y^\lambda)}(x - \hat{x})$$

Since f is continuous in $B_{a,b}$, $\max f$ exists, and let'sdefine the maximum as M. Also $f_y \geq \frac{\alpha}{2}$ in $B_{a,b}$. Hence:

$$|y(x) - y(\hat{x})| = \left|\frac{f_x(x^\lambda, y^\lambda}{f_y(x^\lambda, y^\lambda}\right||x - \hat{x}| \leq \frac{2M}{\alpha}|x - \hat{x}|$$

As $x \to \hat{x}$, $y(x) \to y(\hat{x})$ and therefore $y(\cdot)$ is continuous.

The final step is to prove that the implicit function is differentiable. Since:

$$y(x) - y(\hat{x}) = y - \hat{y} = -\frac{f_x(x^\lambda, y^\lambda)}{f_y(x^\lambda, y^\lambda)}(x - \hat{x})$$

Dividing both sides by $x - \hat{x}$ and taking the limit $x \to \hat{x}$, we have:

$$\lim_{\hat{x}\to x} \frac{y(x) - y(\hat{(x)}}{x - \hat{x})} = -\lim_{\hat{x}\to x} \frac{f_x(x^\lambda, y^\lambda)}{f_y(x^\lambda, y^\lambda)} = \frac{f_x(x,y)}{f_y(x,y)}$$

Because, as $\hat{x} \to x$, $(x^\lambda, y^\lambda) \to (x, y)$. □

Theorem 9.2.10 (Implicit Function Theorem: general version)**.** *Let $f : X \times Y \to \mathbb{R}^n$ where $X \subseteq \mathbb{R}^m$ and $Y \subseteq \mathbb{R}^m$ and $Y \subseteq \mathbb{R}^n$. Let $(x^0, y^0) \in \mathbb{R}^{m\times n}$ be a point such that $f(x^0, y^0) = c$, where f is C^1. If $\mathbf{D}_y f(x^0, y^0)$ is a invertible $n \times n$. Then there are open sets U, V with $x^0 \in U \subseteq X$ and $y^0 \in V \subseteq Y$ and a C^1 onto function $y : U \to V$ such that:*

$$f(x, y(x)) = c, \forall x \in U$$

and, at (x_0, y_0)

$$\mathbf{D}_x y(x) = -[\mathbf{D}_y f(x, y)]^{-1} \mathbf{D}_x f(x, y)$$

or:

$$\underbrace{\begin{bmatrix} \frac{\partial y_1(x)}{\partial x_1} & \cdots & \frac{\partial y_1(x)}{\partial x_m} \\ \vdots & \vdots & \vdots \\ \frac{\partial y_n(x)}{\partial x_1} & \cdots & \frac{\partial y_n(x)}{\partial x_m} \end{bmatrix}}_{n\times m} = -\underbrace{\begin{bmatrix} \frac{\partial f_1(x,y)}{\partial y_1} & \cdots & \frac{\partial f_1(x,y)}{\partial y_n} \\ \vdots & \vdots & \vdots \\ \frac{\partial f_n(x,y)}{\partial y_1} & \cdots & \frac{\partial f_n(x,y)}{\partial y_n} \end{bmatrix}^{-1}}_{n\times n} \underbrace{\begin{bmatrix} \frac{\partial f_1(x,y)}{\partial x_1} & \cdots & \frac{\partial f_1(x,y)}{\partial x_m} \\ \vdots & \vdots & \vdots \\ \frac{\partial f_n(x,y)}{\partial x_1} & \cdots & \frac{\partial f_n(x,y)}{\partial x_m} \end{bmatrix}}_{n\times m}$$

Proof. □

Chapter 10

Concave Functions

10.1 Convex sets

Definition 10.1.1. The set $A \subseteq \mathbb{R}^n$ is **convex** if for any $x, y \in A$ and $\lambda \in (0,1)$, $\lambda x + (1-\lambda)y \in A$. It is **strictly convex** if $\lambda x + (1-\lambda)y \in int(A)$ for $\lambda \in (0,1)$.

Theorem 10.1.1. *If $A_1, A_2, \ldots, A_n \in \mathbb{R}^n$ are convex, then $A_1 \times A_2 \times A_n$ is convex*

Proof. □

Theorem 10.1.2. *The intersection of convex sets is convex*

Proof. □

The union of two separated convex sets is not convex.

Definition 10.1.2. For any $x_1, x_2, \ldots, x_n \in A$, y is their **convex combination** if for some $\lambda_i \in [0,1], i = 1, 2, \ldots, n$ such that $\sum_{i=1}^{n} \lambda_i = 1$ and:

$$y = \sum_{i=1}^{n} \lambda_i x_i$$

Then, if A contains every convex combination of its elements, then it must be convex.

Theorem 10.1.3. *A set is convex if and only if it contains every convex combination of its elements.*

Proof. □

10.2 Concave Functions

Definition 10.2.1. Suppose $E \subseteq \mathbb{R}^n$ is convex. A function $f : E \to \mathbb{R}$ is **concave** if:

$$f(\lambda x + (1-\lambda)x') \geq \lambda f(x) + (1-\lambda)f(x'), \quad \forall \lambda \in [0,1], \forall x.x' \in E.$$

If the equality is strict, the function is said to be **strictly concave**.

The definition of a convex function is symmetric:

Definition 10.2.2. Suppose $E \subseteq \mathbb{R}^n$ is convex. A function $f : E \to \mathbb{R}$ is **convex** if:

$$f(\lambda x + (1-\lambda)x') \leq \lambda f(x) + (1-\lambda)f(x'), \quad \forall \lambda \in [0,1], \forall x, x' \in E.$$

If the equality is strict, the function is said to be **strictly convex**.

Another way of defining concavity is to say:

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

Where the first term is the **supporting affine function** at $f(x)$ and $x, x' \in E$.

Besides, the graph of a function can be written as:

$$\Big\{(x, y) \in \mathbb{R}^{n+1} : x \in E, y = f(x)\Big\}$$

Definition 10.2.3. Let $f : E \to \mathbb{R}$, E is convex. We call the **hypograph** of f as:

$$\Big\{x, y) \in \mathbb{R}^{n+1} : x \in E, y \leq (f(x)\Big\}$$

and the **epigraph** as:

$$\Big\{x, y) \in \mathbb{R}^{n+1} : x \in E, y \geq f(x)\Big\}$$

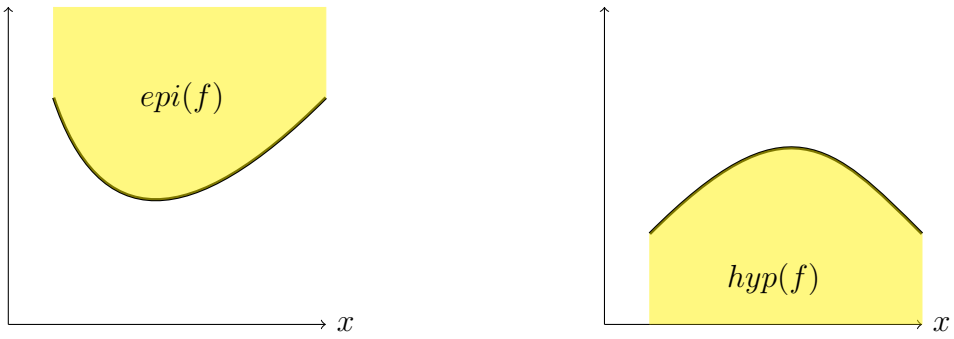


Figure 10.1: The epigraph and the hypograph

Theorem 10.2.1. *Let $f : E \to \mathbb{R}$, $E \subseteq \mathbb{R}^n$ is convex. Then:*

1. *f is concave if and only if its hypograph is convex*
2. *f is convex if and only if its epigraph is convex*

Proof. Let's see 1) ($\Rightarrow$) Take any $(x, y), (x', y')$ in hypograph, and $\lambda \in [0, 1]$. f is concave, then:

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

Then, $y^\lambda \leq f(x^\lambda)$. $(x^\lambda, y^\lambda) \in$ hypograph, and the hypograph is convex.

($\Leftarrow$) Take $x, x' \in E$. Then:

$$(f(x), x), f(x'), x') \in \text{hypograph}$$

Hypograph is convex, then we can write:

$$(\lambda f(x') + (1-\lambda)f(x), \lambda x' + (1-\lambda)x) \in \text{hypograph}$$

So:

$$\lambda f(x') + (1-\lambda)f(x) \leq f(\lambda x' + (1-\lambda)x)$$

then f is concave.

The proof of 2) is symmetric. □

Theorem 10.2.2. *Let $f : E \to \mathbb{R}$ is C^1 where $E \subseteq \mathbb{R}^m$ is convex.f is concave if and only if:*

$$f(y) \leq f(x) + \mathbf{D}f(x)(y - x) \ \ \forall (x, y) \in E.$$

f is strictly concave if and only if the inequality is strict.

Proof. □

This is a generalization of what was said above. Indeed, $\mathbf{D}f(x)$ is a $1 \times n$ matrix, x is a $n \times 1$ vector. If $n = 1$, we have:

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

Theorem 10.2.3. *Let $f : E \to \mathbb{R}$, f is C^2, $E \subseteq \mathbb{R}^n$ is convex. Then:*

1. *f is concave if and only if $\mathbf{D}^2 f(x)$ is **negative semidefinite** for all $x \in int(E)$*
2. *If $\mathbf{D^2} f(x)$ is **negative definite** for all $x \in int(E)$, then f is concave.*

Proof. □

When $n = 1$, this simply means that:

- $f''(x) \leq 0$, f is concave
- $f"(x) < 0$, f is strictly concave.

To check for negative-(semi)definiteness, a way developed by Gerard Debreu in 1952 is that of looking at the determinant of the n-leading principal minors (across the diagonal) of a square matrix, namely, the n $k \times k$ submatrices of A, that contains the first k rows and columns.

Theorem 10.2.4. *A $n \times n$ symmetric matrix A is:*

1. *Negative Definite if and only if $(-1)^k|A_k| > 0$ for all $k = 1 \ldots, n$.*
2. *Positive Definite if and only if $|A_k| > 0$ for all $k = 1 \ldots, n$*

Proof. □

Notice that for negative definiteness, this means that $|A_1| < 0, |A_2| > 0, |A_3| < 0$ and so on...

For semi-definiteness, things are more complicated. If, for example, a 3×3 matrix has $|A_1| > 0, |A_2| > 0, |A_3| = 0$, this matrix can be positive definite or indefinite. In this case, one has to look to **all** the principal minors, not only the leading one. If one of them is negative, the matrix is indefinite. If none is negative, the matrix is **positive semi-definite**.

For **negative semi-definiteness**, if none of the minors of odd order is positive and none of the minors of even order is negative, then the matrix is negative semi-definite. Otherwise, it is indefinite.

Finally, two important properties of concave functions:

Theorem 10.2.5. *If a function $f : E \to \mathbb{R}$ is concave, then it is continuous on $int(E)$*

Proof. □

Theorem 10.2.6. *If a function $f : E \to mathbbR$ is concave, then f is differentiable at every point on E except possibly at a set of points of Lebesgue measure zero.*

Proof. □

10.3 Quasi-Concave Functions

Quasi-concavity is a property less strict than concavity but which maintains some desirable properties and behavior of Concave functions.

Definition 10.3.1. Let $f : E \subseteq \mathbb{R}^n$ be convex. A function $f : E \to \mathbb{R}$ is **quasi-concave** if:

$$f(\lambda x + (1-\lambda)x') \geq \min\{f(x'), f(x)\} \ \ \forall x, x' \in E, \forall \lambda \in [0,1]$$

If the inequality is strict, for all $\lambda \in (0,1)$, then we have **strict quasi-concavity**.

Symmetrically, A function $f : E \to \mathbb{R}$ is **quasi-convex** if:

$$f(\lambda x + (1-\lambda)x') \leq \max\{f(x'), f(x)\} \ \ \forall x, x' \in E, \forall \lambda \in [0,1]$$

If the inequality is strict, for all $\lambda \in (0,1)$, then we have **strict quasi-convexity**.

Then, if a function is quasi-concave, whenever $f(x), f(x') \geq t$, we have:

$$f(\lambda x + (1-\lambda)x') \geq t$$

A monotonic function is both quasi-concave and quasi-convex.

Theorem 10.3.1. *Let $f : E \to \mathbb{R}$, $E \subseteq \mathbb{R}^n$ is convex. Then:*

1. *If f is concave, then it is quasi-concave*
2. *If f is convex, then it is quasi-convex*

Proof. Let's see that concavity implies quasi-concavity. If f is concave, then:

$$\begin{aligned} f(\lambda x + (1-\lambda)x') &\geq \lambda f(x) + (1-\lambda) f(x') \;\; \forall x, x' \in E, \forall \lambda \in [0,1] \\ &\geq \lambda \min\{f(x), f(x)'\} + (1-\lambda)\{f(x), f(x')\} \\ &\geq \min\{f(x), f(x')\} \end{aligned}$$

This means □

The converse of the result is not true

Theorem 10.3.2. *f is quasi-concave and $\phi : \mathbb{R} \to \mathbb{R}$ is a non-decreasing function, then $\phi \circ f$ is also quasi-concave. In particular, any monotone transformation of a concave function is a quasi-concave function.*

Proof. Pick x, y and $\lambda \in [0,1]$. Because f is quasi-concave, we have:

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

Because ϕ is non-decreasing, then:

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

So $\phi \circ f$ is quasi-concave. □

This last result is important because concavity may not be preserved under monotone transformation.

Another way of defining quasi-concavity involves the notion of **contour set**.

Definition 10.3.2. Take $f : E \to \mathbb{R}$. For each $\alpha \in \mathbb{R}$, the **Upper contour set** of f is:

$$U_\alpha = \Big\{ x \in E : f(x) \geq \alpha \Big\}$$

Symmetrically, the **Lower Contour Set** of f is:

$$L_\alpha = \Big\{ x \in E : f(x) \leq \alpha \Big\}$$

This last definition is extremely useful in Microeconomics. Defining a consumer's utility by $u(x,y)=\alpha$, then:

- $(x,y)\in U_\alpha$ if and only if $u(x,y)\geq\alpha$
- $(x,y)\in L_\alpha$ if and only if $u(x,y)\leq\alpha$
- $L_\alpha\cap U_\alpha=\{(x,y)\in\mathbb{R}^2_+ : u(x,y)=\alpha\}$

Theorem 10.3.3. *Suppose that $f:E\to\mathbb{R}$, where $E\subseteq\mathbb{R}^n$ is convex. Then:*

1. *f is quasi-concave if and only if the upper contour set of f is convex*
2. *f is quasi-convex if and only if the lower contour set f is convex*

Proof. Let's see 1) ($\Rightarrow$). f is quasi-concave. Fix α such that:

$$f(x),f(x')\geq\alpha$$

Since f is quasi-concave, then:

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

$\forall\lambda\in(0,1)$. Therefore, $\lambda x+(1-\lambda)x'\in U_\alpha$.

($\Leftarrow$) Suppose U_α is convex. For any $x,x'\in U_\alpha$, any convex combination $\lambda x+(1-\lambda)x'\in U_\alpha$. Define $\alpha=\min\{f(x),f(x')\}$. Then:

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

□

Theorem 10.3.4. *Suppose $f:E\to\mathbb{R}$ is C^1, where $E\subseteq\mathbb{R}^n$ is convex. Then:*

1. *f is quasi-concave if and only if:*

$$f(y)\geq f(x)\Rightarrow\mathbf{D}f(x)(y-x)\geq 0$$

2. *If, $x\neq y$ and $f(y)\geq f(x)\Rightarrow\mathbf{D}f(x)(y-x)>0$, then f is strictly quasi-concave*

Proof. □

Theorem 10.3.5. *(Arrow and Enthoven Theorem) Suppose that $f:E\to\mathbb{R}$ is C^2 where $E\subseteq\mathbb{R}^n$ is convex. Then:*

1. *If f is quasi concave on E, we have $(-1^k|C_k(x)|\geq 0$ for every $k=1,2,\ldots,n$.*
2. *If $(-1^k(C_k(x)|>0$, for every $k=1,2,\ldots,n$, then f is quasi-concave on E*

And:

$$C_k(x)=\begin{bmatrix}0 & \frac{\partial f(x)}{\partial x_1} & \cdots & \frac{\partial f(x)}{\partial x_k}\\ \frac{\partial f(x)}{\partial x_2} & \frac{\partial^2 f(x)}{\partial x_1^2} & \cdots & \frac{\partial^2 f(x)}{\partial x_1 x_k}\\ \vdots & \vdots & \ddots & \vdots\\ \frac{\partial f(x)}{\partial x_k} & \frac{\partial^2 f(x)}{\partial x_k x_1} & \cdots & \frac{\partial^2 f(x)}{\partial x_k x_k}\end{bmatrix}$$

Is called ***Bordered Hessian Matrix****.*

Proof. □

Chapter 11

Optimization I

An optimization problem (max/min) has the following form:

$$v(\theta) = \max_{x \in \Gamma(\theta)} f(x, \theta)$$

Where $\Gamma(\theta) \subseteq \mathbb{R}^n$ is the constraint correspondence, θ are parameters, x are variables, f is the objective function, and v is the value function. To denote the set of optimal solutions, we write:

$$\arg\max_{x \in \Gamma(\theta)} f(c, \theta) = \{x \in \Gamma(\theta) : f(x, \theta) \geq f(y, \theta), \forall y \in \Gamma(\theta)\}$$

The questions to be addressed are:

- The existence and the uniqueness of the solutions
- the characterization of the solutions
- how to find a solution
- Comparative statics

The standard assumptions are usually:

- f is continuous
- $\Gamma(\theta)$ is non-empty
- $\Gamma(\theta)$ is compact, so, by Weierstraß theorem, a solution exists
- Since $\Gamma(\theta) \subseteq \mathbb{R}^n$, this is closed and bounded.

11.1 Unconstrained Optimization

Unconstrained optimization simply means that constraints are not binding. Namely, it is possible to move away from x^* without leaving Γ.

Definition 11.1.1. Let f be a real-values continuous function on $E \subseteq \mathbb{R}^n$. x^* is a local maximum of f if there exists a $\delta > 0$ such that $f(x^*) \geq f(x)$, $\forall x \in N_\delta \cap \Gamma$. x^* is a global maximum if $f(x^*) \geq f(x)$, $\forall x \in \Gamma$.

Theorem 11.1.1. *Let f be a real-valued continuous function on $E \subseteq \mathbb{R}$. If f has a local maximum (or minimum) at $x_0 \in int(\Gamma)$, and if $f(x_0)$ exists, then the **First-order conditions** holds, i.e. $f'(x_0) = 0$*

Proof. □

Notice that this condition characterizes only **interior** optima. Furthermore, this can be generalized to $\mathbf{R}^n$.

Theorem 11.1.2. *Let f be a real-valued C^1 on $E \subseteq \mathbb{R}^n$, If f has a local maximum or minimum at $x^* \in int(\Gamma)$, then $\mathbf{D}f(x^*) = 0$.*

Proof. Suppose x^* is an interior local max. Then, it exists $\delta > 0$ such that, $\forall x \in N_\delta(x^*), f(x^*) \geq f(x)$. For each $i = 1, 2, \ldots, n$ define:

$$h(t) = f(x^* + t\mathbf{e}^i)$$

$h(0) = f(x^*)$ and $h(0) \geq h(t)$ for $|t| < \delta$ for some $\delta > 0$. Take a sequence $t_k \to 0$, then:

$$\frac{h(t_k) - h(0)}{t_k} = \frac{f(x^* + t_k{}^i) - f(x^*)}{t_k}$$

and:

$$h'(0) = \frac{\partial(x^*)}{\partial x_i}, \forall i = 1, 2, \ldots, n$$

Because $\mathbf{D}f(x^*)$ exists:

$$\frac{\partial(x^*)}{\partial x_i} = h'(0) = \lim_{t_k \to 0^+} \frac{h(t_k) - h(0)}{t_k} = \lim_{t_k \to 0^-} \frac{h(t_k) - h(0)}{=} 0$$

Then:

$$\mathbf{D}f(x^*) = 0$$

□

FOCs are **necessary but not sufficient** for local maximum and minima. Besides, they cannot identify optima on the boundary set of E.

Second-order necessary conditions for local maxima involve the definiteness of matrices.

Theorem 11.1.3. *Suppose $f : E \to \mathbb{R}$ is C^2 and $x^* \in int(\Gamma)$. Then:*

1. *If x^* is a local maximum, then $\mathbf{D}^2 f(x^*)$ is **negative definite***
2. *If x^* is a local minimum, then $\mathbf{D}^2 f(x^*)$ is **positive definite***

Proof. If x^* is a local max, then it exists a δ such that $\forall y \in N_\delta$, $f(x^*) \geq f(y)$, and $\mathbf{D}f(x^+) = 0$. Take $y = x + \epsilon z$, where $z \neq 0$, and $\epsilon \|z\| < \delta$. By Taylor's expansion:

$$f(x^* + \epsilon z) - f(x^*) = \underbrace{\epsilon \mathbf{D}f(x^*)z}_{0} + \frac{1}{2}\epsilon^2 \underbrace{z^T \mathbf{D}^2 f(x^* + \alpha\epsilon z))z}_{\leq 0, \forall \epsilon < \frac{\delta}{\|z\|}} \leq 0$$

For some $\alpha \in [0, 1]$. As $\epsilon \to 0$:

$$z^T \mathbf{D}^2 f(x^* + \alpha\epsilon z))z \to z^T \mathbf{D}^2 f(x^*)z \leq 0$$

□

Theorem 11.1.4. *Suppose $f : E \to \mathbb{R}$ is C^2 and $x^* \in int(\Gamma)$. Then:*

1. *If $\mathbf{D}f(x^*) = 0$ and $\mathbf{D}^2 f(x^*)$ is **negative definite**, then x^* is a strict local maximum*
2. *If $\mathbf{D}f(x^*) = 0$ and $\mathbf{D}^2 f(x^*)$ is **positive definite**, then x^* is a strict local minimum*

If $\mathbf{D}^2 f(x^*)$ is negative (or positive) semidefinite, we cannot conclude anything on the nature of x^*.

Proof. Because f is C^2, if $z^t \mathbf{D}^2 f(x^*)z < 0$, it exists $\delta > 0$ such that $z^T \mathbf{D}^2 f(x)z < 0$ for $x \in N_\delta(x^*)$. By Taylor's formula and the FOCs:

$$f(x) - f(x^*) = \mathbf{D}f(x^*)(x - x^*) + \frac{1}{2}(x - x^*)^T \mathbf{D}^2 f(x^+ + (1 - \lambda)(x - x^*))(x - x^*) < 0$$

for $x \in N_\delta(x^*)$. Because $\mathbf{D}^2$ is continuous. Then:

$$\mathbf{D}^2 f(x^*) < 0 \Rightarrow \mathbf{D^2} f(x) < 0, \forall x \in N_\delta(x^*)$$

□

The sufficient second-order conditions that allow us to find all interior local optima involve the concavity of functions.

Theorem 11.1.5. *Suppose $f : E \to \mathbb{R}$ is C^2, E is convex, and f is concave. Then:*

1. *$x^* \in int(\Gamma)$ is a global maximum if and only if $\mathbf{D}f(x^*) = 0$*
2. *the set of optimal solutions is also convex*

If f is strictly concave, x^* such that $\mathbf{D}f(x^*)$ is the unique global max.

Proof. Let's see 1). Notice that a global max is a local max. So we need only to prove ($\Rightarrow$). By Taylor's formula:

$$f(x) - f(x^*) = \underbrace{\mathbf{D}f(x^*)(x - x^*)}_{0} + \underbrace{\frac{1}{2}(x - x^*)^T \mathbf{D}^2 f(\lambda x^* + (1-\lambda)x)(x - x^*)}_{\leq 0 \text{ by concavity}} \leq 0$$

$\lambda \in (0, 1)$.

Let's see 2). If x_1, x_2 are optimal, then $f(x_1) = f(x_2) = \max_x f(x)$. By concavity:

$$f(\lambda x_1 + (1-\lambda)x_2) \geq \lambda f(x_1) + (1-\lambda) f(x_2), \forall \lambda \in (0, 1)$$

Since E is convex, $\lambda x_1 + (1-\lambda)x_2$ is feasible and optimal. □

11.2 Optimization with Equality Constraints

This is a problem of the type:

$$\begin{array}{ll} \max\limits_{x} & f(x, y) \\ \text{s.t.} & g_i(x, y) = c \quad i = 1, 2, \ldots, m \end{array} \tag{11.1}$$

or:

$$\begin{array}{ll} \min\limits_{x} & f(x, y) \\ \text{s.t.} & g_i(x, y) = c \quad i = 1, 2, \ldots, m \end{array} \tag{11.2}$$

$f : E \subseteq \mathbb{R}^n$ is the objective function. g is the constraint. Both are C^1. We can write the Jacobian Matrix of the constraints as follows:

$$\mathbf{D}g(x_1, \ldots, x_n) = \begin{bmatrix} \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} & \cdots & \frac{\partial g_1}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g_m}{\partial x_1} & \frac{\partial g_m}{\partial x_2} & \cdots & \frac{\partial g_m}{\partial x_n} \end{bmatrix}$$

Constraints must satisfy the so-called **non-degenerate constraints qualifications** (NDCQ), at $x^* \in \Gamma$. Namely, the $m \times n$ Jacobian matrix must have a $m \times m$ invertible submatrix.

Another way of seeing constraint qualification is the following: the rank of the Jacobian matrix of $g(\cdot)$ must be equal to the number of constraints. This can be easily satisfied in the case of a linear function. In the case of a non-linear function, it means that the matrix, evaluated at x^*, must be different from the zero-matrix (whose rank is zero).

Theorem 11.2.1. *Let f, g be real-valued continuous function on $E \subseteq \mathbb{R}^n$, and C^1. x^* is a local maximum, and NDCQ are satisfied at x^*. Then, there exists $\lambda_1, \ldots, \lambda_m$ (one for each constraint), such that we have the* ***first order conditions****:*

$$\mathbf{D}L(\mathbf{x}) = \mathbf{D}f(x^*) + \sum_{i=1}^{n} \lambda_i g_i(x^*) = (0, \ldots, 0)$$

L is called ***Lagrangian*** *and $\lambda_1, \ldots, \lambda_n$ are called* ***Lagrangian multipliers****.*

Proof. □

A "cookbook" procedure to solve equality-constrained optimization problems is the following. The Lagrangian method involves three steps:

1. The first step is to set up the Lagrangian function:

$$\mathcal{L}(x_1, \ldots, x_n, \lambda) = f(x_1, \ldots, x_n) + \sum_{i=1}^{m} \lambda_i g_i(x_1, \ldots, x_n) \;\; i = 1, \ldots, n$$

2. We find all the **critical points** of $\mathcal{L}$. This is the set of points such that

$$D\,\mathcal{L}(\mathbf{x}, \lambda) = 0$$

   This means to solve the following system with $(n + k)$ equations (n-variables and k constraint):

$$\begin{cases} \frac{\partial \mathcal{L}}{\partial x_j} = 0 \;\; j = 1, \ldots, n \\ \frac{\partial \mathcal{L}}{\partial \lambda_i} = 0 \;\; ß = 1, \ldots, k \end{cases}$$

3. We evaluate f at each critical point. The points at which f attains the maximum (or the minimum) value are the optimal points.

Notice, however, that the FOCs are only **necessary** conditions for local max when NDCQ holds. They are not **sufficient**.

11.3 Optimization with Inequality Constraints

This is a problem of the type:

$$\begin{array}{ll} \max\limits_{x} & f(x, y) \\ \text{s.t.} & g_i(x, y) \geq 0 \quad \text{i=1,2,...,n} \end{array} \tag{11.3}$$

An inequality constraints $g_i(x) \geq 0$ is **binding** at x^* if $g_i(x^*) = 0$. Otherwise, it is **slack**.

If the m constraints are binding at x^*, NDCQ holds if the matrix:

$$\mathbf{D}g(x_1,\ldots,x_n) = \begin{bmatrix} \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} & \cdots & \frac{\partial g_1}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g_m}{\partial x_1} & \frac{\partial g_m}{\partial x_2} & \cdots & \frac{\partial g_m}{\partial x_n} \end{bmatrix}$$

has full rank.

Theorem 11.3.1 (Karush, Kuhn and Tucker)**.** *Suppose f, g_i are C^1 functions and $x^* \in int(E)$ is a local max for the optimization problem. Suppose one of the following conditions hold for x^*:*

1. *All binding constraints functions are affine (linear):*

$$g_i(\boldsymbol{x}) = \sum_{i=1}^{n} a_{ij}x_j + b_i$$

2. *for all binding constraints, NDCQ is satisfied.*

Then there are non-negative $\lambda_1,\ldots,\lambda_n$, (each for any constraint), such that the FOCs:

$$\mathbf{D}L(\mathbf{x}) = \mathbf{D}f(x^*) + \sum_{i=1}^{n} \lambda_i g_i(x^*) = (0,\ldots,0)$$

*And the **complementary slackness conditions**:*

$$\lambda_i g_i(x^*) = 0 \;\; \forall i = 1,\ldots,n$$

*$\lambda_i,\ldots,\lambda_n$ are called **Karush-Kuhn-Tucker multipliers**.*

Proof. □

Example 11.3.1. *A quick cookbook on how to solve a KKT optimization problem.*

$$\max f(x_1,\ldots,x_n)$$

$$s.t$$

$$g_i(x_1,\ldots,x_n) \geq 0, \forall i = 1\ldots,n$$

These are the main steps to solve this problem.

1. *Set up the Lagrangian*
2. *Check if f, g are C^1*
3. *Check for NDCQ. The Jacobian matrix of constraints must be full rank for all the binding constraints.*

If there is an interior solution, then $\lambda_1, \ldots, \lambda_n$ must satisfy:

- *FO(N)C for all $x_1, \ldots, x_n$*
- *KKT multipliers constraint, namely $\lambda_1, \ldots, \lambda_n \geq 0$*
- *Complementary Slackness: $\lambda_i g_i(x^*) \geq 0, \forall i = 1, \ldots, n$*
- *Feasibility Constraints, i.e., the non-negativity of the constraints*

This means that if, for instance, there are 2 variables and 3 constraints, there are 11 inequalities to solve.

11.3.1 Envelope Theorem

Let's write the consumer problem in the following way:

$$v(p, w) = \max_{x \in \mathbb{R}_+} u(x)$$

$$\text{s.t}$$

$$p \cdot x \leq w$$

We are interested in the effect of a change of p and w on $v(p, w)$. We assume f and g are C^1 functions in order to apply KKT. Defining $(p, w) = \theta$, we can generalize the problem as:

$$v(\theta) = \max_{x \in \mathbb{R}_+} f(x, \theta)$$

$$\text{s.t}$$

$$g_i(x, \theta) = 0 \ \ \forall i = 1, 2, \ldots, n$$

$\theta \in \Theta$ is a family of parameters. For each θ, let $x^*(\theta)$ be a global maximum, then:

$$v(\theta) = f(x^*(\theta), \theta)$$

Fixing x, $f(x, \theta)$ is a function of θ. $v(\theta)$, the **value function**, is the **upper envelope** of $\{f(x, \theta\}$. Since $x^*(\theta)$ is continuous, then it changes smoothly as θ changes. So, also $v(x^*(\theta), \theta)$ also changes smoothly.

Assume x^* is differentiable, then v is differentiable, and:

$$\frac{dv(\theta)}{d\theta} = \underbrace{\frac{\partial f(x^*(\theta), \theta)}{\partial \theta}}_{\text{direct effect}} + \underbrace{\frac{\partial f(x^*(\theta), \theta)}{\partial x} \frac{\partial x^*(\theta)}{\partial \theta}}_{\text{indirect effect}}$$

by the chain rule.

If $\forall \theta$, $x(\theta)$ satisfies FOC, so that:

$$\frac{\partial f(x^*(\theta), \theta)}{\partial x} = 0$$

Then:

$$\frac{dv(\theta)}{d\theta} = \frac{\partial f(x^*(\theta), \theta)}{\partial \theta}$$

Theorem 11.3.2 (Envelope Theorem)**.** *Suppose $\theta \in int(\Theta)$ and the global max $x^*(\cdot)$ is characterized by FOCs in a neighborhood of θ. If $x^*(\cdot)$ is differentiable at θ, then:*

$$\frac{\partial v(\theta)}{\partial \theta_j} = \frac{\partial L(x^*, \lambda, \theta)}{\partial \theta_j} = \frac{f(x^*, \theta)}{\partial \theta_j} + \sum_{i=1}^{m} \lambda_i \frac{\partial g_i(x^*, \theta)}{\partial \theta_j}, \quad \forall j = 1, \ldots, n$$

Proof. Let's consider the special case of $m = 1$, $n = 2$ and $\theta \in \mathbb{R}$. Let's take the FOCs at $(x^*(\theta), y^*(\theta))$:

$$\frac{\partial L}{\partial x} = \frac{\partial f(x^*(\theta), y^*(\theta), \theta}{\partial x} + \lambda \frac{\partial g(x^*(\theta), y^*(\theta), \theta}{\partial x} = 0$$

$$\frac{\partial L}{\partial y} = \frac{\partial f(x^*(\theta), y^*(\theta), \theta}{\partial y} + \lambda \frac{\partial g(x^*(\theta), y^*(\theta), \theta}{\partial y} = 0$$

The value function is:

$$v(\theta) = f(x^*(\theta), y^*(\theta), \theta)$$

If the constraint is binding at θ, then:

$$h(\theta) = g(x^*(\theta), y^*(\theta), \theta) = 0$$

Where h is a composition function. Since $h(\theta') = g(x^*(\theta'), \theta') \geq 0$, $\forall \theta'$, θ is a local min of a function $h(\cdot)$ and so $h'(\theta) = 0$.

Differentiating $h(\theta)$ with respect to θ yields:

$$\frac{dh(\theta)}{d\theta} = \frac{\partial g(x^*(\theta), y^*(\theta), \theta)}{\partial x} \frac{\partial x^*(\theta)}{\partial \theta} + \frac{g(x^*(\theta), y^*(\theta), \theta}{\partial y} \frac{\partial y^*(\theta)}{\partial \theta} + \frac{g(x^*(\theta), y^*(\theta), \theta)}{\partial \theta} = 0$$

Differentiating $v(\theta)$ with respect to θ yields:

$$\begin{aligned}
\frac{dv(\theta)}{d\theta} &= \frac{\partial f(x^*(\theta), y^*(\theta), \theta)}{\partial x} \frac{\partial x^*(\theta)}{\partial \theta} + \frac{\partial f(x^*(\theta), y^*(\theta), \theta)}{\partial y} \frac{\partial y^*(\theta)}{\partial \theta} + \frac{\partial f(x^*(\theta), y^*(\theta), \theta)}{\partial \theta} \\
&= -\lambda \Big[\frac{\partial g(x^*(\theta), y^*(\theta), \theta)}{\partial x} + \frac{\partial g(x^*(\theta), y^*(\theta), \theta)}{\partial y} \frac{\partial y^*(\theta)}{\partial \theta} \Big] + \frac{\partial f(x^*(\theta), y^*(\theta), \theta)}{\partial \theta} \\
&= \lambda \frac{\partial g(x^*(\theta), y^*(\theta), \theta}{\partial \theta} + \frac{\partial f(x^*(\theta), y^*(\theta), \theta)}{\partial \theta} = \\
&= \frac{\partial L(x^*(\theta), y^*(\theta), \theta}{\partial \theta}
\end{aligned}$$

Notice, finally, that if the constraint is not binding, then $\lambda = 0$, and therefore we have simply:

$$\frac{dv(\theta)}{\theta} = \frac{f(x^*(\theta), \theta)}{\partial \theta}$$

□

Example 11.3.2. *Consider the following minimization problem:*

$$c(w) = \min_x w_1 x_1 + w_2 x_2$$

$$s.t.$$
$$x_1^{\frac{1}{3}} x_2^{\frac{2}{3}} \geq y$$

where $w \in \mathbb{R}^2_{++}$.

Set up the Lagrangian:

$$\mathcal{L}(x_1, x_2, \lambda) = w_1 x_1 + w_2 x_2 + \lambda(y - x_1^{\frac{1}{3}} x_2^{\frac{2}{3}})$$

Let's write down the FOCs:

- $\frac{\partial \mathcal{L}}{\partial x_1} = w_1 - \frac{1}{3} x_1^{-\frac{2}{3}} x_2^{\frac{1}{3}} = 0$
- $\frac{\partial \mathcal{L}}{\partial x_2} = w_2 - \frac{2}{3} x_1^{\frac{1}{3}} x_2^{-\frac{1}{3}} = 0$

Let's solve the FOCs:

$$\begin{aligned} \frac{w_1}{w_2} =& \frac{\frac{1}{3} x_1^{-\frac{2}{3}} x_2^{\frac{1}{3}}}{\frac{2}{3} x_1^{\frac{1}{3}} x_2^{-\frac{1}{3}}} \\ &= \frac{1}{2} \frac{x_1}{x_2} = \\ x_2 &= \frac{2 w_1 x_1}{w_2} \end{aligned}$$

Plugging in the constraint function:

$$x_1^{\frac{1}{3}} \Big(\frac{2 w_1 x_1}{w_2} \Big)^{\frac{2}{3}} = y$$

and solving for x_1^* *and* x_2^**, we have:*

$$x_1^* = y \Big(\frac{2 w_1}{w_2} \Big)^{-\frac{2}{3}}$$

$$x_2^* = y \Big(\frac{2 w_1}{w_2} \Big)^{\frac{1}{3}}$$

$$\lambda^* = 2^{-\frac{2}{3}} w_1^{\frac{1}{3}} w_2^{\frac{2}{3}}$$

Plug in the value function:

$$v(w) = w_1 y \Big(\frac{2 w_1}{w_2} \Big)^{-\frac{2}{3}} + w_2 y \Big(\frac{2 w_1}{w_2} \Big)^{\frac{1}{3}}$$

Take the partial derivatives with respect to w_1, w_2, and y. We have:

$$\frac{\partial c(w)}{\partial w_1} = y\Big(\frac{2w_1}{w_2}\Big)^{-\frac{2}{3}} = x_1^* = \frac{\partial \mathcal{L}}{\partial w_1}$$

$$\frac{\partial c(w)}{\partial w_2} = y\Big(\frac{2w_1}{w_2}\Big)^{\frac{1}{3}} = x_2^* = \frac{\partial \mathcal{L}}{\partial w_2}$$

$$\frac{\partial c(w)}{\partial y} = 2^{-\frac{2}{3}} w_1^{\frac{1}{3}} w_2^{\frac{2}{3}} = \lambda = \frac{\partial \mathcal{L}}{\partial y}$$

11.4 Concave Optimization

Theorem 11.4.1. *Let $E \subseteq \mathbb{R}^n$ be convex and $f : E \to \mathbb{R}$ be concave. Then:*

1. *any local maximum of f is a global max of f*
2. *the set $\arg\max\{f(x) : x \in E\}$ of maximizers of f on E is convex*

Proof. To prove 1), suppose x^* is a local max but not a global max. Then, there exists $\epsilon > 0$ s.t.:

$$f(y) \leq f(x^*), \forall y \in N_\epsilon(x^*)$$

But since x^* is not a global max, it exists a $z \in E$, such that $f(z) > f(x^*)$. Since E is convex, then we can write $\lambda x^* + (1-\lambda)z \in E$ for $\lambda in (0,1)$. If $\lambda \approx 1$. Then, we have:

$$f(\lambda x^* + (1-\lambda)y > \lambda f(x^*) + (1-\lambda)f(y) > f(x^*)$$

Which is a contradiction.

If f is concave, and x, y are both maximizers, then also their convex combination is a maximizer. Indeed:

$$f(\lambda x + (1-\lambda)x') \geq \lambda f(x) + (1-\lambda)f(y) = f(x) \;\; \forall \lambda \in (0,1)$$

□

Theorem 11.4.2. *Let $E \subseteq \mathbb{R}^n$ be convex and $f : E \to \mathbb{R}$ be strictly concave. Then $\arg\max\{f(x) : x \in E\}$ is either empty or a singleton.*

Proof. □

Theorem 11.4.3. *Let f and g_i, $i = 1, \ldots, n$ be concave. If x^* satisfies:*

1. $\nabla f(x^*) + \sum_{i=1}^n \lambda_i g_i(x^*) = 0$
2. $\lambda_i g_i(x^*) = 0$
3. $g_i(x^*) \geq 0 \quad \forall i = 1, \ldots, n$

For some multipliers $\lambda_1, \ldots, \lambda_n \geq 0$, then it is a global max.

Proof. Suppose x^*, λ satisfy KKT conditions (FOCs and CS). Let's define:

$$h(x) = f(x) + \sum_{i=1}^{n} \lambda_i g_i(x)$$

As a sum of concave functions, $h(x)$ is still concave. By FOCs, $\nabla h(x^*) = 0$, so h is maximized at x^*. Since $\lambda_i g_i(x^*) = 0$, then:

$$h(x^*) = f(x^*) \geq h(x) = f(x) + \sum_{i=1}^{n} \lambda_i g_i(x)$$

Thus, x^* is a global max. □

11.4.1 Quasi-Concave Programming

Theorem 11.4.4. *Suppose f is strictly quasi-concave and Γ is convex. Then:*

1. *Any local max of f is also a global max*
2. *The set $\arg\max\{f(x) : x \in E\}$ is either empty or a singleton*

Proof. Let's see 1). Suppose x^* is a local max. If x^* is not a global max, it exists a z such that $f(x^*) < f(z)$. Define $y = \lambda x^* + (1-\lambda)z$, since Γ is convex, then $y \in \Gamma$. Hence, we can write:

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

And then x^* is not a local max if $\lambda \approx 1$.

To see 2), suppose $z \neq x^* \in \arg\max f(x)$. Then, taking $\lambda \in (0,1)$ and $y = \lambda x^* + (1-\lambda)z$, we have:

$$f(y) > \min\{f(x^*), f(z)\}$$

Which is a contradiction. □

Theorem 11.4.5. *Suppose $g_i : E \to \mathbb{R}, \ \forall i = 1, \ldots, m$ is quasi-concave. Then:*

$$\Gamma = \Big\{x \in E : g_i(x) \geq 0, \forall i = 1, \ldots, m\Big\}$$

is convex.

Proof. Take $x, y \in \Gamma$. For each $\lambda \in [0,1]$, $\lambda x + (1-\lambda)y \in E$. Since g_i is quasi-concave, then:

$$g_i(\lambda x + (1-\lambda)y) \geq \min\{g_i(x), g_i(y)\} \geq 0$$

So, $\lambda x + (1-\lambda)y \in \Gamma$. □

Theorem 11.4.6. *(Arrow and Enthoven) Suppose f and g_i be C^1 quasi-concave functions mapping $E \subseteq \mathbb{R}^n \to \mathbb{R}$, where E is open and convex. Define:*

$$\Gamma = \Big\{ x \in E : g_i \geq 0, i = 1, \ldots, n \Big\}$$

Suppose there exists $x^* \in \Gamma$ and $\lambda \in \mathbb{R}^k$ such that the KKT conditions are met.

$$\mathbf{D}f(x^*) + \sum_{i=1}^{n} \lambda \mathbf{D}g_i(x^*) = 0$$

$$\lambda_i \geq 0, \ \ \forall i = 1, \ldots, n$$

$$\lambda_i g_i(x^*) \geq 0, \ \ \forall i = 1, \ldots, n$$

Then x^* maximizes f over Γ, provided at least one of the following conditions:

1. $\mathbf{D}f(x^*) \neq 0$
2. *f is concave*

Proof. □

Chapter 12

Convexity

In this section, we focus on Convex sets in Euclidean spaces.

Definition 12.0.1. A set X in R^n is convex if it contains a line segment connecting any two of its elements. I.e., $\forall x, y \in C$, the $(\lambda x + (1-\lambda)x') \in X$.

Definition 12.0.2. A set is called **cone** if, $\forall x \in C$, there is $\alpha x \in C$, $\forall \alpha \geq 0$.

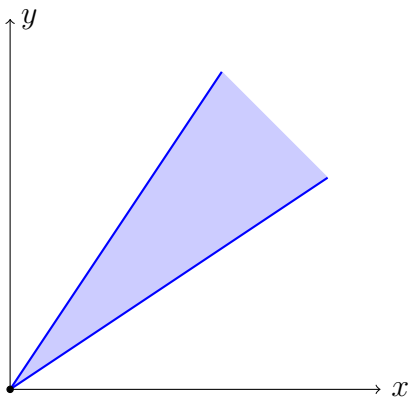


Figure 12.1: A convex cone

Definition 12.0.3. C is a **convex cone** if it is a cone, and it is convex. I.e.,$\forall x, y \in C$, $\alpha x + \beta y \in C$

Why do we need the definition of a convex cone? Because a cone may be not a convex set (think of just two lines connected at the origin)

Definition 12.0.4. A **conic combination** is a linear combination of x_i,

$$\sum_{i=1}^{n} \alpha_i x_i$$

where $\alpha_i \geq 0$.

Notice that we have a restriction on α, contrary to the linear combination (which is, therefore, a weaker condition)

Definition 12.0.5. An **hyperplane** is a set of vectors such that:

$$H = \Big\{ x \in \mathbb{R}^n ; \alpha^T x = b \Big\}$$

Where $\alpha \neq 0$ and $b \in \mathbb{R}$.

Definition 12.0.6. A **closed half-space** is:

$$H \Big\{ x \in \mathbb{R}^n ; \alpha^T x \leq b \Big\}$$

In $\mathbb{R}^2$, a hyperplane is a line. In constrained optimization, if there is more than one constraint, we are interested in their intersection.

Definition 12.0.7. A **polyhedron** is a set:

$$C = \Big\{ x \in \mathbb{R}^n : Ax \leq b \Big\}$$

A is a matrix, b is a vector.

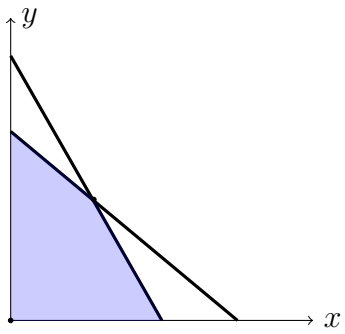


Figure 12.2: A polyhedron

Theorem 12.0.1 (Separating Hyperplane Theorem 1)**.** *Let C_1 and C_2 be two convex sets in $\mathbb{R}^m$, and $C_1 \cap C_2 = \emptyset$. Then, there exists $a \in \mathbb{R}^n_+$ and $b \in \mathbb{R}$, such that:*

$$ax \geq b \quad \forall x \in C_1$$

$$ax \leq b \quad \forall x \in C_2$$

Proof. □

Notice that this is a weaker version of the SHT. Indeed, take a point on the boundary of an open set. Recall that a point is a convex set. In this case, we cannot separate them.

Furthermore, being a closed convex set, it is not sufficient to guarantee separability.

Under what conditions can we strictly separate convex sets?

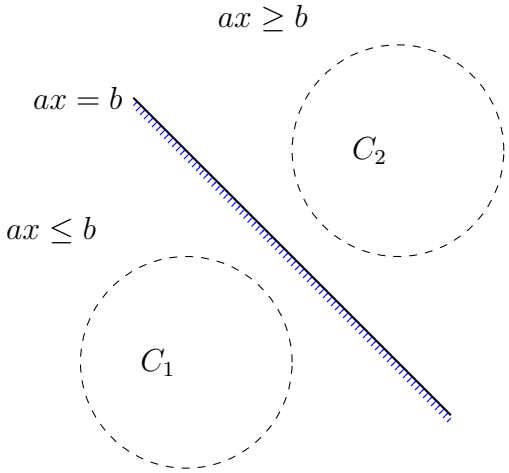


Figure 12.3: Separating Hyperplane Theorem 1

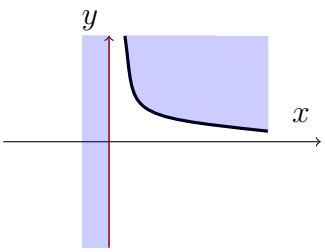


Figure 12.4: Two closed convex sets that are not separated

Theorem 12.0.2 (Separating Hyperplane Theorem 2)**.** *Let C be a closed convex set, and $x_0 \notin C$. Then there exists $a \in \mathbb{R} \setminus 0$ and $b \in \mathbb{R}$, such that:*

$$ax_0 > b > ax \quad \forall x \in C$$

How can we describe a convex set C?
A **primal description** is the list of any element in C.

$$C = \{x : x \in C\}$$

A **dual description** is what is not in C.

$$C = \{x : x \in C^C\}$$

Let's focus on the dual description. Suppose we have a Convex set C. We have a hyperplane, which separates a Convex set. Suppose we find **all** the hyperplanes that separate a Convex Set. Can we describe C? The answer is yes.

Theorem 12.0.3. *Any closed convex set in $\mathbb{R}^m$ is the intersection of the half-spaces containing it.*

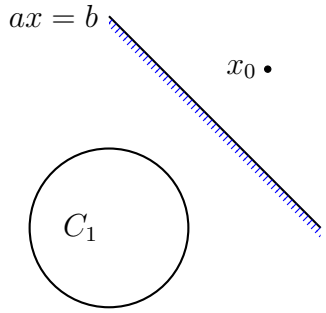


Figure 12.5: Separating Hyperplane Theorem 2

Proof. Let's define H as the intersection of all half-spaces containing C

$$H = \bigcap = \left\{ H' = C \subseteq H' \right\}$$

We want to show that $H = C$. That is: $H \subset C$ and $C \subset H$. By definition, we know that $C \subset H$ Now, suppose that there exists a $x \notin C$ but $x \in H$. Since $x \notin C$, by SHT2, we can find a plane that separates x from C. But $x \notin H'$. So, $x \notin H$. Therefore, we have reached a contradiction. □

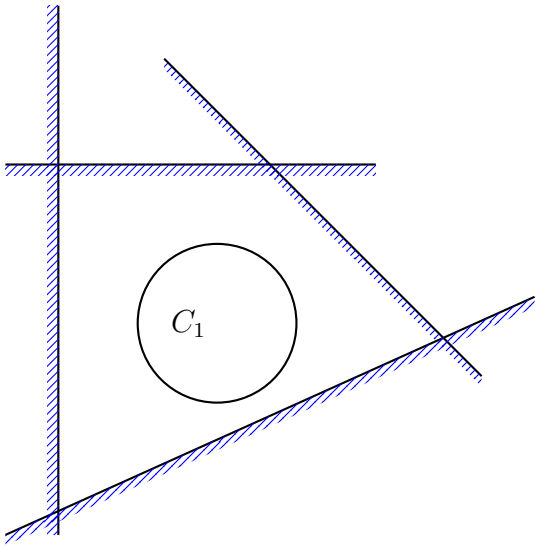


Figure 12.6: Any closed convex set as the intersection of all half-spaces containing it

12.1 The Farkas' Lemma

Using the notion of convex sets and separating hyperplanes, we can prove the Farkas' Lemma.

Theorem 12.1.1. *(Farkas' Lemma or Theorem of Alternatives) Let $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$. One and only one of the following is true:*

1. *$Ax = b$, $x \geq 0$ has a solution*
2. *$y^T A \geq 0$ and $y^T b < 0$ has a solution*

Before proving this lemma, let's discuss the geometrical intuition behind that. Recall that a cone can be written as:

$$C = \Big\{ y : \theta_1 a_1 + \theta_2 a_2, \theta_1, \theta_2 > 0 \Big\}$$

Define y as:

$$y = \Big\{ Ax = y : x \geq 0 \Big\}$$

And $A = (a_1, a_2)$ is a matrix containing two vectors a_1, a_2. Suppose there is b in the plane. We have two possibilities:

- $b \in C$
- $b \notin C$

Suppose $b \in C$. Then $Ax = b$, for all $x > 0$ (by above). Suppose $b \notin C$. C is closed and convex, so we can use the separating hyperplane theorem II. Then, there exists a hyperplane:

$$H = \Big\{ x : y^T x = \alpha \Big\}$$

such that:

$$y^T b < \alpha$$
$$y^T x \geq \alpha, \ \ \forall x \in C$$

More specifically, we can take a hyperplane passing through the origin so that:

$$y^T b < 0$$

$$y^T x \geq 0$$

If $x \in C$, we can express x as a conic combination of a_i. So, let's write:

$$y^T x = y^T (a_1 \theta_1) + y^T (a_2 \theta_2).$$

Choosing $(\theta_1, \theta_2) = (1, 0)$, we have:

$$y^T a_1 \geq 0$$

Choosing $(\theta_1, \theta_2) = (1, 0)$, we have $y^T a_2 \geq 0$. Therefore, we have:

$$y^T A \geq 0$$

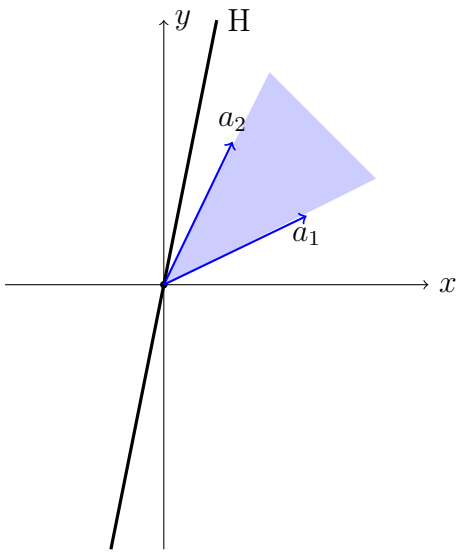


Figure 12.7: A geometrical representation of the Farkas' Lemma

Proof. (i) First, we show that (1) and (2) does not hold simultaneously. Suppose $y^T A \geq 0$ and $x \geq 0$. Then, we have $y^T Ax \geq 0$ (right multiplying by x). Since $Ax = b$, we can write $y^T b \geq 0$. But this contradicts (2).

(ii) We show that at least (1) or (2) must hold. If (1) fails, b is outside the cone. So we can use the separating hyperplane theorem. Then we have:

$$y^T b < \alpha \leq \underbrace{y^T a_i}_{\in C}, \quad \forall y \in C$$

Since, by definition of Cone, $0 \in C$, we must have:

$$y^T b < \alpha \leq 0$$

So $y^T b < 0$. This is (2).

We now must prove $y^T A \geq 0$. This can be written as:

$$y^T A \equiv y^T (a_1 a_2 \dots a_n) =$$

$$y^T a_1 + y^T a_2 + \dots + y^T a_n \geq 0$$

(recall that A is a $(m \times n)$ matrix, y is a $(n \times 1)$ vector, and y^T is a $(1 \times m)$ vector. Then $y^T A$ is a $(1 \times n)$ vector). By contradiction, suppose $y^T A \ngeq 0$. Hence, there must exist a $y^T a_k < 0$. Without loss of generality, we can assume $y^T a_1 < 0$, and we can write:

$$y^T A \underbrace{(x_1, 0, 0, \dots, 0)}_{\beta}$$

$y^T a_1 x_1 \leq 0$. Then:

$$y^T a_1 \beta < 0$$

and

$$y^T a_1 \beta < \alpha$$

This contradicts the SHT2. Hence:

$$y^T b < \alpha \leq y^T x, \forall x \in C \Rightarrow y^T A \geq 0.$$

□

To conclude, a recap of some concepts seen previously can be helpful.

Definition 12.1.1. The subgraph (or hypograph) of f is:

$$subf = \{(x, y) \in \mathbb{R}^n \times \mathbb{R} : f(x) \geq y\}$$

The supergraph (or epigraph) of f is:

$$supf = \{(x, y) \in \mathbb{R}^n \times \mathbb{R} : f(x) \leq y\}$$

Recall a function f is concave if and only if the subgraph is convex. Similarly, f is convex if and only if the supergraph is convex.

Finally, we see the concept of **hemi-continuity**.

Definition 12.1.2. We say that a function $f : \mathbb{R} \to \mathbb{R}$ is **upper-hemi continuous** at x_0 if $\lim_{x \to x_0} supf(x) \leq f(x_0)$

We say that a function $f : \mathbb{R} \to \mathbb{R}$ is **lower-hemi continuous** at x_0 if $\lim_{x \to x_0} supf(x) \geq f(x_0)$

Theorem 12.1.2. *A function f is upper hemi-continuous if and only if its subgraph is closed.*

A function f is lower hemi-continuous if and only if its supergraph is closed.

Proof. □

Chapter 13

Linear Programming

Let's consider the following problem:

$$V_p(b) = \max c^T x$$

$$\text{s.t.}$$

$$Ax \leq b, x \geq 0$$

where A is a $(m \times n)$ matrix, x is $(n \times 1)$ vector, and b is $(n \times 1)$.

This is called the **primal problem**. This form is called **canonical form**. The objective function is $c^T x$, The constraint set is:

$$C = \Big\{ \ x : Ax \geq b, a, x \geq 0 \Big\}$$

A solution is any $x \in \mathbb{R}^n$. A feasible solution is any $x \in C$. The optimal solution is any x^* that solves the problem.

Another form of writing this problem is:

$$V_p(b') = \max c' x$$

$$\text{s.t}$$

$$A'x = b, x' \geq 0$$

This is called the **standard form**.

We can express a standard form using a canonical form. Indeed:

$$A'x \geq b \quad \text{and} \quad A'x \leq b$$

Combining these two inequalities, we have:

$$Ax \geq b \ \text{ and } \ Ax' \leq b$$

We can write:

$$\begin{cases} -Ax \leq -b \\ Ax \leq b \end{cases}$$

In matrix form:

$$\begin{bmatrix} A \\ -A \end{bmatrix} x \leq \begin{bmatrix} b \\ -b \end{bmatrix}$$

We can also express a canonical form as a standard form, which means passing from inequality to equality. This can be done by using **slack variables**. Then:

$$Ax \leq b \Rightarrow b - Ax \geq 0$$

So, we can write $A(x+z) = b$, $x, z > 0$. Therefore:

$$\max c^T x$$

$$\text{s.t.}$$

$$(A|I) \cdot \begin{bmatrix} x_1 \\ \vdots \\ x_n \\ z_1 \\ \vdots \\ z_n \end{bmatrix} = b$$

Example 13.0.1.

$$\max x_1 + x_2$$

$$s.t$$

$$\begin{cases} x_1 + 2x_2 \leq 6 \\ x_1 - x_2 \leq 3 \\ x_1, x_2 \geq 0 \end{cases}$$

This problem can be written, in canonical form as:

$$\max \quad \begin{bmatrix} 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$s.t.$$

$$\begin{bmatrix} 1 & 2 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \leq \begin{bmatrix} 6 \\ 3 \end{bmatrix}$$

The problem in standard form is:

$$\max \begin{bmatrix} 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$s.t$$

$$\begin{bmatrix} 1 & 2 & 1 & 0 \\ 1 & -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ z_1 \\ z_2 \end{bmatrix} \leq \begin{bmatrix} 6 \\ 3 \end{bmatrix}$$

A graphical representation is given in Figure 8.

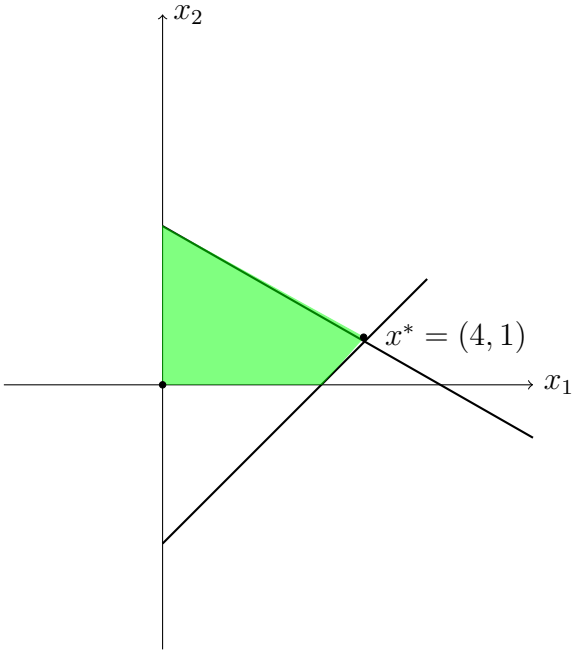


Figure 13.1: The geometrical representation

The problem of finding optimal points can be extremely simplified by looking only at the intersections between the linear constraints. These are the vertices of the polyhedron formed by the linear constraints.

Definition 13.0.1. x is a **vertex** of polyhedron C if there is no exists $y \neq 0$ such that $x + y$ and $x - y$ are both in C.

Theorem 13.0.1. *(Vertex Theorem) For any Linear Programming in* ***standard form*** *(only) with feasible solutions, we have:*

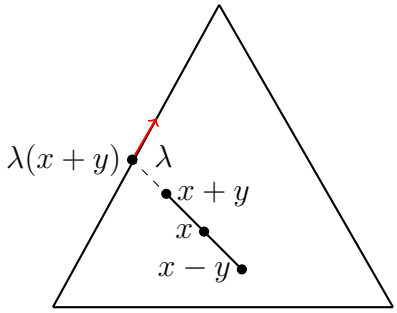
1. *A vertex exists*
2. *If the problem is bounded, $V_p(b) < \infty$, and $x \in C$, (the constrained set) then it exists a vertex x' such that $c \cdot x' \geq c \cdot x$*

Before seeing the proof, let's discuss the intuition.

Take $x \in C$. If it is a vertex, we are done. If it is not, then $x + y$ and $x - y$ belong to C. Each of these two points can be "stretched", i.e., scalar multiplied by a scalar λ up to reach the boundary. If it is a vertex, we are done. If it is on the boundary but not a vertex, we can do the same until we reach a vertex.

Let's see the proof now.

Proof. Let's first prove the existence. Choose $\mathbf{x} \in C$. If $\mathbf{x}$ is a vertex, we are done. If it is not a vertex, then, by definition, for some $\mathbf{y} \neq 0$, we have $\mathbf{x} \pm \mathbf{y} \in C$.



Therefore, $A\mathbf{y} = 0$, since:

$$A\mathbf{x} = \mathbf{b}$$

$$A(\mathbf{x} + \mathbf{y}) = 0$$

Further, if $x_j = 0$, then $y_j = 0$. Recall that $C = \{\mathbf{x} : A\mathbf{x} = \mathbf{b}, \mathbf{x} \geq 0\}$. Then, if $\mathbf{x} \pm \mathbf{y} \in C$, $\mathbf{x} \pm \mathbf{y} \geq 0$. If $x_j = 0$, but $y_j \neq 0$, then we have $x_j - y_j < 0$, and $\mathbf{x} - \mathbf{y} \ngeq 0$. This also implies that $\mathbf{x} \pm \lambda\mathbf{y}$ has fewer 0's than x.[1]

Let λ^* solve $\sup\{\lambda : \mathbf{x} \pm \lambda^*\mathbf{y} \in C\}$. Since $\mathbf{x}$ is not a vertex, $\lambda^* \neq 0$. And since C is closed, $\mathbf{x} \pm \lambda^*\mathbf{y} \in C$ (that is, the boundary is also in C). If $\mathbf{x} \pm \lambda^*\mathbf{y}$ is a vertex, we stop. Otherwise, we repeat the process. A vertex has more zeros than a point in the boundary. This assures that, at a certain point, the process will stop.

Let's see optimality now. We want to prove that $c \cdot \mathbf{x} < c \cdot \mathbf{x}'$, where $\mathbf{x}'$ is a vertex. If $\mathbf{x}$ is not a vertex, then $\mathbf{x} \pm \mathbf{y} \in C$. We have, then $c\mathbf{x} + c\mathbf{y}$. As before, we scalar multiply by $\lambda > 0$ until we reach the boundary. Then, $c(\mathbf{x} + \lambda\mathbf{y}) = c\mathbf{x} + c\lambda\mathbf{y} \geq c\mathbf{x}$.

Notice that it cannot be the case that $y_i \geq 0$ for all j. Indeed, if so, by construction, we have $c \cdot \mathbf{y} > 0$ and $\mathbf{x} + \lambda\mathbf{y} \in C, \forall\lambda \geq 0$. But then $c(\mathbf{x} + \lambda\mathbf{y}) \to \infty$ as $\lambda \to \infty$, thus violating the assumption that of boundedness of $V_p(b)$. Therefore $y_j < 0$ for some j. For large λ, $\mathbf{x} + \lambda\mathbf{y} \ngeq 0$. Let λ^* denote the $\max\lambda$. Then we know $\mathbf{x} + \lambda^*\mathbf{y} \in C$ and has at least one more zero. Then:

$$c(\mathbf{x} + \lambda^*\mathbf{y}) \geq c\mathbf{x}$$

Repeat the process until we reach a vertex. In each process, the value function strictly increases, so at a certain point, we reach a vertex. □

The first statement provides the existence of a vertex. The second statement establishes the optimality of a vertex. Notice that this implies that the number of vertex is finite because we are in $\mathbb{R}^m$. Notice that a vertex may be not the only optimal solution. The theorem states that a vertex is part of the set of solutions.

A vertex may not exist in canonical form.

[1] Another way of seeing it is that a point in the constraint set has fewer zeros than a point on the boundary. Think of the unit simplex in $\mathbb{R}^3$. Each point on the boundary can be written as $(x_1, x_2, 0), (x_1, 0, x_3), (0, x_2, x_3))$.

13.1 Duality

We introduce the notion of **basic solution**. Notice that this holds only for LP in standard form.

Definition 13.1.1. A feasible solution to LP in standard form is **basic** if and only if the columns a_j of A such that $x_j > 0$ are linearly independent.

This means that if we take a submatrix of A, say A_X, consisting of columns a_j for each $x_j > 0$ has full column rank.

Example 13.1.1. *Let's see the following example:*

$$x_1 + x_2 + x_3 = 1$$

$$2x_1 + 3x_2 = 1$$

with $x \geq 0$. Let's write the matrix of the coefficient constraints.

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 3 & 0 \end{bmatrix}$$

This matrix has three bases.

- $B_1 = \begin{bmatrix} 1 & 1 \\ 2 & 0 \end{bmatrix}$ *(notice that this matrix is made up by columns a_1 and a_3). Then, the solution is given by the system:*

$$\begin{cases} x_1 + x_3 = 1 \\ 2x_1 = 1 \end{cases} \Rightarrow \begin{cases} x_1 = 1 - x_3 \\ 2(1 - x_3) = 1 \end{cases} \Rightarrow \begin{cases} x_1 = 1 - x_3 \\ 2 - 2x_3 = 1 \end{cases} \Rightarrow \begin{cases} x_1 = 1 - x_3 \\ x_3 = \frac{1}{2} \end{cases}$$

  Then $x^1 = (\frac{1}{2}, 0, \frac{1}{2})$. This is a basic solution.

- $B_2 = \begin{bmatrix} 1 & 1 \\ 2 & 3 \end{bmatrix}$ *Since $x_2 = -1$, there is no basic solution for this submatrix.*

- $B_3 = \begin{bmatrix} 1 & 1 \\ 3 & 0 \end{bmatrix}$ *$x^3 = (0, \frac{1}{3}, \frac{2}{3})$ is a basic solution.*

Once one finds the basic solutions, she can ask how these are connected with vertices. A theorem which connects basic solutions and vertices is the following.

Theorem 13.1.1. *A solution is basic if and only if it is a vertex*

Proof. Let's see ($\Rightarrow$). Suppose $\hat{x}$ is not a vertex. Then we want to show it is not a basic solution. Since $\hat{x}$ is not a vertex, it exists a $y \neq 0$ such that $x \pm y \in C$ (the constraint set), and if $x_j \neq 0$, then $y_i \neq 0$ (see above). This implies Ay is a linear combination of columns of the submatrix A_X because y only takes positive values in A_X. By definition,

$Ay = 0$, so the columns of A_X must be linearly dependent. So $\hat{x}$ is a basic solution. We have reached a contradiction.

($\Leftarrow$) Suppose $\hat{x}$ is not basic. We want to show it is not a vertex. If $\hat{x}$ is not a basic solution, then the submatrix A_X must be linearly dependent, and so, there is a $\hat{y}$ such that $A_X\hat{y} = 0$. Let y that takes value 0 outside the columns in A_X be equal to $\hat{y}$ for the columns of A_X. Then $Ay = 0$, $x_j = 0$ and $y_j = 0$. This implies $\hat{x} \pm y \in C$. therefore, $\hat{x}$ is not a vertex. □

Theorem 13.1.2. *(The Fundamental Theorem of Linear Programming) If an LP in a standard form has a feasible solution, then it has a basic feasible solution.*

If it has an optimal solution, then it has a basic optimal solution.

Proof. If x is a feasible solution, then, by the Vertex theorem, a vertex exists, and it is a solution. From the theorem above, a vertex is a basic solution. □

This means that, graphically, if the LP problem has a solution, we only need to check for basic solutions, i.e., vertices. However, two important caveats must be recalled:

1. The problem must be bounded
2. We are in standard form

However, the following proposition (without proof) links feasible solutions in standard forms and feasible solutions in canonical form.

Proposition 15. *The feasible solution x to a canonical problem is a vertex if and only if there exists some slack variable z such that (x, z) is a vertex for the corresponding standard form.*

Before going on, let's see visually what it means to pass from standard form to canonical form.

A constraint set of a problem in standard form is given by the following figure.

This can be written as:

$$C_3 = \Big\{(x, y, z) : x + y + z = 1, x, y, z \geq 0\Big\} \subset \mathbb{R}^3_+$$

However, in canonical form, we have one variable less (no slack variables). So, $C \subset \mathbb{R}^2_+$

Now, we can see more in detail the dual problem. Recall that the primal of an LP problem in the canonical form is:

$$V_p(b) = \max c^T x$$

$$\text{s.t.}$$

$$Ax \leq b, x \geq 0$$

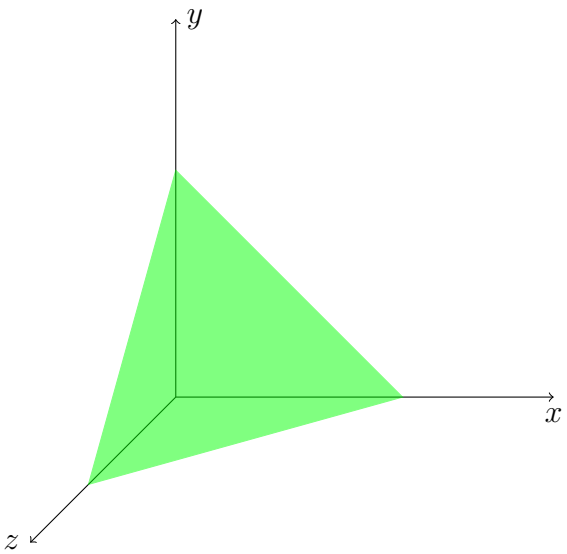


Figure 13.2: A C of LP in standard form

The **dual** is:

$$V_d(c) = \min y^T b$$

$$\text{s.t.}$$

$$A^T y \geq c, y \geq 0$$

Example 13.1.2. *Let's see the following example:*

$$V_p(b) = \max x_1 + x_2$$

$$s.t$$

$$\begin{cases} x_1 + 2x_2 \leq 6 \\ x_1 - x_2 \leq 3 \\ x_1, x_2 \geq 0 \end{cases}$$

In the matrix from this becomes:

$$\max \quad \begin{bmatrix} 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$s.t.$$

$$\begin{bmatrix} 1 & 2 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \leq \begin{bmatrix} 6 \\ 3 \end{bmatrix}$$

From the above, we know that the optimal solution is $(4, 1)$.

Let's write down the dual now:

$$V_d(c) = \min 6y_1 + 3y_2$$

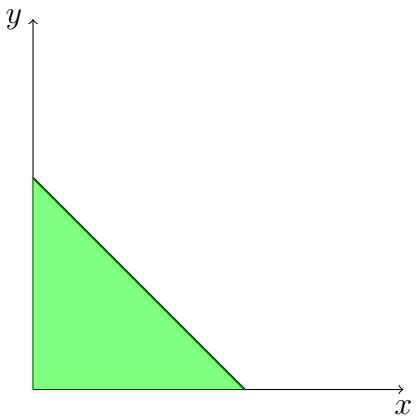


Figure 13.3: A C of LP in canonical form

$$s.t.$$

$$\begin{cases} y_1 + y_2 \geq 1 \\ 2y_1 - y_2 \geq 3 \\ y_1, y_2 \geq 0 \end{cases}$$

In matrix form, this is equal to:

$$\min \quad \begin{bmatrix} 6 & 3 \end{bmatrix} \cdot \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

$$s.t.$$

$$\begin{bmatrix} 1 & 1 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \geq \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Solving this system, $y^* = (\frac{2}{3}, \frac{1}{3})$.

Plugging x^* *in* $V_p(d)$ *we obtain:*

$$V_p(b) = 4 + 1 = 5$$

Plugging y^* *in* $V_d(c)$ *we obtain:*

$$V_d(c) = \frac{2}{3}6 + 3\frac{1}{3} = 4 + 1 = 5$$

Then:

$$V_p(b) = V_d(c)$$

What if the primal problem is written in standard form? Recall that we can pass from a standard form into a canonical form rewriting the constraints $Ax = b$ as:

$$Ax \leq b$$

$$-Ax \leq -b$$

Therefore the problem in canonical form is:

$$V_p(b) = \max c^T x$$

$$\text{s.t.}$$

$$\begin{bmatrix} A \\ -A \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \leq \begin{bmatrix} b \\ -b \end{bmatrix}$$

The dual problem is then:

$$V_d(c) = \min y^T \begin{bmatrix} b \\ -b \end{bmatrix}$$

$$\text{subject to}$$

$$y^T \begin{bmatrix} A & -A \end{bmatrix} \geq c^T$$

In the dual problem in standard form, the non-negativity constraint of y can be waived.

What if the primal has both inequalities and equalities? Assume the following primal problem:

$$V_p(b, b') = \max c^T x$$

$$\text{subject to}$$

$$Ax \leq b$$

$$Ax' = b'$$

$$x \geq 0$$

We can pass from canonical to standard forms. So we can write the constraints of the problem above as:

$$\begin{bmatrix} A \\ A' \\ -A' \end{bmatrix} x \leq \begin{bmatrix} b \\ b' \\ -b' \end{bmatrix}$$

Then, the dual is:

$$\min \begin{bmatrix} y^T & z_1^T & z_2^T \end{bmatrix} \begin{bmatrix} b \\ b' \\ -b' \end{bmatrix}$$

$$\text{subject to}$$

$$\begin{bmatrix} y^T & z_1^T & z_2^T & \end{bmatrix} \begin{bmatrix} A \\ A' \\ -A' \end{bmatrix} \geq c^T$$

$$y^T, z_i^T \geq 0$$

Taking $z = z_1 - z_2$, we can write:

$$y^T b + z^T b$$

$$\text{subject to}$$

$$y^T A + z^T A \geq c^T$$

However, it is not always the case that the values of the primal and the dual are the same. See the following example.

$$\max x_1 + x_2$$

$$\text{subject to}$$

$$-2x_1 - x_2 \leq 1$$

$$-x_1 - 2x_2 \leq 1$$

The dual is:

$$\min y_1 + y_2$$

$$\text{subject to}$$

$$-2y_1 - y_2 \geq 1$$

$$-y_1 - 2y_2 \geq 1$$

$$y_i \geq 0$$

In this case, the dual problem is infeasible since there is no $y \in C$. In some cases, it can be that both are infeasible.

Example 13.1.3. *See the following example:*

$$V_p(b) = \max x_1 + x_2$$

$$\textit{subject to}$$

$$\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \leq \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

This gives the following unbounded constrained set:

Is it possible that both the problems are unbounded? Recall that unbounded means that $V_p(b), V_d(c)$ may be equal to $\pm\infty$. If $x^* \in C$ is feasible, then $V_p(b) = \infty$ is not possible. The only possibility for both being unbounded is if $V_d = -\infty$, $V_p = \infty$. But these cases are not possible.

An important result is contained in the following theorem.

Theorem 13.1.3 (Weak Duality)**.** *$V_p(b) \leq V_d(c)$ if both are feasible.*

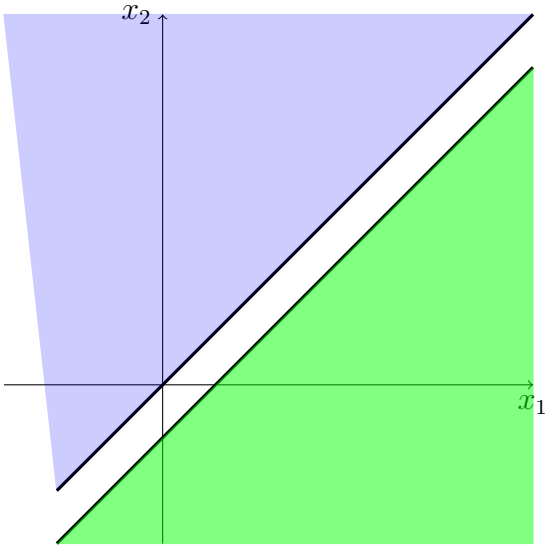


Figure 13.4: A problem with no solution

Proof. Let x, y be feasible solutions of the primal and the dual problem (namely $x \in C_P$, $y \in C_D$). Since $y \in C_D$, then $y^T A \geq c^T$. Similarly, since $x \in C_P$, then $x \geq 0$. Then we have $(y^T A - c^T)x \geq 0$ and $y^T A \geq c^T x$.

Following the same logic as above, since $x \in C_P$, then $Ax \leq b$ and, since $y \in C_D$, then $y \geq 0$. So $y^T Ax \leq y^T b$. Combining these two results, we have:

$$c^T x \leq y^T Ax \leq y^T b \quad \forall x, y \in C$$

And then $V_p \leq V_d$. □

According to this result, it is not possible for both V_p and V_d to be unbounded. So we can have only three possibilities:

1. V_p is unbounded and V_d is unfeasible
2. V_d is unbounded and V_p is unfeasible
3. Both are feasible and bounded

Then, we can write down the following result.

Theorem 13.1.4 (Strong Duality)**.** *For primal and dual problems, exactly one of the following three alternatives must hold:*

1. *Both are feasible, and both have optimal solutions, with $V_p = V_d$*
2. *One is unbounded, the other is unfeasible*
3. *Both are unfeasible*

Proof. We want to show that $V_p = V_d$. Let's express the primal problem using the standard form. Then we can write:

$$V_p(b) = \max c \cdot x$$

$$\text{subject to}$$

$$Ax = b$$

$$x \geq 0$$

Consider the following system of inequalities:

$$\begin{bmatrix} A \\ C \end{bmatrix} x = \begin{bmatrix} b \\ V_p + \epsilon \end{bmatrix} \tag{13.1}$$

With $x \geq 0$ (notice that this is another way of saying that exists an x such that the primal problem has a solution). Suppose V_p has a solution. If $\epsilon = 0$, this constraint must have a feasible solution ($Ax = b$ is the constrained set of the standard form, if x is feasible, a solution exists). If $\epsilon > 0$, you cannot find an optimal solution that achieves the max value. If 4) has not a solution, x is outside the cone spanned by $\begin{bmatrix} A \\ c \end{bmatrix}$. By Farkas' Lemma, if $\epsilon > 0$, the following must have a solution:

$$\begin{bmatrix} y^T & \alpha \end{bmatrix} \cdot \begin{bmatrix} A \\ c \end{bmatrix} \geq 0 \tag{13.2}$$

$$\begin{bmatrix} y^T & \alpha \end{bmatrix} \cdot \begin{bmatrix} b \\ V_p + \epsilon \end{bmatrix} < 0 \tag{13.3}$$

If $\alpha = 0$, if y is feasible to 5) and 6), and $\epsilon > 0$, it must be feasible also when $\epsilon = 0$. But this is not possible because, by Farkas' Lemma, if 4) has a solution, then 5) and 6) don't have.

If $\alpha > 0$, if y satisfies 4) and 5), when $\epsilon > 0$, it must satisfy them also when $\epsilon = 0$. And this is not possible.

Then $\alpha < 0$ is the only possibility. Without loss of generality, we can put $\alpha = -1$. Then, from 4), we have $y^T A \geq c^T$, $y \in C_d$. From 5) we have $y^T b < V_p + \epsilon$. In other words, we find a feasible solution for the dual problem, such that $y?y^T b < V_p + \epsilon$ and $V_d < V_p + \epsilon$. From the weak duality theorem, V_d is bounded from below by V_p. So, $V_p \leq V_d < V_p + \epsilon$. Since ϵ can be made arbitrarily small, we have $V_p = V_d$. This completes the proof. □

Notice, however, that the duality theorem does not provide a solution to a linear optimization problem but only establishes a way to find it. Then, we need the following result.

Theorem 13.1.5 (Complementary Slackness)**.** *Suppose x^* and y^* are feasible solutions for primal and dual problems, respectively. Then, they are optimal solutions if and only if the following holds:*

1. *For each constraint i in the primal problem*

$$y_i^*(b_i - A_i x^*) = 0$$

(i^{th} coordinate unknown times i^{th} constraint.

2. *For each y_j in the dual problem*

$$(y^{*^T} A_j - c_j)x_j^* = 0$$

Proof. To see 1), suppose x^* and y^* are feasible and satisfy the complementary slackness condition. Then:

$$\underbrace{y^{*^T} b = y^{*^T} A x^*}_{\text{from 1)}} = c^T x^*$$

So y^* and x^* are optimal solutions if $V_p = V_d$.

To see 2), if x^* and y^* are optimal, they must be feasible. $Ax^* \leq b$, $y^* \geq 0$ imply $y^{*^T} Ax^* \leq y^{*^T} b$. Similarly, $c^T x^+ \leq y^{*T} Ax^*$ imply:

$$y^{*^T} b = y^{*^T} Ax^* = c^T x^*$$

Where $y^{*T} b = y^{*T} Ax^*$ derives from 1), and $y^{*T} Ax^* = c^T x^*$ from 2). Then, complementary slackness holds. □

In a nutshell, what the 2 conditions of complementary slackness say, is that:

- If $A_i x^* = 0$, then $b_i = 0$ and $y_i^* \geq 0$
- If $A_j y^* = 0$, then $c_j = 0$ and $x_j^* \geq 0$

Notice that complementary slackness is a necessary and sufficient condition for optimal solutions.

Example 13.1.4. *Let's see a numerical example:*

$$V_p(b) = \max x_1 - x_2$$

$$s.t.$$

$$-2x_1 + x_2 \leq 2$$

$$x_1 - 2x_2 \leq 2$$

$$x_1 + x_2 \leq 5$$

$$x_1, x_2 \geq 0$$

The dual problem is:

$$V_d(c) = \min 2y_1 + 2y_2 + 5y_3$$

$$s.t$$

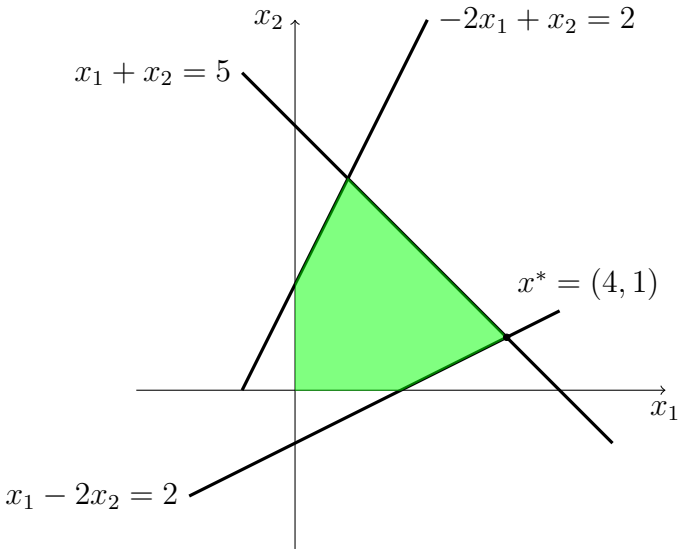


Figure 13.5: Geometrical solution

$$-2y_1 + y_2 + y_3 \geq 1$$
$$y_1 - 2y_2 + y_3 \geq -1$$
$$y_1, y_2, y_3 \geq 0$$

See the figure for a graphical representation of this problem. The optimal solution is $x^* = (4, 1)$. *Then* $V_p(b) = 3$. *Plugging into the constraints, we have:*

$$-2(4) + 1 < 2$$
$$(4) - 2(1) = 2$$
$$(4) + (1) = 5$$

Then, we have one ***slack constraint*** *and two binding constraints. Since the first constraint is slack, this means that* y_1 *is equal to* 0 *by Complementary Slackness. So, we need only to find* y_2 *and* y_3. *Applying Complementary Slackness, we have:*

$$x_1^*(-2y_1^* + y_2^* + y_3^* - 1) = 0$$
$$x_2^*(y_1^* - 2y_2^* + y_3^* + 1) = 0$$

Since $x_1^*, x_2^* > 0$, *this means that:*

$$-2y_1^* + y_2^* + y_3^* - 1 = 0$$
$$y_1^* - 2y_2^* + y_3^* + 1 = 0$$

But the first constraint is slack, so $y_1^* = 0$. *Then:*

$$y_2^* + y_3^* = 1$$

$$-2y_2^* + y_3 = -1$$

Solving the system, we have:

$$\begin{cases} y_2^* + y_3^* = 1 \\ -2y_2^* + y_3 = -1 \end{cases} \Rightarrow y_2^* = \frac{2}{3}, y_3^* = \frac{1}{3}$$

Then, the optimal solution of the dual is $(y_1^*, y_2^*, y_3^*) = (0, \frac{2}{3}, \frac{1}{3})$ *and* $V_d(c) = 3 = V_p(b)$

Chapter 14

Non-Linear Programming

In this section, we explore the topic of Non-linear programming. Most results have already been established in the previous section about optimization. However, now they will be treated in connection with some results from Linear programming.

The main difference with the results in the previous chapters is that it is not assumed linearity anymore. The objective function is non-linear, and the constraints are too (but these can also be linear as well).

A non-linear optimization problem is:

$$\max f(\mathbf{x})$$

$$\text{s.t.}$$

$$g_i(\mathbf{x}) \geq 0 \quad i \in M = 1, \ldots, m$$

Where $f, g : \mathbb{R}^n \to \mathbb{R}$, continuous and differentiable. Recall that a $\epsilon-$**neighborhood** is the set of point enough close to x, that is:

$$N_\epsilon(x) = \Big\{ y : |x - y| < \epsilon \Big\}$$

Further, we can define the **feasible set** as the set of values for x which satisfies the constraints:

$$F = \Big\{ x \in \mathbb{R}^n : g_i(x) \geq 0, ; \quad \forall i \in M \Big\}$$

Recall that x is a **local max** for a linear programming problem if it exists $\epsilon > 0$ such that $f(x) > g(y)$ for all $y \in N_\epsilon(x)$. x is a **global max** if it is an optimal solution for the problem.

The simplest case is that of unconstrained optimization, namely when $M = \emptyset$. Therefore, we have the following result.

Theorem 14.0.1 (First Order Conditions)**.** *If x^* is a local max, then:*

$$\nabla f(x^*) = 0$$

where $\nabla f = \begin{bmatrix} f_1' & f_2' & \ldots & f_n' \end{bmatrix}$ is the gradient of $f(\cdot)$.

Proof. Suppose $h = \nabla f(x^*) \neq 0$. Then:

$$h^T h = h^T \nabla f(x^*) > 0$$

Therefore, we can write $f(x^* + \varepsilon h) \approx f(x^*) + \varepsilon h^T \nabla f(x^*)$, where:

$$\varepsilon h^T \nabla f(x^*) > 0$$

Where ε is small, then:

$$f(x^* + \varepsilon h) > f(x^*)$$

thus contradicting x^* being an optimal point. □

The intuition is that the derivative is the slope of the function so that it reaches its maximum or minimum when the slope is equal to 0. Since this is a necessary condition for local max, it is a necessary condition also for global max. However, it is not a sufficient condition. That is, the derivative can be equal to 0, still the point not be a max or min. The standard example is the function $y = x^3$.

Further, this result does not say anything about the optimal point being a max or min. We need to impose further restrictions on the sign of the second derivatives. These are the Second Order Conditions:

- $f''(x^*) < 0$ is a max.
- $f''(x^*) > 0$ is a min.

The intuition is that in the $\epsilon-$neighborhood of x^*, if x^* is a max, then $f'(x^*) < 0$, and therefore $f''(x^*)$ (for x^* being a min, the sign is reversed). Let's see the formal statement.

Theorem 14.0.2 (Second Order Conditions)**.** *Suppose f is C^2 on $\mathbb{R}$. Then:*

1. *if f has a local max in x^*, $\nabla^2 f(x^*)$ is **negative semidefinite**[1]*
2. *if f has a local min in x^*, $\nabla^2 f(x^*)$ is **positive semidefinite***

These are the necessary conditions. The sufficient conditions are:

1. *If $\nabla f(x^*) = 0$ and $\nabla^2 f(x^*)$ is **negative definite**, then x^* is a strictly local max*
2. *If $\nabla f(x^*) = 0$ and $\nabla^2 f(x^*)$ is **positive definite**, then x^* is a strictly local min*

[1]Recall that there are many definitions for negative/positive definitiness/semi-definiteness. One is

Definition 14.0.1. Let A be a $n \times n$ symmetric matrix, and A_k denote the $k \times k$ submatrix. Then:

- A is **positive definite** if and only if $(-1)^k |A_k > 0|$, for all k
- A is **negative definite** if and only if $(-1)^k |A_k| < 0$, for all k.

The definition for semidefiniteness is more involute. See p.62 of these notes)

Proof. Let's see only the necessary and sufficient conditions for local max (the argument for local min is specular).

Suppose x^* is a local max, we want to show that $\nabla^2 f(x^*)$ is negative semi-definite.

We can write (using Taylor expansion):

$$f(x^* + \varepsilon h) \approx f(x^*) + \varepsilon h^T \nabla f(x^*) + \frac{1}{2}\varepsilon^2 h^T \nabla^2 f(x^*) h + \mathcal{O}(\varepsilon^T)$$

Since $\varepsilon h^T \nabla f(x^*) = 0$, and rearranging, we have:

$$f(x^* + \varepsilon h) - f(x^*) \approx \frac{1}{2}\varepsilon^2 h^T \nabla^2 f(x^*) h$$

Since $\frac{\varepsilon^2}{2} > 0$, then:

$$h^T \nabla^2 f(x^*) h \leq 0$$

$\nabla^2 f(x^*)$ is a Negative Semidefinite Matrix.

Let's see the case for strict local max. Suppose $\nabla f(x^*) = 0$, and $\nabla^2 f(x^*)$ is negative definitive. We want to show that x^* is a local max. We can write;

$$f(x) - f(x^*) = \nabla f(x^*)(x - x^*) + \frac{1}{2}(x - x^*)^T \nabla^2 f(x^* + \mathcal{O}(x - x^*))(x - x^*)$$

where x is enough close to x^*. Since $\nabla f(x^*) = 0$, we can look at the second term. Because of continuity, we can find x sufficiently close to x^* such that this term is less than 0. This completes the proof. □

The FOCs are necessary conditions for local max but not sufficient. The Second Order Conditions are sufficient, but only for local max. For global max, we need a further result.

Theorem 14.0.3. *(Sufficient Conditions for Global Max) Let f be a **concave, continuous** and **differentiable** function of an open convex set C. Then f has a max at x^* if and only if we have the FOCs, i.e., $\nabla f(x^*) = 0$.*

Proof. Notice that this theorem states an if and only if condition. Then, ($\Rightarrow$) is trivial since a Global Max is also a Local Max.

Let's see ($\Leftarrow$). By contrapositive, we want to show that if y is not a global max, then $\nabla f(y) \neq 0$. y is not a global max, thus we can find $x \neq y$ such that $f(x) > f(y)$. Since $f(\cdot)$ is concave:

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

Rearranging, we have:

$$f(y + \lambda(x - y)) - f(y) \geq \lambda(f(x) - f(y))$$

Let $h = x + y$, and $\theta = f(x) - f(y) > 0$ then:

$$f(y + h) - f(y) \geq \lambda\theta$$

For sufficiently small λ, we have $h^T \nabla f(y) \geq \theta > 0$. So $h^T \nabla f(y)$ is strictly positive and $\nabla f(y) \neq 0$. □

So far, we have recapped the main results from the simplest case of Non-linear optimization, the unconstrained case. Let's now see the more general case, where there are one or more constraints.

14.1 Constrained Optimization

A general version of this problem is the following:

$$\max_{x \in \mathbb{R}^n} f(\mathbf{x}) \tag{14.1}$$

s.t.

$$g_i(\mathbf{x}) \geq 0 \quad i \in M = 1, \ldots, m$$

We want to show that if x^* is a local max, then it satisfies the FOCs. Therefore, in this section, I will provide several results and variants of the original, powerful, **Karush-Kuhn-Tucker Theorem**.

To begin with, let's state that if x^* solves the problem (7), then, we can express the constraint $g(x^*)$ as a linear approximation:

$$g(x^* + \varepsilon g) \approx g^(x^*) + \varepsilon h^T \nabla g(x^*)$$

Therefore, we can rewrite the problem as follows.

$$\max_h f(x^*) + \varepsilon h^T \nabla f(x^*)$$

s.t.

$$g_i(x^*) + \varepsilon h^T \nabla g_i(x^*) \geq 0 \quad ß \in M$$

If x^* is a local max, then h^* must be zero. Since we have linearized both the value function and the constraints, we can rewrite this problem as a Linear Programming problem.

$$V_p(b) = \max_h \begin{bmatrix} \varepsilon \frac{\partial f}{\partial x_1} & \ldots & \varepsilon \frac{\partial f}{\partial x_n} \end{bmatrix} \begin{bmatrix} h_1 \\ \vdots \\ h_n \end{bmatrix} \tag{14.2}$$

s.t.

$$\begin{bmatrix} -\varepsilon \frac{\partial g_1}{\partial x_1} & \cdots & -\varepsilon \frac{\partial g_1}{\partial x_n} \\ \vdots & \cdots & \vdots \\ -\varepsilon \frac{\partial g_n}{\partial x_1} & \cdots & -\varepsilon \frac{\partial g_n}{\partial x_n} \end{bmatrix} \begin{bmatrix} h_1 \\ \vdots \\ h_n \end{bmatrix} \leq \begin{bmatrix} g_1 \\ \vdots \\ g_n \end{bmatrix}$$

Then, the dual is:

$$V_d(c) = \min_\mu \begin{bmatrix} \mu_1 & \ldots & \mu_n \end{bmatrix} \begin{bmatrix} g_1 \\ \vdots \\ g_n \end{bmatrix} \tag{14.3}$$

$$\text{s.t.}$$
$$\begin{bmatrix} -\varepsilon\frac{\partial g_1}{\partial x_1} & \cdots & -\varepsilon\frac{\partial g_n}{\partial x_1} \\ \vdots & \cdots & \vdots \\ -\varepsilon\frac{\partial g_1}{\partial x_n} & \cdots & -\varepsilon\frac{\partial g_n}{\partial x_n} \end{bmatrix} \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_n \end{bmatrix} = {}^{2} \begin{bmatrix} \varepsilon\frac{\partial f}{\partial x_1} \\ \vdots \\ \varepsilon\frac{\partial f}{\partial x_n} \end{bmatrix}$$

The value of the primal problem is 0, since the value of the dual problem is 0 by complementary slackness. Then:

$$\sum_{i=1}^{n} h_i \varepsilon \frac{\partial f}{\partial x_i} = 0$$

Because:

$$\mu_i g_i(x^*) = 0$$

Since $h_i \geq 0$ and $\frac{\partial f}{\partial x_i} \geq 0$, then the sum of any non-negative terms is equal to 0 if and only if each term is equal to 0.

The following result formally establishes what we have stated.

Lemma 14.1.1 (Fritz John Conditions)**.** *Let's have a constrained optimization problem where x^* is a local max. Then, there exist non-negative multipliers $\{\mu_0, \mu_1, \ldots, \mu_n\}$ not all zeros, such that:*

$$\mu_0 \nabla f(x^*) + \sum_{i=1}^{n} \mu_i \nabla g_i(x^*) = 0$$

and

$$\mu_i f(x^*) = 0 \quad \forall i \in M$$

where μ_0 can be zero or not.

Notice that we can normalize the multipliers as follows:

$$\frac{\mu_0}{\sum_{i=1}^{n} \mu_i} \nabla f(x^*) + \ldots \frac{\mu_n}{\sum_{i=1}^{n} \mu_i} \nabla g_m(x^*) = 0$$

In other words, we can say that 0 belongs to the **Convex Hull**[3] spanned by the conditions of the Lemma above.

Proof. Suppose the first $r > 0$ constraints bind. Then, for $i > r$ and $\mu_i = 0$, Complementary Slackness holds. We want to show that:

$$\frac{\mu_0}{\sum_{i=1}^{n} \mu_i} \nabla f(x^*) + \frac{\mu_r}{\sum_{i=1}^{n} \mu_i} \nabla g_m(x^*) = 0$$

and:

$$0 \in Conv\Big\{\nabla f(x^*), \ldots, \nabla g_m(x^*))\Big\}$$

[2]Since the primal does not have non-negativity constraints, the dual has the equal sign

[3]Given a set $X = (a, b)$, the Convex Hull is the smallest convex set containing (a, b)

Suppose not. Then, by the Separating Hyperplane Theorem, we can find a vector $h \in \mathbb{R}^n$ such that $h \cdot y > 0, \forall y \in Conv\Big\{\nabla f(x^*), \ldots, \nabla g_m(x^*))\Big\}$ if and only if $h\nabla g^i(x^*) > 0$, for all $i \in \{i, \ldots, r\}$. Recall that:

$$g_i(x^* + \varepsilon h) - g_i(x^*) = \varepsilon h \nabla g_i(x^*) + \mathcal{O}(\varepsilon)$$

and for small enough ε, we have:

$$g_i(x^* + \varepsilon h) > g_i(x^*) \forall i \in \{0, 1, \ldots, r\}$$

and for $i > r$, we have:

$$g_i(x^* + \epsilon h) \approx g_i(x^*) + \varepsilon h^T \nabla g_i(x^*) \geq 0$$

Therefore, $x^* + \varepsilon h$ is feasible. Then:

$$f(x^* + \varepsilon h) > f(x^*)$$

implies that x^* is not a local max. Then we have reached a contradiction. □

We can see the following example.

Example 14.1.1.

$$\max_{x,y} x^2 - y^2$$

$$s.t.$$

$$g(x, y) = (x - 1)^3 - y^2$$

The optimal solution is $(x^*, y^*) = (1, 0)$. *The gradients are:*

$$\nabla f(x^*) = \begin{bmatrix} -2x \\ -2y \end{bmatrix} = \begin{bmatrix} -2 \\ 0 \end{bmatrix}$$
$$\nabla g(x^*) = \begin{bmatrix} 3(x-1)^2 \\ -2y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

By the Fritz John Conditions:

$$\mu_0 \begin{bmatrix} -2 \\ 0 \end{bmatrix} + \mu_1 \begin{bmatrix} 0 \\ 0 \end{bmatrix} = 0$$

This is true if and only if $\mu_0 = 0$.

Therefore, according to these conditions, we can have two situations, where $\mu_0 = 0$ and where $\mu_0 \neq 0$. We can generalize this by imposing some constraint qualifications, namely that $\nabla g_i(x^*)$, for all $i \in M$ must be **linearly independent**.

This is the first of several results, which will be presented in the next pages, and that represent the necessary conditions for solving constrained optimization problems.

14.2 Constrained Optimization: necessary conditions

Theorem 14.2.1. *(Karush-Kuhn-Tucker Theorem I) Suppose $M \neq \emptyset$, x^* is a local maximum for an optimization problem. If $\{\nabla g_i(x^*)\}$ are **linearly independent** (constraint qualification), then we can find **non-zero multipliers** $\{\lambda_i\}_{i\in M}$ such that:*

1. *First Order Conditions*

$$\nabla f(x^*) + \sum_{i=1}^{M} \lambda_i g_i(x^*) = 0$$

2. *and Complementary Slackness*

$$\lambda_i g_i(x^*) = 0; \qquad \forall i \in M$$

Are satisfied.

A more general problem can expressed with inequality and equality constraints as follows:

$$\begin{gathered} \max_x f(x) \\ \text{s.t.} \\ g_i(x) \geq 0; \quad \forall i \in M \\ g_j(x) = 0 \qquad \forall j \in N \end{gathered}$$

Where $f, g_i, g_j : \mathbb{R}^n \to \mathbb{R}$. In this example, we have two sets of constraints. Then we have the following definition.

Definition 14.2.1. $g_i(x) \geq 0$ is **effective** of x^* if the constraints bind, i.e. $g_i(x^*) = 0$. We define M^e as the set of effective constraints.

Thus, we have the following result.

Theorem 14.2.2 (Karush-Kuhn-Tucker II)**.** *Let x^* be a local max for an optimization problem. If $\{\nabla g_i(x^*) : i \in M^e \cup N\}$ are **linearly independent**, then it exists a multiplier $\{\lambda_i\}$ such that:*

$$\nabla f(x^*) + \sum_{i\in M\cup N} \lambda_i \nabla g_i(x^*) = 0 \tag{14.4}$$

$$\begin{aligned} &\lambda_i g_i(x^*) = 0 \\ &\forall i \in M, \lambda_i \geq 0, g_i(x^*) \geq 0 \end{aligned} \tag{14.5}$$

$$\begin{aligned} &\lambda_i g_i(x^*) = 0 \\ &\forall i \in N, \lambda_i \text{ is unrestricted} \end{aligned} \tag{14.6}$$

We have unrestricted λ_i for equality constraints because of the representation as a primal/dual problem seen before. Then, when we have x unrestricted in the primal, we need equality in the dual.

Another version of the KKT theorem, with linear constraints, is the following.

Theorem 14.2.3 (Karush-Kuhn-Tucker Theorem (Linear))**.** *Let x^* be a local max for the following problem:*

$$\max_x f(x)$$
$$s.t.$$
$$Ax \geq b$$

Where A is a $(m \times n)$, b is $(n \times 1)$ and f is continuously differentiable.

Then it exists a non-zero $\lambda \in \mathbb{R}^m_+$ such that:

1. *First Order-Conditions*
$$\nabla f(x^*) + A^T\lambda = 0$$
2. *and Complementary Slackness*
$$\lambda_i A x^* = 0$$

are satisfied.

Proof. We can express the constraint in matrix form:

$$g_i(x) = \sum_{i=1}^{n} a_{ij}x_j - b_i$$

Then:

$$\frac{\partial g_i(x)}{\partial x_j} = a_{ij}$$

and:

$$\nabla f(x) = \begin{bmatrix} \nabla g_1(x) & \dots & \nabla g_n(x) \end{bmatrix}$$

and:

$$A^T = \begin{bmatrix} a_{11} & \dots & a_{m1} \\ \vdots & \vdots & \vdots \\ a_{1n} & \dots & a_{mn} \end{bmatrix}$$

We want to show that $\nabla f(x^*) + \sum_{i=1}^{n} \lambda_i \nabla g_i(x^*) = 0$. Suppose that the first r constraints bind. We want to show that exists a $\lambda \in \mathbb{R}_+$ such that:

$$\nabla f(x^*) + \sum_{i=1}^{n} \lambda_i \nabla g_i(x^*) = 0$$

In other words, that $-\nabla f(x^*)$ is in the Cone spanned by $[\nabla g_1(x), \dots, \nabla g_m(x^*)]$. Suppose not, then, by Farkas' Lemma, there is some vector $h \in \mathbb{R}^n$ such that:

- $h^T \nabla g_i(x^*) \geq 0$, for all $i \in \{1, \cdots, r\}$
- $h^T \nabla f(x^*) > 0$

Consider $x^* + \varepsilon h$, then:

- $x^* + \varepsilon h$ leads to an higher value

$$f(x^* + \varepsilon h) = f(x^*) + \underbrace{\varepsilon h^T \nabla f(x^*)}_{>0}$$

which implies:

$$f(x^* + \varepsilon h) > f(x^*)$$

- Further, we want to show that $x^* + \varepsilon h$ is feasible when ε is small. That is:

$$A(x^* + \varepsilon h) \geq b$$

Implies:

$$\sum_{i=1}^{n} a_{ij} x^* + \varepsilon \nabla g_i(x^*) h \geq b$$

Then we have two cases: (i) $i = r$:

$$\sum_{i=1}^{L} a_{ij} x_j^* = b_i$$

Since we also have $\nabla g_i(x^*) h \geq 0$, the i^{th} constraint holds; (ii) $i >$:

$$\sum_{i=1}^{n} a_{ij} x_j^* > b_i$$

So the i^{th} constraint holds if ε is small. But this contradicts the fact that x^* is a local max.

□

The following theorem establishes other necessary conditions when the constraints are concave functions.

Theorem 14.2.4 (Karush-Kuhn-Tucker-Slater)**.** *Let x^* be a local max for the problem:*

$$\max_x f(x)$$

$$s.t.$$

$$g_i(x) \geq 0 \quad \forall i \in M$$

All $g_i(x)$ are concave functions for $i \in M$, and f, g_i are continuously differentiable. Suppose that there exist a $x^o \in \mathbb{R}^n$ such that $g_i(x^o) > 0$, for all $i \in M$. Then, there exist multipliers $\lambda_i \geq 0$ such that:

1. *First Order Conditions*

$$\nabla f(x^*) + \sum_{i=1}^{M} \lambda_i \nabla g_i(x^*) = 0$$

2. *and Complementary Slackness*

$$\lambda_i g_i(x^*) = 0; \qquad \forall i \in M$$

Are satisfied.

Proof. Recall, from Fritz John Conditions, that:

$$\lambda_0 \nabla f(x^*) + \sum_{i=1}^{n} \lambda_i \nabla g_i = 0$$

We want to show that $\lambda_0 > 0$. Suppose not. Then:

$$\sum_{i=1}^{n} \lambda_i \nabla g_i(x^*) = 0$$

Since $g_i(\cdot)$ is concave, then:

$$g_i(x^*) + \nabla g_i(x^*)(x^o - x^*) > g_i(x^*)$$

This can be written as:

$$\sum_{i=1}^{n} \underbrace{\lambda_i g_i(x^*)}_{=0 \text{ by CS}} + \sum_{i=1}^{n} \underbrace{\lambda_i \nabla g_i(x^*)}_{=0 \text{ by CS}} (x^o - x^*) \geq \sum_{i=1}^{n} \lambda_i g_i(x^o) > 0$$

Therefore $0 > 0$, we have reached a contradiction. □

Here is a recap of the **necessary** conditions for local max. If x^* is a local max, then we can use the Fritz John Conditions when $\lambda_0 \neq 0$. Otherwise, we need one of the following constraints qualifications:

1. All gradients must be **linear independent**
2. The constraints must be **linear**
3. The constraints must be **concave**, and therefore we have interior points.

14.3 Constrained Optimization: sufficient conditions

Theorem 14.3.1 (Karush-Kuhn-Tucker: sufficient I)**.** *Consider the following problem:*

$$\max_x f(x)$$

$$s.t.$$

$$Ax \geq b$$

*Where f is **concave**, and continuously differentiable, and A is $(m \times n)$ and b is $(m \times 1)$. Let x^* be a feasible solution of this problem. Then x^* is an optimal solution if and only if it exists $\lambda \in \mathbb{R}^m_+$ such that:*

1. *First Order Conditions*
$$\nabla f(x^*) + A^T\lambda = 0$$
2. *and Complementary Slackness*
$$\lambda^T(Ax^* - b) = 0$$

Are satisfied.

Proof. the necessary part has already been proven in KKT linear. Let's see the sufficient part. We define a function:

$$h(x) = f(x) + \lambda^T(Ax - b)$$

$h(x)$ is concave because $f(\cdot)$ is concave and this is a linear function. The FOCs are:

$$\nabla h(x^*) = \nabla f(x^*) + A^T\lambda = 0$$

Where x^* is a global max of $h(x)$. This implies:

$$h(x^*) \geq h(x)$$

For all feasible x. Indeed:

$$f(x^*) = f(x^*) + \underbrace{\lambda^T(Ax^* - b)}_{=0 \text{ by CS}}$$

Therefore:

$$h(x^*) \geq h(x)$$

where:

$$h(x) = f(x) + \lambda^T(Ax - b) \geq f(x)$$

for all feasible x, and therefore x^* is a global max. □

Let's see some examples now.

Example 14.3.1.

$$\min_{x_1,x_2,x_3} \frac{1}{2}x_1^2 + \frac{1}{2}x_2^2 + \frac{1}{2}x_3^2$$

$$s.t.$$

$$x_1 + x_2 + x_3 = 3$$

Notice that this is a concave function. Then sufficient conditions hold. Taking the FOCs, we can see that the solution is $x_1 = x_2 = x_3 = \lambda = 1$.

Example 14.3.2. *Let's see the following problem:*

$$\max_{x_1,x_2} -x_1^2 - 2x_2^2 - 4x_1x_2$$

$$s.t.$$

$$x_1 + x_2 = 1$$

$$x_1, x_2 \geq 0$$

The constraints are linear. However, the function is not concave. indeed, the hessian matrix:

$$\nabla^2 f(x^*) = \begin{bmatrix} -2 & -4 \\ -4 & -2 \end{bmatrix}$$

is not negative definite but indefinite. So sufficient conditions cannot apply. Therefore, we need to check for necessary conditions. Any x^* *which solves the problem must satisfy the FOCs and the Complementary Slackness conditions. Before finding a solution by solving the FOCs and the CS, let's be sure that a solution exists: a solution exists because the function is continuous and the constraint set is compact (Weierstraß's theorem). By KKT (linear), an optimal* x^* *must satisfy:*

- *First Order Conditions*
- *Complementary Slackness*
- *Feasibility*

Let's write the Lagrangian function:

$$\mathcal{L}(x_1, x_2, \lambda) = -x_1^2 - 2x_2^2 - 4x_1x_2 + \lambda[1 - x_1 - x_2] + \mu_1 x_1 + \mu_2 x_2]$$

The FOCs are:

$$\frac{\partial \mathcal{L}}{\partial x_1} = -2x_1 + 4x_2 - \lambda + \mu_1 = 0$$
$$\frac{\partial \mathcal{L}}{\partial x_2} = -4x_1 - 4x_2 - \lambda + \mu_2 = 0$$

and the Complementary Slackness are:

$$\mu_1 x_1 = 0$$
$$\mu_2 x_2 = 0$$

Notice that we don't need the complementary slackness condition for λ and the equality constraint. Indeed, because of the equality, the constraint is already equal to zero and slackness holds automatically. Complementary Slackness is necessary only when there is the possibility of slack. Further, we have the feasibility conditions:

$$x_1 + x_2 = 1$$
$$x_1 \geq 0$$
$$x_2 \geq 0$$
$$\mu_1, \mu_2 \geq 0$$

Let's see case by case. Case 1, $\mu_1, \mu_2 = 0$. Then the FOCs are:

$$\frac{\partial \mathcal{L}}{\partial x_1} = -2x_1 + 4x_2 - \lambda = 0$$
$$\frac{\partial \mathcal{L}}{\partial x_2} = -4x_1 - 4x_2 - \lambda = 0$$

And $x_1 + x_2 = 1$. Solving, it gives: $\lambda = -4, x_1 = 0, x_2 = 1$).

Case 2, $\mu_1, \mu_2 \geq 0$. In this case, by Complementary Slackness, we have $x_1, x_2 = 0$, which clearly violates $x_1 + x_2 = 1$.

Case 3, $\mu_1 > 0$ and $\mu_2 = 0$. Then we have: $x_1 = 0$ and $x_2 = 1$ (by $x_1 + x_2 = 1$).

Case 4, $\mu_1 = 0$ and $\mu_2 > 0$. Then we have: $x_1 = 1$ and $x_2 = 0$ (by $x_1 + x_2 = 1$). However, for cases 3 and 4 we need to check also the FOCs, in order to exclude negative multipliers μ_1 and μ_2. Then, for the case 3:

$$4 - \lambda + \mu_1 = 0$$
$$-4 - \lambda = 0$$

Then $\lambda = -4$ and $\mu_1 = 8$.

For case 4:

$$-2 - \lambda = 0$$
$$-4 - \lambda + \mu_2 = 0$$

we have $\lambda = -2$ and $\mu_2 = 2$. Then, multipliers feasibility holds. To find the optimal $x^* = (x_1^*, x_2^*)$, we just plug in the objective function. Then:

$$f(0,1) = -2 < f(1,0) = -1$$

Thus $x^* = (1,0)$ is optimal.

There is a second sufficient condition for KKT Theorem holding.

Theorem 14.3.2 (Karush-Kuhn-Tucker: sufficient II)**.** *Let x^* be a feasible solution of:*

$$\max_x f(x)$$

$$s.t$$

$$g(x) \geq 0 \quad \forall i \in M$$

$$g_i(x) = 0 \quad \forall i \in N$$

Where f, g are continuously differentiable and concave, and g_i are affine functions (linear). If there $\lambda_1, \ldots, \lambda_m \geq 0$ and $\mu_1, \ldots, \mu_n \in \mathbb{R}$ such that:

1. *First Order Conditions*

$$\nabla f(x^*) + \sum_{i=1}^{m} \lambda_i \nabla g(x^*) + \sum_{i=1}^{n} \mu_i \nabla g_i(x^*) = 0$$

2. *Complementary Slackness (just for the inequality constraint)*

$$\lambda_i(g(x^*)) = 0$$

Then x^* is a global max.

Proof. Define

$$h(x) = f(x) + \sum_{i=1}^{m} \lambda_i g(x) + \sum_{i=1}^{n} \mu_i g_i(x)$$

$h(x)$ is concave because it is the sum of concave functions and $\nabla h(x^*) = 0$ implies that $h(x^*) \geq h(x)$ for all feasible x.

Then:

$$\tilde{f}(x^*) = f(x^*) + \underbrace{\sum_{i=1}^{m} \lambda_i \nabla g(x^*)}_{=0 \text{ by CS}} + \underbrace{\sum_{i=1}^{n} \mu_i \nabla g_i(x^*) = h(x^*)}_{=0 \text{ by CS}} \geq$$

$$\geq h(x) = f(x) + \underbrace{\sum_{i=1}^{m} \lambda_i g(x)}_{>0} + \sum_{i=1}^{n} \mu_i g_i(x) > f(x)$$

□

Let's see one example.

Example 14.3.3. *Let's see the following problem:*

$$\max_{x_1,x_2} x_1 + x_2 - 4x_1^2 - x_2^2$$

$$s.t.$$

$$2x_1 + x_2 \leq 1$$

$$x_1^2 \leq 1$$

Both the objective function and the constraints are concave. Besides, we have only inequality constraints, so we don't need to check for linearity of equality constraints. We can apply KKT ST2. Further, since Slater conditions hold too, KKT ST2 is both necessary and sufficient.

Let's solve the Lagrangian:

$$\mathcal{L} = x_1 + x_2 - 4x_1^2 - x_2^2 + \lambda_1(1 - x_2 - 2x_1) + \lambda_2(1 - x^2)$$

The KKT conditions are:

$$\frac{\partial \mathcal{L}}{\partial x_1} = 1 - 8x_1 - 2\lambda_1 - 2\lambda_2 x_1 = 0$$
$$\frac{\partial \mathcal{L}}{\partial x_2} = 2 - 2x_2 - \lambda_1 = 0$$

The Complementary Slackness

$$\lambda_1(1 - 2x_1 - x_2) = 0$$
$$\lambda_2(1 - x^2) = 0$$

The feasibility conditions for the Primal

$$2x_1 + x_2 \leq 1$$
$$x_2^1 \leq 1$$

The feasibility of the dual:

$$\lambda_1, \lambda_2 \geq 0$$

We have 4 cases.

1. *Case 1.* $\lambda_1 = \lambda_2 = 0$*, all constraints are slack. So we have:*

$$8x_1 - 1 = 0 \Rightarrow x_1 = \frac{1}{8}$$

$$2x_2 - 2 = 0 \Rightarrow x_2 = 1$$

But then:

$$2\frac{1}{8} + 1 \geq 1$$

Feasibility is violated.

2. $\lambda_1, \lambda_2 > 0$, *all constraints bind. Then:*

$$2x_1 + x_2 = 1$$
$$x_2 = 1$$

and:

$$x_1 = 1 \quad x_2 = -1$$
$$x_2 = -1 \quad x_2 = 3$$

Plug $(1, -1)$ *into the FOCs:*

$$1 - 8x_1 - 2\lambda_1 - 2\lambda_2 x_1 = 0$$
$$2\lambda_1 - 2\lambda_2 = -7$$

But this is not possible, since $\lambda_1, \lambda_2 \geq 0$. *Plug* $(-1, 3)$ *into:*

$$2 - 2x_2 - \lambda_1 = 0$$

then $\lambda = -4$. *Again, this is not possible*

3. $\lambda_1 > 0, \lambda_2 = 0$. *From Complementary Slackness, the first constraint binds. Solving the FOCs, we have the following possible solutions:*

$$\begin{cases} x_1 = \frac{1}{16} \\ x_2 = \frac{7}{8} \\ \lambda_1 = \frac{1}{4} \\ \lambda_2 = 0 \end{cases}$$

This can be a solution. Let's see the last case.

4. $\lambda_1 = 0, \lambda_2 > 0$. *The second constraint binds. Then:*

$$1 - x^2 \Rightarrow x = \pm 1$$

From the second FOCs:

$$2 - 2x_2 = 0 \Rightarrow x_2 = 1$$

But by plugging $(-1, 1)$ *into the first FOCs, we have:*

$$\lambda_1 = -\frac{9}{2}$$

Plugging $(1, 1)$ *into the feasibility constraint:*

$$2 \cdot 1 + 1 > 1$$

So this is violated.

Therefore, the unique solution to this problem is:

$$\begin{cases} x_1 = \frac{1}{16} \\ x_2 = \frac{7}{8} \\ \lambda_1 = \frac{1}{4} \\ \lambda_2 = 0 \end{cases}$$

14.4 Duality in Non-Linear Programming

So far we have seen the following problem (**primal**):

$$V_P = \max_x \in \mathbb{R}_+ f(x)$$

$$\text{s.t}$$

$$g(x) \geq 0$$

But recall that when dealing with linear optimization, the idea of duality was introduced. This simply provided a possibly simpler way to solve an optimization problem, namely by reducing the number of constraints. The same idea applies in the case of non-linear optimization. The idea, in a nutshell, is that of introducing a "penalty" if the optimality conditions are violated.

From the Lagrangian:

$$\mathcal{L}(x, \lambda, \mu) = f(x) + \sum_{i=1}^{m} \lambda_i g(x) + \mu_i^m g(x)$$

we can define the **dual function** as follows:

$$q(\lambda, \mu) = \max_{x \in \mathbb{R}} \mathcal{L}(x, \lambda, \mu)$$

where $\lambda \in \mathbb{R}^m, \mu \in \mathbb{R}^n$. Notice that this is an unconstrained optimization. Suppose that you violate the constraints so that $g(x) \leq 0$ or $g(x) \neq 0$. This is why we need λ, μ. Furthermore, $q(\lambda, \mu)$ is convex, and we can define its domain as follows:

$$Dom(q) = \left\{ (\lambda, \mu) \in \mathbb{R}^m \times \mathbb{R}^n : q(\lambda, \mu) < \infty \right\}$$

Therefore, we can define the **dual problem**:

$$V_D = \min q(\lambda, \mu)$$

$$\text{s.t.}$$

$$(\lambda, \mu) \in Dom(q)$$

Why the constraints will bind the dual problem? The idea is the dual problem must provide an upper bound to the primal problem. Namely, fixing λ, μ, then for all $x \in F$, $f(x) + \sum_{i=1}^n \lambda_i g(x) + \sum_{i=1}^n g(x) \geq f(x)$. If we take the max on the left-hand-side, we have:

$$\max_{x \in \mathbb{R}^n} f(x) + \sum_{i=1}^{n} \lambda_i g(x) + \sum_{i=1}^{n} \mu g(x) \geq f(x) \geq$$

$$\geq \max_{x \in F} \mathcal{L}(x, \lambda, \mu) \geq \max_{x \in F} f(x) = V_P$$

This holds for all possible (λ, μ). So we need to choose λ, μ such that:

$$V_D = \min_{\lambda, \mu \in Dom(q)} q(\lambda, \mu) \geq V_P$$

Let's see the following important results.

- Weak Duality: $V_P \leq V_D$
- Strong Duality: if, in the primal, the objective function and the non-linear constraints are concave, the linear constraint is affine, and, finally the primal problem has a finite value, then the dual is also finite and $V_D = V_P$.

Before offering a formal description of these results, and their proofs, let's see some examples.

Example 14.4.1.

$$V_P = \max -x_1^2 - x_2^2 - 2x_1$$
$$x_1 + x_2 = 0$$

The value of the primal is $x^* = \Big(-\frac{1}{2}, \frac{1}{2} \Big)$ $V_P = \frac{1}{2}$. *Let's solve the dual problem:*

$$\mathcal{L}(x, \lambda) = -x_1^2 - x_2^2 - 2x_1 + \lambda(-x_1 - x_2)$$

Solving for the dual function, we have:

$$q(\lambda) = \max_x \mathcal{L}(x, \lambda) = -x_1^2 - x_2^2 - 2x_1 + \lambda(-x_1 - x_2)$$

Taking the FOCs:

$$\frac{\partial \mathcal{L}}{\partial x_1} = -2x_1 - 2 - \lambda = 0 \Rightarrow x_1 = -\frac{\lambda}{2} - 1$$
$$\frac{\partial \mathcal{L}}{\partial x_2} = -2x_2 - \lambda = 0 \Rightarrow x_2 = -\frac{\lambda}{2}$$

Therefore:

$$q(\lambda) = -\Big(-\frac{\lambda}{2} - 1 \Big)^2 - \Big(-\frac{\lambda}{2} \Big)^2 - 2\Big(\frac{\lambda}{2} - 1 \Big)$$

Which can be simplified as:

$$q(\lambda) = \frac{\lambda^2}{4} - \lambda + 1 + \frac{\lambda^2}{4} - 2\lambda - 2 =$$
$$\frac{\lambda^2}{2} + \lambda + 1$$

This is the dual function. So the dual problem is:

$$V_D \min_{\lambda \in Dom(q)} = \frac{\lambda^2}{2} + \lambda + 1$$

Therefore $\lambda^* = -1$ *and the solution to the dual problem is:*

$$V_D \Rightarrow \frac{-1^2}{2} - 1 + 1 = \frac{1}{2} = V_P$$

Example 14.4.2.

$$V_P = \max c^T x$$

$$s.t.$$

$$Ax \leq b$$

$$x \geq 0$$

The dual can be written as:

$$q(\lambda, \mu) = \max_x \left[c^T x + \lambda^T (b - Ax) + \mu^T x \right] =$$
$$\max \left[\lambda^T b + (c^T - \lambda^T A + \mu^T) x \right]$$

The objective function is linear so:

$$x^* = \begin{cases} +\infty & if\ c^T - \lambda^T A + \mu^T \neq 0 \\ \lambda^T b & if\ c^T - \lambda^T A + \mu^T = 0 \end{cases}$$

The domain of the dual function is:

$$Dom(q) = \left\{ (\lambda, \mu) : c^T - \lambda^T A + \mu^T = 0 \right\}$$

The dual problem is:

$$V_D = \min q(\lambda, \mu)$$

$$s.t$$

$$(\lambda, \mu) \in Dom(q)$$

This can be written as:

$$\min \lambda^T b$$

$$s.t$$

$$c^T = \lambda^T A - \mu^T$$

But if we get rid of μ (and we can since $\mu > 0$), we have:

$$\min \lambda^T b$$

$$s.t$$

$$c^T \leq \lambda^T A$$

$$\lambda \geq 0$$

This is the statement of the dual in linear programming.

Let's see now the Strong Duality Theorem.

Theorem 14.4.1 (Strong Duality Theorem)**.** *If*

1. $f, g_i, \forall i \in M$ *are concave*
2. $g_i, i \in N$ *are affine.*
3. *it exists* $x \in F$ *such that* $f^i(x) > 0, \forall i \in M$

Then $V_P = V_D$*, where* V_D *is the solution to the dual problem:*

$$V_D = \min_{\lambda,\mu} q(\lambda, \mu) = \min \Big[\max \mathcal{L}(x, \lambda, \mu) \Big]$$

$$s.t.$$

$$(\lambda, \mu) \in Dom(q)$$

Before seeing the proof of this theorem, let's see the following lemma.

Lemma 14.4.2. *Under the conditions of the Strong Duality Theorem, let* $c \in \mathbb{R}$*, then the following are equivalent:*

1. $x \in \mathbb{R}^n$*,* $g_i(x) \geq 0, \forall i \in M$*,* $g_i(x) \forall i \in N$*, then* $f(x) \leq c$
2. *It exists a* $\lambda \in \mathbb{R}_+^M$ *and* $\mu \in \mathbb{R}^N$ *such that:*

$$g(\lambda, \mu) = \max \left\{ f(x) + \sum_{i=1}^{M} \lambda_i g_i(x) + \sum_{i=1}^{N} g_i(x) \right\} \leq c$$

This lemma states that if x is feasible (1) for a primal problem, then c is an upper bound to the primal problem. Then, we can always find (λ, μ) such that the dual problem can achieve values $\leq c$. Roughly speaking, if the primal problem is bounded by c, the dual problem takes values less than c. Further, the strong duality holds.

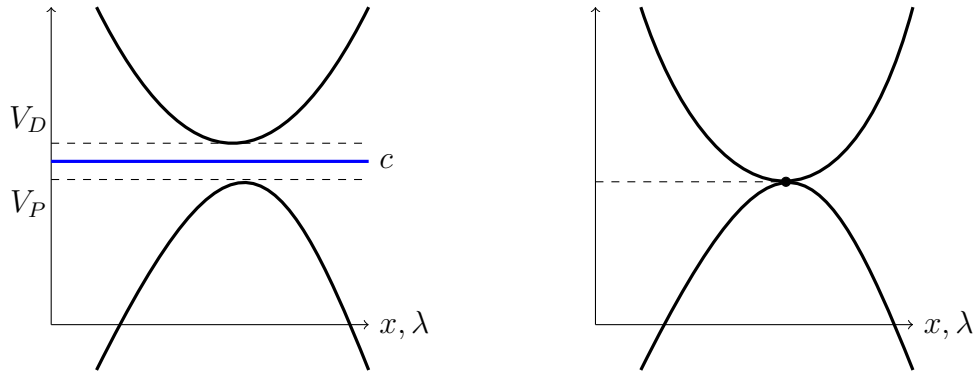


Figure 14.1: Duality

In the figure on the left, we see that the primal achieves a maximum value that is less than c. c is the upper bound of the dual, but cannot be achieved by that. In other words, there is a **duality gap**. Therefore, the only possible case is when the two problems achieve the same value, as depicted in the figure on the right.

Let's see the proof of the lemma and then that of the Strong Duality Theorem.

Proof of the Lemma. Let's see 2) $\Rightarrow$ 1). If it exists (λ, μ) such that:

$$\max_x \left\{ f(x) + \left\{ f(x) + \sum_{i=1}^{M} \lambda_i g_i(x) + \sum_{i=1}^{N} g_i(x) \right\} \leq c \right.$$

Then, we can take $\hat{x} \in F$, and therefore:

$$f(\hat{x}) \underbrace{\sum_{i=1}^{M} \lambda_i g_i(\hat{x})}_{\geq 0} + \underbrace{\sum_{i=1}^{N} g_i(\hat{x})}_{=0} \leq c$$

Therefore $f(\hat{x}) \leq c$.

Let's see now 1) $\Rightarrow$ 2). Assume for simplicity that $N \neq \emptyset$. We want to show that if c is an upper bound of the primal, then it can be achieved by the dual. Define:

$$S = \left\{ u = (u_0, u_1, u_2, \ldots, u_m) : \text{it exists } x \in \mathbb{R}^n, \text{ s.t. } f(x) \geq u_0, g_i(x) \geq u_i, \forall i \in M \right\}$$

Define further:

$$T = \left\{ u = u_= \geq c, u_1 \geq 0, \ldots, u_n > 0 \right\}$$

such that S, T are not empty and convex. Convex because S is the intersection of the subgraphs of concave functions (which is, the intersection of convex sets is convex itself). Furthermore, their intersection is empty. Therefore, we can apply the Separating Hyperplane Theorem. It exists $a = (a_0, a_1, \ldots, a_m) \neq 0$ that separates S and T such that:

$$\min_{u \in S} a \cdot u \geq \sup_{u \in T} a \cdot u \tag{14.7}$$

(we use sup instead of max because T is not a closed set). We can claim that $(a_0, a_1, \cdot, a_m) < 0$. Suppose not, then $a_i > 0$. We can let $u_i \to \infty$ so that the righthand-side above is $+\infty$. But this contradicts S being not empty and therefore finite. So, the right-hand-side of (14.7) is $a_0 \cdot c$. We want to show that $a_0 \neq 0$. Suppose not, then $a_0 = 0$. We have:

$$\min_{u \in S} a \cdot u = \min_{u \in S} \sum_{i=1}^{m} a_i \cdot u_i$$

Let $u_i = f(x_0)$, where x_0 satisfies Slater conditions. So $u_i > 0$ implies $\sum_{i=1}^{m} a_i \cdot u_i < 0$. Then the left-hand side of (14.7) is less than 0, and the right-hand side equals 0. We have reached a contradiction. Then $a_0 \neq 0$.

We can write:

$$\min_{u \in S} u \cdot u \geq a_0 \cdot c$$

This implies that:

$$\max_{u \in S} \frac{a}{a_0} \cdot u \leq c$$

Let $\lambda_i = \frac{a}{a_0} \geq 0 \quad \forall i \in M$. Then:

$$\max \left[u_0 + \sum_{i=1}^{m} \lambda_i \cdot u_i \right] \leq c$$

□

Proof of the Strong Duality Theorem. Define the set $\hat{S}$ as:

$$\hat{S} = \left\{ u = (u_0, u_1, \cdot, u_n) : f(x) = u_0, f_i(x) = u_i, \forall i \in M \right\}$$

Then $\hat{S} \subset S$. This implies that:

$$\max_{\lambda \in \hat{S}} \left[u_0 + \sum_{i=1}^{n} \lambda_i u_i \right] \leq \max_{u \in S} \left[u_0 + \sum_{i=1}^{m} \lambda_i u_i \right]$$

And:

$$\max_{x} \left[f(x) + \sum_{i=1}^{n} \lambda_i g_i(x) \right] \leq c$$

Therefore, it exists a λ such that $\min q(\lambda) \leq c$. □

Chapter 15

Some Elements of Dynamic Programming

So far, we have considered a series of problems that do not involve time. However, especially when dealing with choices, what people do at a certain moment influences what comes after. Therefore, an agent must solve a sequential problem to find the sequence of values that optimize a function for a potentially infinite period of time, where at each time period, the function remains the same, but at the same time, the optimal value is influenced by choices made before. In other words, an optimal solution for an optimization problem must be optimal for all the periods t, with, in the most general case, $t \to \infty$.

The way of solving these types of problems is through a **recursive** approach, using Dynamic Programming. This chapter will be divided into two sections: the first will present a recap of some results concerning the continuity of correspondences, useful to establish some useful results, namely three fixed point theorems; in the second section, the main aspects of Dynamic Programming will be discussed.

15.1 Some Fixed Point Theorems

To establish the existence of some important results in economics, a technique often used is that of using a fixed point theorem, namely to show that a value of x such that

$$f(x^*) = x^*$$

exist. This, in general, can be done for any mapping, not only functions. Therefore, before discussing the three most important fixed point results used in economics, namely the **Brouwer Fixed Point Theorem**, the **Kakutani Fixed Point Theorem**, and the **Banach Fixed Point Theorem**, a recap of the notion of correspondences and continuity is needed.

Definition 15.1.1. Consider $X \subseteq \mathbb{R}^n$ and $Y \subseteq \mathbb{R}^n$. A **correspondence** $\Gamma : X \rightrightarrows Y$ is a set-valued mapping from X into $2^Y \setminus \emptyset$.

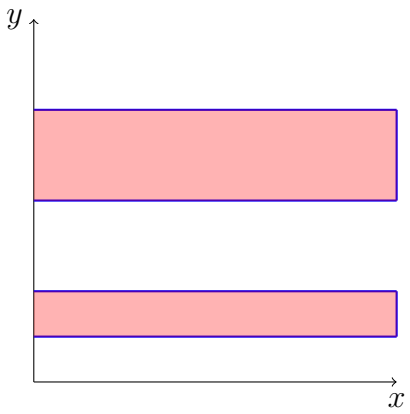


Figure 15.1: A non-convex valued correspondence

Let's see some examples:

Example 15.1.1. *An example in economics is that of* ***budget sets****. For any* $n \in \mathbb{N}, p \in \mathbb{R}^n_{++}$*, and* $w \in \mathbb{R}_{++}$*, we define* $B_{p,w}$ *as:*

$$B_{p,w} = \left\{ x \in \mathbb{R}^n : p \cdot x \leq w \right\}$$

This is a correspondence. Indeed it can be written as:

$$B_{p,w} : \mathbb{R}^m_{++} \rightrightarrows 2^{\mathbb{R}^n_+} \setminus \emptyset$$

A correspondence can have different characterizations.

Definition 15.1.2. A correspondence is:

1. **Single-valued** at x if $\Gamma(x)$ is a singleton
2. **Closed-valued** at x if $\Gamma(x)$ is a closed set
3. **Compact-valued** at x if $\Gamma(x)$ is compact
4. **Convex-valued** at x if $\Gamma(x)$ is a convex set

Recall that a very important feature of functions is their continuity. Functions are a special kind of correspondence. However, to define the continuity of correspondences, we need a more complicated idea, resting upon the notions of **upper** and **lower hemi-continuity**.

Definition 15.1.3. A correspondence $\Gamma : X \rightrightarrows Y$ is **upper hemi-continuous** at x is for all open subsets $O \subseteq Y$, with $O \supseteq \Gamma(x)$, there exists some $\epsilon > 0$ such that $\Gamma(N_\epsilon(x) \cap X) \subseteq O$ (the image of the $\epsilon-$neighborhood is also in O).

Γ is upper hemi-continuous on $S \subset X$, if it is upper hemi-continuous for all $x \in S$.

The basic idea is that a small perturbation of x does not change its image "a lot," namely, it does not suddenly expand. Then, we have the following result.

Proposition 16. *Γ is upper hemi-continuous at x if, for any sequence $\{x_n\}$ and $\{y_n\} \subset Y$ with $x_n \to x$ and $y_n \in \Gamma(x_n)$ for each n, there exists a subsequence of $\{y_n\}$ that converges to a point in $\Gamma(x)$. If Γ is compact-valued, then the converse is also true.*

Graphically:

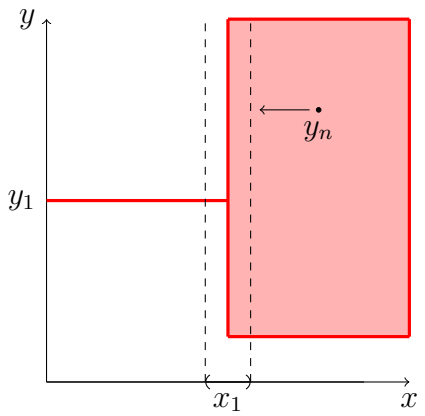


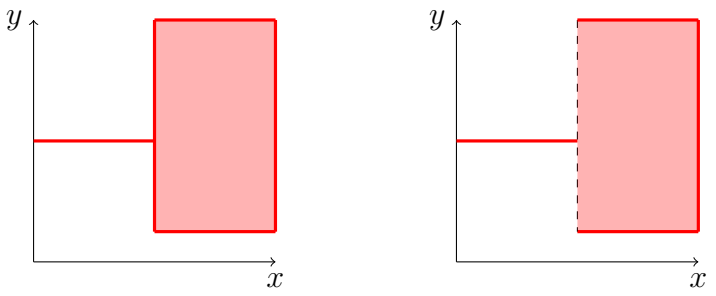
Figure 15.2: A upper hemi-continuous correspondence

Since the graph is closed, any sequence into $\Gamma(x)$ converges inside $\Gamma(x)$.
A related concept is the **closed graph property**,

Definition 15.1.4. The graph of Γ is:

$$Gr(\Gamma) = \Big\{(x, y) \in X \times Y : y \in \Gamma(x)\Big\}$$

Γ has the closed graph property if the graph is closed.



(a) Closed graph (b) Non-closed graph

Using this property, we have a simpler characterization of upper hemi-continuity.

Proposition 17. *Let $\Gamma : X \rightrightarrows Y$ be a correspondence. If Γ has the closed graph and Y is compact, then it is upper hemi-continuous (sufficient condition). Further, if Γ is upper hemi-continuous and closed-valued, then it has a close graph.*

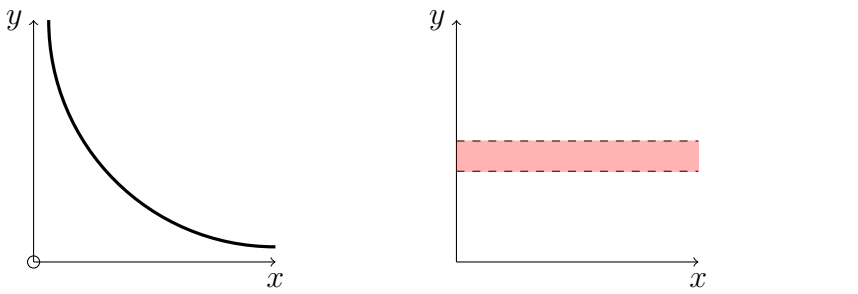
However, some examples show that we can have the closed-graph property but not upper hemi-continuity.

Example 15.1.2. *Let's see a first example. Take the function:*

$$\Gamma(x) = \begin{cases} \frac{1}{x} \ \textit{for } x \in [0,1] \\ \{0\} \ \textit{for } x = 0 \end{cases}$$

The graph of this function is closed, but not upper hemi-continuos, since it is not compact (it is not bounded). Notice, however, that even without applying the definition, it is clear that this cannot be upper hemi-continuous since the function is not continuous.

A second example is the correspondence $\Gamma(x) = (0.5, 1.5)$. This is upper hemi-continuous since it is a constant mapping, but the graph is not closed.



(a) Closed graph but not u.h.c. (b) Non-closed graph but u.h.c.

A related definition is that of **lower hemi-continuity**.

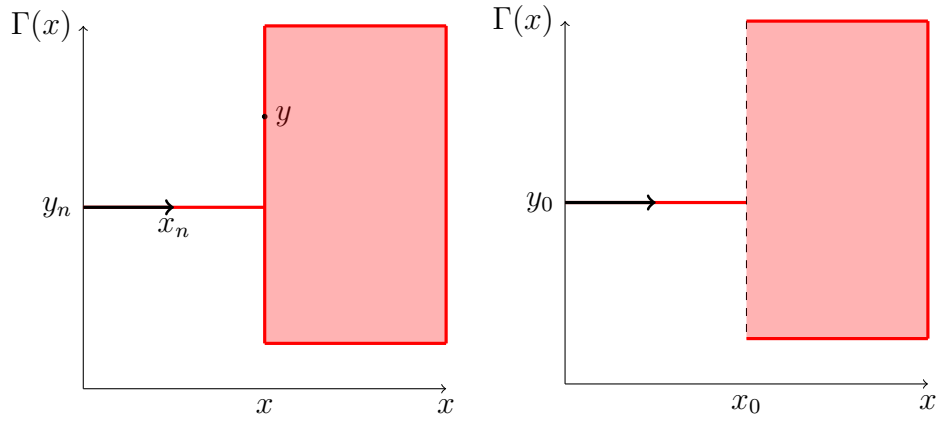
Definition 15.1.5. A correspondence is **lower hemi-continuous** if for all open sets $O \subset Y$ with $\Gamma(x) \cap O \neq$, then there exists some $\epsilon > 0$ such that $\Gamma(x') \cup O \neq \emptyset$ for all $x' \in N_\epsilon(x) \cap X$.

In the single-valued case, $\Gamma(x) \cup O \neq \emptyset$ just means inclusion. The general idea of Lower Hemi-continuity is that the image of the neighboring points to x must not go "too far away" from $\Gamma(x)$.

As in the case of u.h.c., we can have a sequential characterization for lower hemi-continuity.

Proposition 18. *Γ is lower hemi-continuous at $x \in X$ if and only if, for all $\{x_n\} \subset X$, with $x_n \to x$, and $y \in \Gamma(x)$, then, there exists $y_n \in Y$, such that $y_n \to y$ and $y_n \in \Gamma(x_n)$ for each n.*

A graphical example:



(a) not l.h.c. correspondence (b) l.h.c. correspondence

In the second figure, we can construct a sequence converging toward x_0. Take for instance $y_n = y_0$, and $y \in \Gamma(x_0)$, then $y = y_n$.

Given the definition of u.h.c. and l.h.c., we can now define the continuity for correspondences.

Definition 15.1.6. A correspondence $\Gamma : X \rightrightarrows Y$ is **continuous** at x if it is both upper hemi-continuous and lower hemi-continuous.

Let's see a further example.

Example 15.1.3. *Recall that we have defined the budget set as a correspondence.*

$$B_{p,w} = \Big\{ x \in \mathbb{R}^n : p \cdot x \leq w \Big\}$$

With $p \gg 0, w > 0$. Now, we want to show that the budget set is continuous. the first step is to show that $B_{p,w}$ is upper hemi-continuous. We use sequential characterization. We want to show that for any $(p_n, w_n) \in B_{p,w} \to (p, w)$ and $x_n \in B(p_n, w_n)$ for each n. Then, there exists a convergent subsequence $\{x_{n_k}\} \to x$, where $x \in B_{p_n,w_n}$. Let $\underline{p}_i = \inf\{p_n\}$ and $\bar{w} = \sup\{w_n\}$. Then $B_{p_n,w_n} \subset B_{\underline{p}_i,\bar{w}}$ and $\{x_n\} \in B_{\underline{p}_i,\bar{w}}$ so there must be a convergent subsequence $\{x_{n_k}\} \to x$. Further, $p_{n_k} \cdot x_{n_k} \leq w_{n_k}$, $p_{n_k}, w_{n_k} \to p, w$ and $x_{n_k} \to x$. Then $p \cdot x \leq w$, since the inequality is preserved in the limit, and $x \in B_{p,w}$. So $B_{p,w}$ is upper hemi-continuous.

Let's see lower hemi-continuity now. We want to show that, for all $(p_n, w_n) \to (p, w)$ and $x \in B_{p,w}$, there exists $x_n \in B_{p_n,w_n}$ such that $x_n \to x$. If $x = 0$, let $x_n = 0, \forall n$. If $x \neq 0$, define:

$$x_n = \frac{w_n}{w} \cdot \frac{p \cdot x}{p_n \cdot x} \cdot x$$

multiplying by $-p_n$, we have:

$$-p_n \cdot x_n = \frac{w_n}{w} \cdot \frac{p \cdot x}{p_n \cdot x} \cdot -p_n x$$

and then $p_n = w_n \cdot \frac{p \cdot x}{w} \leq w_n$. $x_n \in B_{p_n,w_n}, \forall n$, and therefore $x_n \to x$, so $B_{p,w}$ is lower hemi-continuous. Since $B_{p,w}$ is l.h.c and u.h.c., then it is continuous.

Continuity of correspondences is necessary to prove an important result, which will be useful in establishing key results in Dynamic Programming, other than in economics, namely Berge's **Maximum Theorem**.

Theorem 15.1.1. *Let $\Theta \subset \mathbb{R}^n$ and $X \subseteq \mathbb{R}^n$, and:*

$$\Gamma : \Theta \rightrightarrows X$$

be a compact-valued correspondence, and $\varphi \in C(X \times \Theta)$ (namely, it is a continuous function on $X \times \Theta$). Further, define $\sigma(\theta) = \arg\max_\theta \left\{\varphi(x, \theta) : x \in \Gamma(\theta)\right\}$, $\forall \theta \in \Theta$, and $\varphi^*(\theta) = \max_\theta \left\{\varphi(x, \theta) : x \in \Gamma(\theta)\right\}$, $\forall \theta \in \Theta$.

If Γ is continuous at some $\theta \in \Theta$, then:

1. *$\sigma : \Theta \rightrightarrows X$ is upper hemi-continuous and compact-valued*

2. *$\varphi^* : \Theta \to \mathbb{R}$ is continuous.*

Example 15.1.4. *Let's see the Demand Correspondence. Define $\Theta = \mathbb{R}^n_{++} \times \mathbb{R}_{++}$: $(p, w) \in \Theta$ in other words, Θ is the space of parameters (p, w). X is the commodity space. $\varphi(x, \theta)$ is the utility function $u(x)$, $\Gamma(\theta)$ is the Budget correspondence. We know that $x^*(p, w)$ (the Marshallian demand) be $\arg\max\{u(x) : x \in B_{p,w}\} \equiv \sigma(\theta)$ (in the notation used above), and $v(p, w)$ (the indirect utility function) be $\max\{u(x) : x \in B_{p,w}\} \equiv \varphi^*(\theta)$. Further, we have seen that $B_{p,w}$ is continuous. If $u(x)$ is also continuous, then, from the maximum theorem, $x^*(p, w)$ is upper hemi-continuous and compact-valued, $v(p, w)$ is continuous. If $x^*(p, w)$ is a function, then it is continuous, and if $u(x)$ is strictly quasi-concave, then $x^*(p, w)$ is a function.*

We can finally present some Fixed Point Theorems. The first is the Contraction Mapping Theorem, also known as the Banach Fixed Point Theorem. Let's start with a definition of contraction mapping.

Definition 15.1.7. A function $f : S \to S$ is a contraction mapping if $d(f(x), f(y)) < \beta d(x, y)$ for some fixed $\beta \in [0, 1)$, for all $x, y \in S$.

Example 15.1.5. *Let's take a function $f(x) = \frac{1}{2}x$. This is a contraction mapping, as apparent from the figure 15.6, with $\beta = \frac{1}{2}$.*

Let's see the following result.

Theorem 15.1.2 (Banach Fixed Point Theorem)**.** *Let $S \subseteq \mathbb{R}^n$ be closed and $f : S \to S$ be a contraction mapping. Then there exists a unique fixed point $f(x) = x$*

Proof. Choose any $x_0 \in S$. Let $x_n = f(x_{n-1})$. If $x_n \to x^*$, then $x^+ \in S$ and $x^* = f(x^*)$. This is because S is closed, so it contains the limit. Furthermore, x^* is a fixed point because $f(\cdot)$ is continuity. But how do we know that $f(\cdot)$ is continuous? A contraction is always continuous. To see this take a sequence $|f(x) - f(y)| \leq \theta|x - y|$. If $|x - y|$ is very small, also $|f(x) - f(y)|$ becomes very small. Take $k \in \mathbb{R}_+$. This still implies continuity

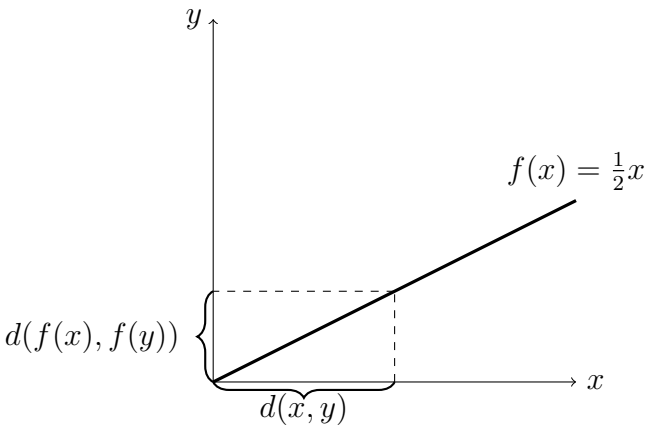


Figure 15.6: A contraction mapping

(Lipschitz continuity). We want to show that x_n converges. To do so, we must show that $\{x_n\}$ is a Cauchy Sequence. Then, since any Cauchy Sequence converges, $\{x_n\}$ converges too. By definition of Cauchy Sequence, take $m, n > M, \epsilon > 0$, such that $|x_m - x_n| < \epsilon$. Suppose $m > n$. Then:

$$|x_m - x_n| \leq \sum_{k=1}^{m} -1|x_{k+1} - x_k|$$

Notice that

$$\begin{aligned} &|x_{k+1} - x_k| = |f(x_k) - f(x_{k-1})| \leq \theta|x_n - x_{n-1}| \\ &\cdots \leq \theta^n|x_1 - x_0| \end{aligned} \tag{15.1}$$

Therefore $\theta^n|x_1 - x_0|$ is a bound for $|x_m - x_n|$, so:

$$|x_m - x_n| \leq \sum_{k=1}^{n-1} |x_{k+1} - x_k| \leq |x_1 - x_0|\frac{\theta^n}{1-\theta}$$

So $\{x_n\}$ is a Cauchy Sequence, and then it is a fixed point in the limit. Since the limit is unique, then also the fixed point is unique. □

To get the main intuition behind this result (in $\mathbb{R}^2$), see the figure 15.7.

Start with x_0. Then $f(x_0) = x_1$. Then, from $f(x_1)$, we obtain x_2. From x_2, we have $f(x_2) = x_3$. And so on, until we reach the fixed point $f(x^*) = x^*$ (follow the red arrow in the graph). In this example, we have assumed $x_0 > x^*$. The same intuition for $x_0 < x^*$. In other words, this result shows that we can find a fixed point by doing iteration $x_0, f(x_0), f(f(x_0))$ and so on...

A suggestive example of Contraction Mapping is that of imagining a geographical map of North Carolina. Suppose you can display the map on the floor of the Garner Hall. Then there is a point on the map (no matter how infinitesimal), which also corresponds to the point where the map has been actually displayed.

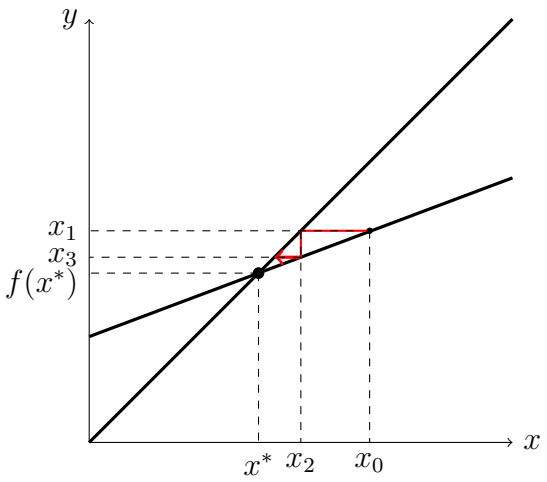


Figure 15.7: A contraction mapping and a fixed point

Two other fixed point results used by economists do not involve contraction mappings but just continuity, compactness, and convexity of the domain. These are the Brouwer Fixed Point Theorem and the Kakutani Fixed Point Theorem.

Theorem 15.1.3 (Brouwer Fixed Point Theorem)**.** *Let $S \subseteq \mathbb{R}^n$ be compact, convex and $f : S \to S$ be a continuous function. Then there exists a fixed point $x^* \in S$ such that $f(x^*) = x^*$.*

Notice that in this case, contrary to the Contraction Mapping, the fixed point is not unique. The trivial example is given by the function $f(x) = x$, which is entirely made up of infinite fixed points.

Let's see the following examples when the theorem fails.

Example 15.1.6. *See $f(x) = \frac{x}{2}$, where $f : (0,1) \to (0,1)$. Notice that this function maps its domain into itself. However, a fixed point does not exist because $(0,1)$ is not compact (it is not closed).*

Take another example, the function $f : [0,1] \to [0,1]$:

$$f(x) = \begin{cases} 1 & \text{if } x = 0 \\ \frac{x}{2} & \text{if } x \in (0,1) \\ 0 & \text{if } x = 1 \end{cases}$$

The domain of this function is compact. However, the function is not continuous, so a fixed point does not exist.

This theorem can be used to prove a famous result in Linear Algebra, namely the **Perron-Frobenius theorem**.

Theorem 15.1.4. *Any matrix $A \gg 0$ has a positive eigenvalue and eigenvector.*

Proof. Define the simplex:

$$S = \{x \in \mathbb{R}^n_+ : \sum_i x_i = 1\} \equiv \Delta$$

and the function $f : S \to S$ as:

$$f(x) = \frac{Ax}{\sum_i Ax}$$

f is continuos and S is compact. So that it exists a fixed point x^* such that:

$$f(x^*) = \frac{Ax^*}{\sum_i (Ax^*}$$

Rearranging, we can write:

$$Ax^* = \sum_i (Ax^*)x^*$$

Notice that $\sum_i (A(x^*))$ is a number, so we can define it as λ. Therefore $Ax^* = \lambda x^*$, and this is the characteristic equation, where x^* is an eigenvector and λ is the eigenvalue. Furthermore, both x^* and λ are positive. □

This result should not be a surprise since an eigenvector associated with an eigenvalue $\lambda = 1$ is a fixed point. indeed:

$$Ax = \underbrace{\lambda}_{=1} x$$

The third fixed point theorem we present is the Kakutani's Theorem.

Theorem 15.1.5 (Kakutani Fixed Point Theorem)**.** *Let $S \subseteq \mathbb{R}^n$ be a compact and convex set. Let $\Gamma : S \rightrightarrows S$ be an upper hemi-continuous and convex-valued correspondence. Then there exists $X \in S$ such that $x^* \in \Gamma(x^*)$*

With these results, we can now formally treat Dynamic Programming.

15.2 Dynamic Programming

A problem of programming along an infinite temporal horizon can be written as follows:

$$\max_{\{x_n\}_n^\infty} \varphi(x_0, x_1) + \sum_{i=1}^{\infty} \delta^i \varphi(x_i, x_{i+1})$$

s.t.

$$x_{i+1} \in \Gamma(x) \quad i = 0, 1, \ldots$$

Where x_0 is the initial state, and $x_i \in X$, X is the state space, most often assumed to be $X = \mathbb{R}$. $\Gamma(x)$ is the **transition correspondence** that describes which states are possible tomorrow, given the state of the system today.

A **feasible plan** is a sequence of states $\{x_i\} \subset X$ such that $x_1 \in \Gamma(x_0), x_2 \in \Gamma(x_1), \ldots$. Instead, $\delta \in (0,1)$ is a discount factor.

The main difference with non-linear programming is that we are operating in an infinite horizon. Namely, we must choose an infinite dimension vector $\{x_n\}_{n=1}^{\infty}$ that is made up of states that maximize the objective function at each period.

There are several assumptions regarding the structure of the problem.

Definition 15.2.1. The main assumptions are:

A.1 For any feasible plan $\{x_i\}$, we have:

$$\lim_{t\to\infty} \sum_{i=1}^{k} \varphi(x_i, x_{i+1}) \in \bar{\mathbb{R}}$$

A.2 φ is continuous and bounded

A.3 Γ is compact-valued and continuous

Assumption 1 states that for any feasible plan, the solution of the problem either diverges or converges. It rules out oscillations.

We can further define the total objective function as:

$$F(\{x_i\}, x) = \varphi(x, x_i) + \sum_{i=1}^{\infty} \delta^i \varphi(x_i, x_{i+1})$$

and the set of feasible plans as:

$$\Omega(x) = \left\{ (x_i) \in X^{\infty} : x_i \in \Gamma(x) \text{ and } x_{i+1} \in \Gamma(x_i), i = 1, 2, \ldots \right\}$$

The value function can be written as:

$$V(x) = \sup \left\{ F(\{x_i\}_{i=1}^{\infty}, x) : \{x_i\} \in \Omega(x) \right\}$$

Since we are in an infinite horizon, we are interested in knowing if a solution exists. To tackle this problem, we use an iterative approach. The problem is transformed into a **recursive problem**. Recursive means that we want to summarize the future as a value function of tomorrow, and therefore the infinite horizon problem is transformed into a two-period problem: today and "tomorrow" (the future). We can write:

$$V(x_i) = x_i + V(x_{i+1}, x_i)$$

We have the following lemma:

Lemma 15.2.1. *Given x_0 (the initial value), $V(x_0) = F(\{x_n^*\}, x_0)$ (namely $\{x_n^*\}$ solves the optimization problem) if and only if:*

$$V(x_0) = \varphi(x_0, x_i^*) + \delta V(x_1^*)$$

and

$$V(x_n^*) = \varphi(x_n^*, x_{n+1}^*) + \delta V(x_{n+1}^*) \ \ \forall n = 1, 2, \dots$$

Proof. ($\Leftarrow$) If $\{x_n\}_{n=1}^{\infty} \Rightarrow$ recursive optimality, then we can write:

$$\begin{aligned} V(x_0) =& F(\{x_n\}_{n=1}^{\infty}, x_0) = \\ & \varphi(x_0, x_1^*) + \sum_{i=1}^{\infty} \delta^i \varphi(x_i^*, x_{i+1}^*) \geq \\ & \varphi(x_0, x_1) + \sum_{i=1}^{\infty} \delta^i \varphi(x_i, x_{i+1}) \ \ \forall \{x_i\} \in \Omega(x) \end{aligned}$$

So, we can take the next period $\{x_2, x_3, \dots\} \in \Omega(x_i^*)$:

$$\begin{aligned} & \varphi(x_0, x_1^*) + \delta \varphi(x_1^*, x_2^*) + \sum_{i=1}^{\infty} \varphi(x_i^*, x_{i+1}^*) \geq \\ & \varphi(x_0, x_1^*) + \delta \varphi(x_1^*, x_2) + \sum_{i=1}^{\infty} \varphi(x_i, x_{i+1}) \end{aligned}$$

We can cancel the first terms at right and left, and we have:

$$\varphi(x_1^*, x_2^*) + \sum_{i=1}^{\infty} \varphi(x_i^*, x_{i+1}^*) \geq \varphi(x_1^*, x_2) + \sum_{i=1}^{\infty} \varphi(x_i, x_{i+1})$$

Then, the left hand-side is $V(x^*)$ and $V(x_0) = \varphi(x_0, x_1^*) + \delta V(x_i^*)$.

($\Rightarrow$) Let's write:

$$\begin{aligned} V(x_0) =& \varphi(x_0, x_i^*) + \delta V(x_i^*) = \\ & \varphi(x_0, x_i^*) + \delta \varphi(x_1^*, x_2) + \delta^2 V(x_2^*) = \\ & \dots \\ & \varphi(x_0, x_i^*) + \sum_{i=1}^{k} \delta^i \varphi(x_i^*, x_{i+1}^*) + \delta^{k+1} V(x_{i+1}^*) \end{aligned}$$

Since the problem is bounded (Assumption 2), then $\delta^k V \to 0$.Therefore:

$$V(x_0) = F(\{x_i^*\}, x_0)$$

□

This lemma says that once you have an optimal solution, this has a recursive structure. You can summarize tomorrow's continuation value using today's optimal value.

Another Lemma is the **principle of optimality**, which tells how to find the value function.

Lemma 15.2.2. *For any function $W \in B(X)$ (where B is the set of bounded functions on $X = \mathbb{R}$), we have:*

$$W(x) = \max\left\{\varphi(x,y) + \delta W(y) : y \in \Gamma(x)\right\} \quad \forall x \in X$$

Then we have:

$$W(x) = \max\left\{F(\{x_i\}, x) : \{x_i\} \in \Omega(x)\right\} \quad \forall x \in X$$

Proof. For all $\{x_i\} \in \Omega(x)$, we have:

$$\begin{aligned} W(x) &\geq \varphi(x,x_1) + \delta W(x_1) \geq \\ &\quad \varphi(x,x_1) + \delta \underbrace{\left[\varphi(x_1,x_2) + \delta^2 W(x_2)\right]}_{W(x_1)} \geq \\ &\quad \cdots \\ &\quad \varphi(x,x_1) + \sum_{i=1}^{k} \delta^i \varphi(x_i, x_{i+1}) + \delta^{k+1} W(x_{i+1}) \end{aligned}$$

Since W is bounded, as $k \to \infty$, the last term of the right hand-side disappears. So we have:

$$W(x) = F(\{x_i^*\}, x)$$

□

$W(x)$ is also called **Bellman Equation**.

The next question to be tackled is that of the existence of a solution.

Theorem 15.2.3 (Existence of a solution)**.** *Under assumptions $A1 - A3$, there exists a solution to the dynamic programming problem*

Proof. The basic idea of the proof is that if we find a value function, then the optimal solution exists. Define a mapping:

$$\Phi : CB(x) \to \mathbb{R}^x$$

where $CB(x)$ is the set of all continuous and bounded functions.

$$\Phi(w)(x) = \max\left\{\varphi(x,y) + \delta W(y) : y \in \Gamma(x)\right\}$$

For this function, $\Phi(w)(x_i), \varphi(\cdot), W(\cdot)$ are continuous ($\Phi(\cdot)$ by Maximum's theorem). Besides, it is also bounded so that:

$$\Phi(w) : CB(x) \to CB(x)$$

If exists a w^* such that $\Phi(w^*) = w^*$, then w^* is the value function. So, we want to show that $\Phi(\cdot)$ has a fixed point. We use the Banach Fixed Point Theorem (in a more general version, on any metric space, not just Euclidean). But before, we need to show that $\Phi(\cdot)$ is a contraction. Therefore we need the following lemma.

Lemma 15.2.4 (Blackwell's sufficient conditions for a contraction)**.** *Suppose* $\Phi(cdot)$ *is a mapping into itself. Then, if:*

1. $\Phi(w) \geq \Phi(w')$*, for all* $w \geq w'$ *(monotonicity)*
2. *it exists a delta* $\in (0,1)$ *such that* $\Phi(w + \alpha) \leq \Phi(w) + \delta\alpha$ *(discounting)*

Then, $\Phi(\cdot)$ *is a contraction.*

$\Phi(\cdot)$ is monotonic in w, and can be bounded by a δ, so $\Phi(\cdot)$ is a contraction. Then, by the Banach Fixed Point Theorem, it exists a w^* such that $\Phi(w^*) = w^* \in CB(x)$.

The problem has a solution. □

So far, we have shown:

- the structure of optimal solution (namely that they are recursive), and optimal values (the Bellman equation)
- the existence of an optimal solution

We need to know how to find a solution. That is, to find an optimal policy. Given the existence of a solution, we can construct an optimal solution correspondence:

$$P(x) = \arg\max_y \Big\{ \varphi(x,y) + \delta V(y) : y \in \Gamma(x) \Big\}$$

we know that $\{x_i^*\}$ must satisfy:

$$x_1^* \in P(x_0^*), \quad x_2^* \in P(x_1^*) \quad \ldots \text{ and so on...}$$

If we know the value function, we can find a solution. Indeed, the value function is smooth, so we can take the derivative.

We also want to determine if the solution is unique. Let's make a further assumption ($A4$), namely that $Gr(\Gamma)$ is convex and $\varphi(\cdot)$ is strictly concave on $Gr(\Gamma)$. Then, we have the following theorem:

Theorem 15.2.5. *(Uniqueness) Let* $X \subseteq \mathbb{R}^n$ *be a non-empty and convex set. Under assumptions* $A1 - A4$*, the dynamic programming problem has a unique solution.*