Chapter 1

Typology of $\mathbb{R}^n$ and Fixed Point Theorems

Typology means the "properties of subsets".

1.1 Fundamental concepts

Sets

- $\mathbb{N}$: natural numbers (1, 2, 3, ...)
- $\mathbb{Z}$: integers (..., -2, -1, 0, 1, 2, ...)
- $\mathbb{Q}$: rational numbers (fractions of integers), same as $\mathbb{Z} \cup \{\text{fractionary numbers}\}$
- $\mathbb{R}$: real numbers (all rational and irrational numbers), same as $\mathbb{Q} \cup \{\text{irrational numbers}\}$
- $\mathbb{C}$: complex numbers (numbers with real and imaginary parts), $\{a + bi, a, b \in \mathbb{R}\}$

Symbols

- $\forall$: for all
- $\exists$: there exists

Terminology

- $\in$: belongs to
- $\subset$: is contained
- $\supset$: contains

A, B, 2 subsets of $\mathbb{R}$

- $A \cup B = \{x \in \mathbb{R} : x \in A \text{ or } x \in B\}$: union of A and B, all elements in A or B
- $A \cap B = \{x \in \mathbb{R} : x \in A \text{ and } x \in B\}$: intersection of A and B, all elements in both A and B
- $A^c = \{x \in \mathbb{R} : x \notin A\}$: complement of A, all elements not in A
- $A \setminus B = \{x \in \mathbb{R} : x \in A \text{ and } x \notin B\}$: difference of A and B, all elements in A but not in B
- $\emptyset$: empty set, set with no elements

Definition 1.1.1: Disjoint

$A, B \subset \mathbb{R}$, A and B are disjoint iff $A \cap B = \emptyset$.

Properties

- Commutative: $A \cup B = B \cup A$, $A \cap B = B \cap A$
- Associative: $A \cup (B \cup C) = (A \cup B) \cup C$, $A \cap (B \cap C) = (A \cap B) \cap C$
- Distributive: $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$, $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
- De Morgan's Laws: $(A \cup B)^c = A^c \cap B^c$, $(A \cap B)^c = A^c \cup B^c$

Terminology

I is a set of indixes, $\forall_j \in I, A_j \subset \mathbb{R}$

- $\bigcap_{j \in I} A_j = \{x \in \mathbb{R} : x \in A_j, \forall_j \in I\}$: intersection of all A_j, all elements in every A_j
- $\bigcup_{j \in I} A_j = \{x \in \mathbb{R} : x \in A_j \text{ for some } j \in I\}$: union of all A_j, all elements in at least one A_j

Example 1.1.1 (Union and Intersection)
$I_n = (-\frac{1}{n}, \frac{1}{n}), n \in \mathbb{N}$
$\bigcap_{j \in \mathbb{N}} I_n = \{0\}$
$\bigcup_{j \in \mathbb{N}} I_n = (-1, 1)$

From now on, extending the domain where the sets live to $\mathbb{R}^n$. $\mathbb{R}^n = \{(x_1, x_2, \cdots, x_n) : x_i \in \mathbb{R}, i = 1, 2, \cdots, n\}$. $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R} = \{(a, b) : a \in \mathbb{R}, b \in \mathbb{R}\}$. $\mathbb{R}^n = \mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R} = (x_1, x_2, \cdots, x_n)$.

Example 1.1.2 ($\mathbb{R}^2$)
$\mathbb{R} \times \{0\} = \{(x, 0), x \in \mathbb{R}\}$, the order of the pair matters.

1.2 Functions

A, B sets. Correspondence between two sets, A and B, in such a way that to each element of A corresponds one and only one element of B.

Note:
element of A: objects
elements of B that receives arrow: image
$f : \underbrace{A}_{domain} \rightarrow \underbrace{B}_{codomain} \Rightarrow x \rightarrow \underbrace{f(x)}_{analytical\ expression}$

Example 1.2.1 (Domain-1)
$f : D_f \rightarrow \mathbb{R} \Rightarrow x \rightarrow \sqrt{x}$
$D_f = [0, +\infty)$ or $D_f = \mathbb{R}_0^+$, also called the maximal domain

Example 1.2.2 (Domain-2)
$g : D_g \rightarrow \mathbb{R} \Rightarrow x \rightarrow \frac{1}{x}$
$D_g = (-\infty, 0) \cup (0, +\infty)$ or $D_g = \mathbb{R} \setminus \{0\}$

Example 1.2.3 (Domain-3)
$h : D_h \to \mathbb{R} \Rightarrow x \to \ln(x)$
$D_h = (0, +\infty)$ or $D_h = \mathbb{R}^+$

Example 1.2.4 (Domain-4)
$A \subset \mathbb{R}, l : A \to A \Rightarrow x \to x$
$D_l = \mathbb{R}$, also known as the identity map

Definition 1.2.1: Graph

$f : A \to B \Rightarrow x \to f(x)$
Graph of f is defined as $\mathrm{Gr}(f) = \{(x, y) \in A \times B : y = f(x)\}$.

Example 1.2.5 (Graph)
$f : \mathbb{R}^3 \to \mathbb{R} \Rightarrow (x_1, x_2, x_3) \to x_1x_2x_3$
$\mathrm{Gr}(f) = \{(x_1, x_2, x_3, y) \in \mathbb{R}^3 \times \mathbb{R} : y = f(x_1, x_2, x_3) = x_1x_2x_3\}$

$f : A \to B$ where $A, B \subset \mathbb{R}$

Definition 1.2.2: Injective

$f(x_1) = f(x_2) \Leftrightarrow x_1 = x_2$

An injective function, also known as one-to-one function, is a function where distinct elements in the domain map to distinct elements in the codomain. This means that no two different inputs can produce the same output.

Example 1.2.6 (Injective?)
$f : \mathbb{R} \to \mathbb{R} \Rightarrow x \to x^2$
This is not injective because $f(1) = f(-1) = 1$ but $1 \neq -1$.
But, if we restrict the domain to $\mathbb{R}_0^+$, then it is injective.

Definition 1.1: Surjective

$\mathrm{Image}(f) = B$

A surjective function, also known as onto function, is a function where every element in the codomain has at least one element from the domain mapping to it. This means that the function covers the entire codomain.

Example 1.2.7 (Surjective?)
$f : \mathbb{R} \to \mathbb{R} \Rightarrow x \to x^2$
This is not surjective because there is no $x \in \mathbb{R}$ such that $f(x) = -1$.
But, if we restrict the codomain to $\mathbb{R}_0^+$, then it is surjective.

Definition 1.2.3: Bijective

f is bijective iff it is injective and surjective.

Compontion of maps

$A, B, C \subset \mathbb{R}$, and $\begin{cases} f : A \longrightarrow B \\ g : B \longrightarrow C \end{cases}$ Then, the composition of f and g is defined as $g \circ f : A \longrightarrow C \Rightarrow x \longrightarrow g(f(x))$. The map is well defined if $\text{Image}(f) \subset D_g$.

Example 1.2.8 (Composition of maps)

$f : \mathbb{R} \to \mathbb{R} \Rightarrow x \longrightarrow x^2$
$g : \mathbb{R} \to \mathbb{R} \Rightarrow x \longrightarrow x + 1$
$g \circ f : \mathbb{R} \to \mathbb{R} \Rightarrow x \longrightarrow g(f(x)) = x^2 + 1$
$f \circ g : \mathbb{R} \to \mathbb{R} \Rightarrow x \longrightarrow f(g(x)) = (x+1)^2 = x^2 + 2x + 1$

In general, $g \circ f \neq f \circ g$ unless linear.

Definition 1.2.4: Inverse

f, g are maps. If $f \circ g = g \circ f = I_d$, then f and g are inverses of each other, denoted as $f = g^{-1}$ and $g = f^{-1}$, one with respect to other.

Example 1.2.9 (Inverse?)

$g : \mathbb{R} \to \mathbb{R}^+ \Rightarrow x \longrightarrow \exp^x$
$f : \mathbb{R}^+ \to \mathbb{R} \Rightarrow x \longrightarrow \ln(x)$
$f \circ g = f(e^x) = \ln(e^x) = x$
$g \circ f = g(\ln(x)) = e^{\ln(x)} = x$
$f \circ g = g \circ f = I_d \Rightarrow$ f and g are inverses of each other.

Corollary 1.2.1 Invertibility

f is invertible iff f is bijective.
$f : \mathbb{R} \to \mathbb{R}$ is not bijective $\Rightarrow$ f is not invertible.

<u>Cardinal of sets</u>

$\Omega \subseteq \mathbb{R}^n, n \in \mathbb{N}$

Definition 1.2: Finite

Ω is finite if $\#\Omega \in \mathbb{N}$.
There exists a bijection with Ω and $\{1, 2, \cdots, n\}$ for some $n \in \mathbb{N}$.

Example 1.2.10 (Cardinality)

$A = \{a, b, c\}$ where $a \neq b \wedge b \neq c \wedge c \neq a$
$\#A = 3$

And of course, Ω is infinite if it is not finite.

Note:

$\#\emptyset = 0$

Infinite sets can be further classified into countable and uncountable sets.

- Ω is countable if there exists a bijection between Ω and $\mathbb{N}$.
- Ω is uncountable if it is not countable.

Note:
In this course, finite sets are countable sets.

Example 1.2.11 (Countable)
$A = \{2, 4, 8\}$ is finite $\Rightarrow$ A is countable.

Example 1.2.12 (Countable)
$B = \{4, 5, 6, 7, \cdots\} = \{n \in \mathbb{N} : x \geq 4\}$
$f : B \to \mathbb{N} \Rightarrow n \to n - 3$ is a bijection $\Rightarrow B$ is countable.

Example 1.2.13 (Countable)
$\mathbb{Z} = \{\cdots, -3, -2, -1, 0, 1, 2, 3, \cdots\}$
$f : \mathbb{N} \to \mathbb{Z}$

$$n \to \begin{cases} 0, & n = 1 \\ k, & n = 2k \\ -k, & n = 2k + 1 \end{cases}, k \in \mathbb{N}, \text{ is a bijection} \Rightarrow \mathbb{Z} \text{ is countable.}$$

Example 1.2.14 (Uncountable)
$[0, 1]$ is uncountable.
Proof by contradiction: assume $[0, 1]$ is countable. Then, there exists a bijection $f : \mathbb{N} \to [0, 1]$. Let $f(n) = 0.a_{n1}a_{n2}a_{n3}\cdots$ be the decimal representation of $f(n)$. Construct a number $b = 0.b_1b_2b_3\cdots$ where $b_n \neq a_{nn}$ and $b_n \in \{0, 1, 2, \cdots, 9\}$. Then, $b \in [0, 1]$ but $b \neq f(n)$ for all $n \in \mathbb{N}$, contradicting the assumption that f is a bijection. Therefore, $[0, 1]$ is uncountable.

Corollary 1.2.2 Countability
$\mathbb{R}$ is uncountable.

1.3 Metric spaces

A metric space is a set equipped with a metric, which is a function that defines a distance between any two elements in the set. $\Omega \neq \emptyset$.

Definition 1.3.1: Metric

A metric or distance is a map $d : X \times X \to \mathbb{R}_0^+$ that satisfies the following properties

- (Non-negativity) $d(x, y) \geq 0$
- (Identity of indiscernibles) $d(x, y) = 0 \Leftrightarrow x = y$
- (Symmetry) $d(x, y) = d(y, x)$
- (Triangle inequality) $d(x, z) \leq d(x, y) + d(y, z), \forall x, y, z \in \Omega$

Example 1.3.1 (Metric)
$\Omega = \mathbb{R}$
$d(x, y) = |x - y|$ is a metric.

Example 1.3.2 (Metric)
$\Omega = \mathbb{R}^2$
$d(x, y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}$ is a metric.

Example 1.3.3 (Metric)
$\Omega = \mathbb{R}^n$
$d(x, y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \cdots + (x_n - y_n)^2} = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}$ is a metric.

The metric defined above is called the Euclidean metric or Euclidean distance. There are also other metrics such as the Manhattan metric and the distance along the surface.

Example 1.3.4 (Manhattan distance)
$A \longrightarrow (x_A, y_A), B \longrightarrow (x_B, y_B)$
$d(A, B) = |x_A - x_B| + |y_A - y_B|$

Definition 1.3.2: Bounded map

(X, d) is a metric space, $A \subset X$.
$f : A \longrightarrow \mathbb{R}$ is bounded if there exists $a, b \in \mathbb{R}$ such that $a \leq f(x) \leq b, \forall x \in A$.
$f : A \longrightarrow X$ is bounded iff $\exists a \in X, \forall x \in A, d(f(x), a) \leq M$.

Example 1.3.5 (Unbounded)
$f(x, y) = e^{x^2+y^2}, (x, y) \in \mathbb{R}$
$f : \mathbb{R}^2 \longrightarrow \mathbb{R}$
f is unbounded because $\lim\limits_{x^2+y^2 \to \infty} e^{x^2+y^2} = \infty$.

Example 1.3.6 (Bounded)
$g(x, y) = e^{-x^2-y^2}, (x, y) \in \mathbb{R}$
$g : \mathbb{R}^2 \longrightarrow \mathbb{R}$
$-x^2 - y^2 \in \mathbb{R}_0^-$
$0 < e^{-x^2-y^2} \leq 1$
$\therefore$ g is bounded.

Definition 1.3.3: Diameter

$A, B \subset X$ where (X, d) is a metric space.
The diameter is defined as $\operatorname{diam}(A, B) = \sup\{d(x, y), x \in A, y \in B\}$ which is the maximum distance between A and B.

Example 1.3.7 (Diameter)
$X = \mathbb{R}^2$
$A = \{(x, 0), x \in \mathbb{R}\}$
$B = \{(x, y) \in \mathbb{R}^2 : (x - 4)^2 + (y - 4)^2 \leq 1\}$
$\operatorname{diam}(A, B) = \infty$ because A is unbounded.

Definition 1.3.4: Bounded set

$A \subset X$, A is bounded if $\text{diam}(x, y) \leq M$, for some $M \in \mathbb{R}_0^+$, and for all $x, y \in A$.

Note:
f bounded $\neq$ A bounded. f being the map and A being the set.

Example 1.3.8 (Bounded?)
$A = \mathbb{R}^2$
$f : \mathbb{R}^2 \to \mathbb{R} \Rightarrow (x, y) \to e^{-x^2 - y^2}$
f is bounded, i.e. bounded range, but A is unbounded, i.e. unbounded domain.

Example 1.3.9 (Bounded?)
$A = [-1, 1] \times [2, 3]$
$B = \mathbb{R}^2$
$f : A \to B \Rightarrow (x, y) \to (2x, 2y)$
f is bounded, i.e. bounded range, and A is bounded, i.e. bounded domain.

Example 1.3.10 (Bounded?)
$A = (0, 1)$
$f(x) = \ln(x)$
f is unbounded, i.e. unbounded range, but A is bounded, i.e. bounded domain.

1.4 Definition and typology

$X = \mathbb{R}^n, A \subset X, a \in A, r > 0$

Definition 1.4.1: Open ball

An open ball centered at $a \in X$ and radius $r > 0$ is defined as
$B_r(a) = \{x \in X : d(a, x) < r\}$.

Definition 1.4.2: Closed ball

A closed ball centered at $a \in X$ and radius $r > 0$ is defined as
$D_r(a) = \{x \in X : d(a, x) \leq r\}$. D is for "disk". $D_r(a) = \overline{B_r(a)}$.

Definition 1.4.3: Open set

A is an open set of X iff $\forall x \in X, \exists r > 0 : B_r(x) \subset A$.

Definition 1.4.4: Interior point

X_0 is an interior points of A if there exists $r > 0$ such that $B_r(x_0) \subset A$.
We write int$A$.

Definition 1.4.5: Adherent point

X_0 is an adherent point if for all $r > 0$, $B_r(x_0) \cap A \neq \emptyset$.
Same notation as closure: $\overline{A}$ or $\text{cl}(A)$. Can also be written as $\text{ad}(A)$.

Definition 1.4.6: Frontier point

X_0 is a boundary or frontier point of A if for any $r > 0$, $B_r(x_0) \cap A \neq \emptyset$ and $B_r(x_0) \cap A^c \neq \emptyset$.
We write $\text{fr}(A)$ or ∂A.

Definition 1.4.7: Closed set

$A \subset X$ is a closed set iff A^c is open, $X \rightarrow \mathbb{R}^n$.

Example 1.4.1

$A = [0, 1] \subset \mathbb{R}$
$\text{int}(A) = (0, 1)$
$\text{cl}(A) = [0, 1]$
$\text{fr}(A) = \{0, 1\}$
A is closed but not open becasuse $A^c = (-\infty, 0) \cup (1, +\infty)$ which is open.

Definition 1.4.8: Exterior point

$B \subset \mathbb{R}^n$
y is an exterior point of B iff $y \in \text{int}(B^c)$.

Example 1.4.2

$A = [0, 1)$
$A^c = (-\infty, 0) \cup [1, +\infty)$
$\text{int}(A^c) = (-\infty, 0) \cup (1, +\infty)$
$\text{ext}(A) = (-\infty, 0) \cup (1, +\infty)$

Example 1.4.3

$B = [0, 1) \times [0, 1)$
$B^c = [(-\infty, 0) \times \mathbb{R}] \cup [[0, \infty) \times \mathbb{R}^-] \cup [[1, \infty) \times \mathbb{R}] \cup [[0, 1] \times [1, \infty)]$

Proposition: $A \subset \mathbb{R}^n$

1. Any open ball is an open set.
2. $\text{int}(A) \subset A \subset \text{cl}(A)$
3. A is open ⇔ A = int(A)
4. A is closed ⇔ A = cl(A)
5. $\text{cl}(A) = \text{int}(A) \cup \partial A$
6. ∂A is closed.
7. A countable union of open sets is open.
8. A countable intersection of closed sets is closed.

Example 1.4.4

$A = [0, 1) \cup [2, 3]$

- $\text{int}(A) = (0, 1) \cup (2, 3) \longrightarrow$ interior is always open
- $\text{cl}(A) = [0, 1] \cup [2, 3] \longrightarrow$ closure is the interior and the walls

- $\partial A = \{0, 1, 2, 3\}$
- A is open? No. Because $A \neq \text{int}(A)$
- A is closed? No. Because $A \neq \text{cl}(A)$

Remarks:

1. There are sets that are neither open nor closed.
2. $\mathbb{R}^n, \emptyset$ are either open or closed.

Example 1.4.5
$B = \{(x, y) \in \mathbb{R}^2 : x^2 + (y + 2)^2 \leq 4\} \backslash \{(0, 0)\}$

- $\text{int}(B) = \{(x, y) \in \mathbb{R}^2 : x^2 + (y + 2)^2 < 4\}$
- $\text{ext}(B) = \{(x, y) \in \mathbb{R}^2 : x^2 + (y + 2)^2 > 4\}$
- $\text{cl}(B) = \{(x, y) \in \mathbb{R}^2 : x^2 + (y + 2)^2 \leq 4\}$
- $\partial B = \{(x, y) \in \mathbb{R}^2 : x^2 + (y + 2)^2 = 4\}$
- B is open? No. Because $B \neq \text{int}(B)$
- B is closed? No. Because $B \neq \text{cl}(B)$

Definition 1.4.9: Bounded

$A \subset \mathbb{R}^n$
A is bounded if $A \subset B_r(x_0)$, for some $r > 0$ and some $x_0 \in \mathbb{R}^n$.

Example 1.4.6
$\mathbb{R} \times \{0\} \subset \mathbb{R}^n$
Not bounded because we cannot find a x large enough to contain the entire set.

Definition 1.4.10: Compact

$A \subset \mathbb{R}^n$
A is compact iff A is closed and bounded.

Definition 1.4.11: Association point

$A \subset \mathbb{R}^n$
x_0 is an association point of A if for any $r > 0$, $B_r(x_0) \cap [A \backslash \{x_0\}] \neq \emptyset$

Example 1.4.7
$A = \{\frac{1}{n}, n \in \mathbb{N}\} \subset \mathbb{R}$
$0 \in A$? No. But there are points from A that accumulates in the ball of 0.
$\therefore$ 0 is an accumulation point of A.

Definition 1.4.12: Isolated point

$A \subset \mathbb{R}^n$
x_0 is an isolated point of A if $x_0 \in A$ and x_0 is not an accumulation point of A.

Proposition:
$\text{cl}(A) = \text{int}(A) \cup \partial A = \{\text{accumulation points}\} \cup \{\text{isolated points}\}$, accumulation points can also be written as derivative of A, A'.

Remarks:
$x^2 = 9 \Leftrightarrow x = -3 \vee x = 3$
$\{x \in \mathbb{R} : x^2 = 9\} \longrightarrow$ two points
$\{(x, y) \in \mathbb{R}^2 : x^2 = 9\} \longrightarrow$ two lines
$\{(x, y, z) \in \mathbb{R}^3 : x^2 = 9\} \longrightarrow$ two planes

Definition 1.4.13: Neighborhood

$\mathbb{R}^n, x_0 \in \mathbb{R}^n$
v is a neighborhood of x_0, if there exists $r > 0$ such that $B_r(x_0) \subset v$.
In general, we consider open neighborhood.

Remarks:

1. $g = \frac{1}{f}, D_g = \{x \in \mathbb{R}^n : f(x) \neq 0\}$
2. $g = \log(f), D_g = \{x \in \mathbb{R}^n : f(x) > 0\}$
3. $g = \sqrt{f}, D_g = \{x \in \mathbb{R}^n : f(x) \geq 0\}$

Definition 1.4.14: Convex

$A \subset \mathbb{R}^n$
A is convex if for any two points X and Y in A, then the segment $[X, Y]$ is contained in A.

1.5 Sequences in $\mathbb{R}^n$

Map from a subset of $\mathbb{N}$ into $\mathbb{R}^n$.

Example 1.5.1 (Sequence)
$f(n) = n^2, n \in \mathbb{N}$
$f : \mathbb{N} \longrightarrow \mathbb{R} \Rightarrow n \longrightarrow n^2$
$u_n = f(n) = n^2 \longrightarrow$ general term

Example 1.5.2 (Sequence)
$V_n = (n^2, n^2 + 1)$ is a sequence in $\mathbb{R}^2$.
$V_5 = (25, 26)$

Sometimes, it is important to see the convergence of sequences.
$u_n = \frac{1}{n}$
$\lim_{n\to\infty} u_n = 0 \Rightarrow u_n \longrightarrow 0$

Lemma 1.5.1
If a sequnece u_n in $\mathbb{R}$ converges, the limit is unique.

$u_n = (a_{1n}, a_{2n}, \cdots, a_{kn}) \in \mathbb{R}^k \longrightarrow$ the sequence converges if each component converges.

Example 1.5.3 (Convergence)
$u_n = (\frac{1}{n}, \frac{n^2-2}{n^2})$ is a sequence in $\mathbb{R}^2$

- $\frac{1}{n} \longrightarrow 0$
- $\frac{n^2-2}{n^2} = 1 - \frac{2}{n^2} \longrightarrow 1$

$\therefore u_n \longrightarrow (0, 1)$, u_n converges.

When one of the components does not converge, we say that the sequence diverges.

Example 1.5.4 (Divergence)
$u_n = (n^3 + n; \sqrt{n}; \frac{\sqrt{n}+3}{\sqrt{n}})$ in $\mathbb{R}^3$

- $\lim_{n\to\infty} n^3 + n = \infty$ (diverges)
- $\lim_{n\to\infty} \sqrt{n} = \infty$ (diverges)
- $\lim_{n\to\infty} \frac{\sqrt{n}+3}{\sqrt{n}} = \lim_{n\to\infty} 1 + \frac{3}{\sqrt{n}} = 1$

$\therefore u_n$ diverges.

Remarks:

- $(k \in \mathbb{R})$ $\frac{k}{n} \longrightarrow 0$
- $(k \in \mathbb{R})$ $\frac{p(n)}{q(n)}$, then
  1. $\pm\infty$ if $deg(p) > deg(q)$
  2. quotient of the coefficient associated to the highest degree if $deg(p) = deg(q)$
  3. 0 if $deg(p) < deg(q)$

Example 1.5.5 (Sequences and subsequences)
$u_n = (-1)^n$

- $(u_n)_n$ diverges.
- $(u_{2n})_n \longrightarrow 1$
- $(u_{2n+1})_n \longrightarrow -1$, it is a subsequence (infinite sequence of terms of the original sequence).

Proposition:

1. If (u_n) is defined in a compact set, it admits a convergent sequence.
2. The accumulation point of a set A can be seen as a limit of a sequence in cl(A).

Example 1.5.6 (Accumulation point)
$A = \{\frac{1}{n}, n \in \mathbb{R}\}$
$A' = \{0\}$

Example 1.5.7 (Accumulation point)
$u_n = \frac{(-1)^n}{n}$
$A' = \{0\}$

Example 1.5.8 (Accumulation point)
$u_n = (-1)^n$
A' of the sequence: $\{-1, 1\}$

Definition 1.5.1: Cauchy sequence

$(u_n)_n$ is a sequence in $\mathbb{R}$. We say that $(u_n)_n$ is a Cauchy sequence if $\forall \epsilon > 0, \exists n_0 \in \mathbb{N}$: $\forall m, n \geq n_0, d(x_n, x_m) = |u_n - u_m| < \epsilon$.
The difference between two terms is as small as I want.

Informally, $(u_n)_n$ convergent but when the limit is not in the space.

Example 1.5.9 (Cauchy sequence)
$X = (0, 1)$
$u_n = \frac{1}{n}$
$(u_n)_n$ is a Cauchy sequence that is not (or could be) convergent.

Sequence is very close each order but could converge or not.

Proposition: (X,d) metric space
If X is compact, then any Cauchy sequence converges.

Remark:
X is compact $\Rightarrow$ It "includes" all possible limits of subsequences of $(u_n)_n$.

Definition 1.5.2: Complete

(X,d) metric space
We say that X is complete if any Cauchy sequence in X converges.

Remark:

1. X is compact $\Rightarrow$ X is complete.
2. $\mathbb{R}$ is complete, however it is not compact.

1.6 Continuity

$f : D_f \to \mathbb{R}, D_f \subset \mathbb{R}$
a is accumulation point of f. $a \in D_f \to$ f is continuous if and only if $\lim_{x \to a} f(x) = f(a)$.

Definition 1.6.1: Continuity (Heine)

$(X, d_1), (Y, d_2)$ metric spaces

$f : X \to Y, a \in X$
f is continuous $\Leftrightarrow \begin{cases} (x_n)_n \text{ is a squence in X} \\ x_n \to a \end{cases} \Leftrightarrow f(x_n)$ is a sequence in Y and $f(x_n) \to f(a)$.

Example 1.6.1 (Continuity)
$$f : \begin{cases} \mathbb{R}^2 \to \mathbb{R} \\ (x, y) \to e^x + y \end{cases}$$

f is continuous because e^x and y are continuous. It is the sum of continuous maps.

Example 1.6.2 (Continuity)

$$f : \begin{cases} \mathbb{R}\backslash\{0\} \longrightarrow \mathbb{R} \\ x \longrightarrow \frac{1}{x} \end{cases}$$

f is continous because $\frac{1}{x}$ is continuous in its domain.

Example 1.6.3 (Continuity)

$$f : \begin{cases} \mathbb{R}^2\backslash\{(x,0), x \in \mathbb{R}\} \longrightarrow \mathbb{R} \\ (x,y) \longrightarrow \frac{x^2}{y} \end{cases}$$

$(u_n)_n = (\frac{1}{n}, \frac{1}{n}) \longrightarrow (0,0)$
$(v_n)_n = (\frac{1}{n}, \frac{1}{n^2}) \longrightarrow (0,0)$

$$f(u_n) = \frac{(\frac{1}{n})^2}{\frac{1}{n}} = \frac{1}{n} \longrightarrow 0$$
$$f(v_n) = \frac{(\frac{1}{n})^2}{\frac{1}{n^2}} = 1 \longrightarrow 1$$

f is not continuous at (0,0) because the limit depends on the path.

Proposition: $(X, d_1), (Y, d_2)$ metric spaces
$f : X \longrightarrow Y$ is continous. If $k \subset X$ is compact, then $f(k)$ is a compact set of Y.

Theorem 1.6.1 Weierstrass Theorem

$(\mathbb{R}^n, d)$ metric space

$$f : \begin{cases} \mathbb{R}^n \longrightarrow \mathbb{R} \text{ continuous} \\ k \text{ is a compact set} \end{cases}$$
$\Downarrow$
$f(k)$ has a maximum and a minimum. $f(k)$ is closed and bounded.

Theorem 1.6.2 Intermediate Value Theorem

$f : [a, b] \longrightarrow [c, d]$ is continuous
$\text{Im}(f) = [c, d]$

1.7 Fixed point theorems

(X, d) metric spaace.
$f : X \longrightarrow X$

Definition 1.7.1: Fixed point

$x_0 \in X$ is a fixed point if $f(x_0) = x_0$.
$x_0 \in X$ is a K-periodic point if $f^K(x_0) = x_0$ and $f'(x_0) \neq x_0, \forall j = 1, \cdots, k-1$.

Note: $f^K = \underbrace{f \circ f \circ \cdots \circ f}_{K \text{ times}}$

Example 1.7.1 (Fixed point)

$$f : \begin{cases} \mathbb{R} \longrightarrow \mathbb{R} \\ x \longrightarrow x^2 \end{cases}$$

$x_0 = 0$ is a fixed point because $f(0) = 0^2 = 0$.
$x_0 = 1$ is a fixed point because $f(1) = 1^2 = 1$.

Example 1.7.2 (Fixed point)

$$f : \begin{cases} [0,1] \longrightarrow [0,1] \\ x \longrightarrow 1 - x \end{cases}$$

$x_0 = \frac{1}{2}$ is a fixed point because $f(\frac{1}{2}) = 1 - \frac{1}{2} = \frac{1}{2}$.

$f[f(0.25)] = f[0.75] = 0.25 \longrightarrow f^2(0.25) = 0.25$
$f[f(0.1)] = f[0.9] = 0.1 \longrightarrow f^2(0.1) = 0.1$
$\forall x \in [0,1] \backslash \{\frac{1}{2}\}, f^2(x) = x \longrightarrow$ every point except $\frac{1}{2}$ is a 2-periodic point.
$\text{Per}_2(f) = [0,1] \backslash \{\frac{1}{2}\}$

Remark: $\frac{1}{2}$ is not 2-periodic because it is a fixed point.9

Example 1.7.3 (Fixed point)

$$R_\theta : \begin{cases} \mathbb{R}^2 \longrightarrow \mathbb{R}^2 \\ (x,y) \longrightarrow R_\theta(x,y) \\ \theta \in (0, 2\pi) \end{cases}$$

R_θ is a rotation of angle θ around the origin.
R_θ has a unique fixed point which is the origin $\Rightarrow \text{Fix}(R_\theta) = \{(0,0)\}$.

Example 1.7.4 (Fixed point)

$$f : \begin{cases} \mathbb{R} \longrightarrow \mathbb{R} \\ x \longrightarrow \begin{cases} 3x, x < \frac{1}{2} \\ 3 - 3x, x \geq \frac{1}{2} \end{cases} \end{cases}$$

$f(x) = 3x \Rightarrow x = 3x \Rightarrow x = 0$
$f(x) = 3 - 3x \Rightarrow x = 3 - 3x \Rightarrow x = \frac{3}{4}$
0 and $\frac{3}{4}$ are the only fixed points of f $\Rightarrow \text{Fix}(f) = \{0, \frac{3}{4}\}$

Definition 1.7.2: Lipschitz contraction

$$f : \begin{cases} X \longrightarrow X \text{ is a map} \\ (X, d) \text{ metric space} \end{cases}$$

f is a contraction if there exists $0 \leq k < 1$ such that $d(f(x), f(y)) < K \cdot d(x, y), \forall x, y \in X$.

Remark: K is called the Lipschitz constant, it is the ratio of contraction.

Example 1.7.5 (Contraction)

$$f : \begin{cases} \mathbb{R} \longrightarrow \mathbb{R} \\ x \longrightarrow \frac{1}{2}x \end{cases}$$

$|f(x) - f(y)| = |\frac{1}{2}x - \frac{1}{2}y| = \frac{1}{2}|x - y|$
$\therefore f$ is a contraction with $K = \frac{1}{2}$.

In general, to prove that f is a contraction is a difficult task. That is why the next result will be useful in the sequel.

Lemma 1.7.1 Contraction

$f : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is a differentiable map.
If the eigenvalues of the Jacobian matrix $Df(x)$ have modules less than1, then f is a contraction.

Reminder: The Jacobian matrix is the matrix of partial derivatives.

$$Df(x) = \begin{bmatrix} \frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} \\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} \end{bmatrix}$$

Example 1.7.6 (Contraction)

$$f : \begin{cases} \mathbb{R}^2 \longrightarrow \mathbb{R}^2 \\ (x, y) \longrightarrow (-\frac{1}{2}x, \frac{y}{3}) \end{cases}$$

$$Df(x, y) = \begin{bmatrix} -\frac{1}{2} & 0 \\ 0 & \frac{1}{3} \end{bmatrix}$$

Since the Jacobian matrix is diagonal, the eigenvalues are $\lambda_1 = -\frac{1}{2}, \lambda_2 = \frac{1}{3}$.
And since both have modules less than 1.
$\therefore f$ is a contraction.

Note:
Eigenvalues of $A \in M_{n \times n}(\mathbb{R})$ are the roots of the characteristic polynomial $P_A(\lambda) = \det(A - \lambda I)$.

Theorem 1.7.2 Banach Theorem

If

- (X, d) is a complete metric space.
- $f : X \longrightarrow X$ is a contraction.

then f has a **unique** fixed point.

Proof. **Unicity** of the fixed point
Suppose that we have two fixed points $A \neq B$.

Since f is a contraction, then $d(A,B) = d(f(A), f(B)) < K \cdot d(A,B)$ with $0 \leq K < 1$.
$\therefore\ d(A,B) < d(A,B)$ which is a contradiction. ■

Proof. **Existence** of the fixed point
$x_0 \in X$ define the sequence $x_n = f^n(x_0)$
$(x_n)_n$ is a Cauchy sequence $\underbrace{\Rightarrow}_{x \text{ is complete}}$ $(x_n)_n$ converges to the fixed point. ■

Note:
The distance decreases each time you composite with f.

Example 1.7.7 (Banach)

$$f : \begin{cases} \mathbb{R} \to \mathbb{R} \\ x \to x^2 \end{cases}$$

$|f'(x)| = |2x| \not< 1 \Rightarrow$ not a contraction.
$\therefore$ Banach theorem does not apply.

Example 1.7.8 (Banach)

$$f : \begin{cases} [-\frac{1}{3}, \frac{1}{3}] \to \mathbb{R} \\ x \to x^2 \end{cases}$$

$|f'(x)| = |2x| \leq \frac{2}{3} < 1 \Rightarrow$ f is a contraction.
$\therefore$ Banach theorem applies and f has a unique fixed point.

$x = x^2 \Rightarrow x(x-1) = 0 \Rightarrow x = 0, 1$
Since $1 \notin [-\frac{1}{3}, \frac{1}{3}]$, the unique fixed point is 0.

Example 1.7.9 (Banach)

$$f : \begin{cases} B_r(0,0) \to \mathbb{R}^2 \\ x \to \lambda x, \lambda \in (0,1) \end{cases}$$

- $B_r(0,0)$ is complete because it is compact (closed and bounded).
- f is a contraction because $|f'(x)| = |\lambda| < 1$.

$\therefore$ Banach theorem applies and f has a unique fixed point at the origin. $\Rightarrow \mathrm{Fix}(f) = \{(0,0)\}$

Note:
$f : x \to x$ is a contraction of ratio K. Then f is a contraction of ratio $\tilde{K} \in (K, 1)$.

Remark:
$(x_0, y_0) \in \mathbb{R}^2$. If I take any point Q, then $\lim_{n\to\infty} f^n(Q) = P$ where P is the unique fixed point of f. This means that the fixed point is an attractor of the map.

Theorem 1.7.3 Brouwer fixed point theorem
If

- $f : D_f \to D_f$ is a continuous map
- D_f is compact and convex

then, f has a fixed point, but not necessarily unique.

Example 1.7.10 (Brouwer)

$$f : \begin{cases} [0,1] \longrightarrow [0,1] \\ x \longrightarrow x^2 \end{cases}$$

- $[0, 1]$ is compact and convex.
- f is continuous (polynomial).

$\therefore$ Brouwer theorem applies and f has at least one fixed point.

$x = x^2 \Rightarrow x(x-1) = 0 \Rightarrow x = 0, 1$
$\therefore \text{Fix}(f) = \{0, 1\}$

Sidenote: Interval is always convex.

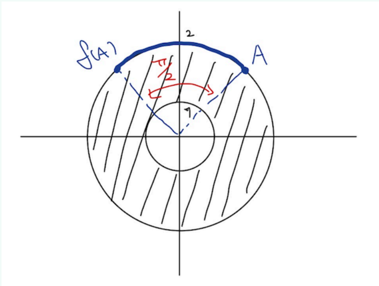
Example 1.7.11 (Brouwer)

$$f : \begin{cases} A \longrightarrow A \\ (x,y) \longrightarrow R_{\frac{\pi}{2}}(x,y) \\ A = \{(x,y) \in \mathbb{R}^2 : 1 \leq x^2 + y^2 \leq 4\} \end{cases}$$

- A is compact but not convex.
- f is continuous (rotation).

$\therefore$ Brouwer theorem does not apply and f has no fixed point.

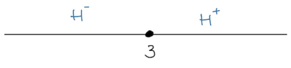
$R_{\frac{\pi}{2}}(x,y) = (x,y)$ has no solution $\Rightarrow \text{Fix}(f) = \emptyset$



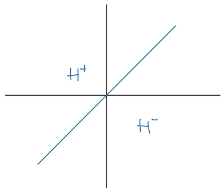
Definition 1.7.3: Hyperplane

Hyperplane in $\mathbb{R}^n$ is a plane of dimension $n-1$ that divides $\mathbb{R}^n$ into semi planes.

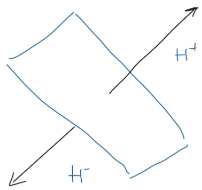
- Hyperplane in $\mathbb{R}^1$ is a point.



- Hyperplane in $\mathbb{R}^2$ is a line that divides the plane into two half-planes.



- Hyperplane in $\mathbb{R}^3$ is a plane that divides the space into two half-spaces.



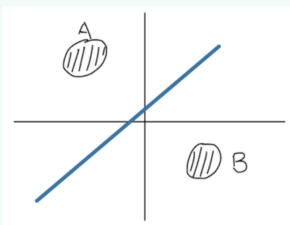
Theorem 1.7.4 Hyperplane theorem

If

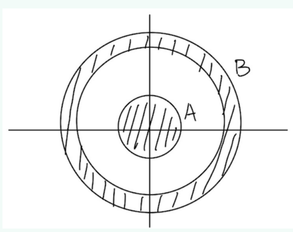
- $A, B \subset \mathbb{R}^n$
- A, B disjoint and convex

then A, B can be separated by a hyperplane.

Example 1.7.12 (Hyperplane)

Separated by a line. 

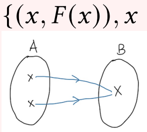
Example 1.7.13 (Hyperplane)



Separation theorem cannot be applied because B is not convex.

Definition 1.7.4: Correspondence

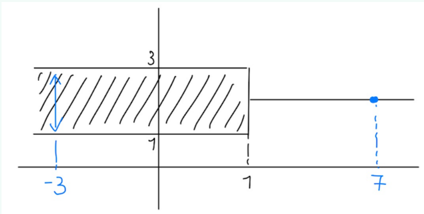
A, B sets $\subseteq \mathbb{R}$
$f : A \longrightarrow B$ map
$F : A \rightrightarrows B$
$\{(x, F(x)), x \in A, F(x) \subset B\}$



Example 1.7.14 (Corrrespondence)

$F : \mathbb{R} \rightrightarrows \mathbb{R}$

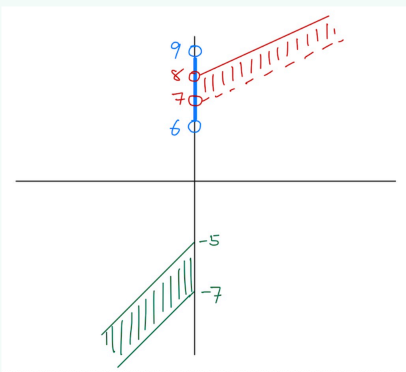
$$x \longrightarrow \begin{cases} [1, 3] & x \leq 1 \\ \{2\} & x > 1 \end{cases}$$



$F(\{7\}) = \{2\}$
$F(\{-3\}) = [1, 3]$

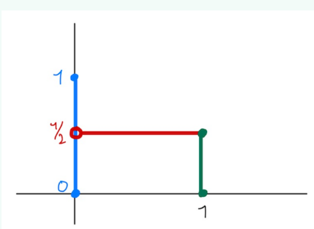
Example 1.7.15 (Correspondence)
$G : \mathbb{R} \rightrightarrows \mathbb{R}$

$$G(x) = \begin{cases} [x-7, x-5] & x < 0 \\ (6, 9) & x = 0 \\ (x+7, x+8] & x > 0 \end{cases}$$

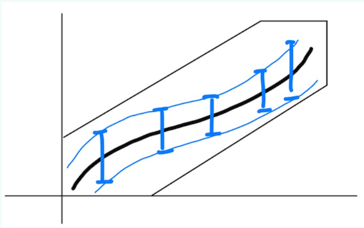


Example 1.7.16 (Correspondence)
$f : [0, 1] \rightrightarrows [0, 1]$

$$x \longrightarrow \begin{cases} [0, 1] & x = 0 \\ \{\frac{1}{2}\} & 0 < x < 1 \\ [0, \frac{1}{2}] & x = 1 \end{cases}$$



Example 1.7.17 (Correspondence)
Important for confidence interval construction of financial data.

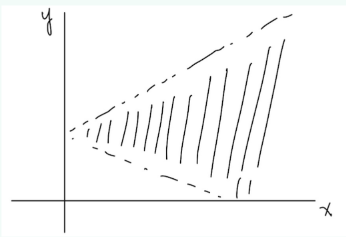


Definition 1.7.5: Closed graph property

$F : A \rightrightarrows B$ a correspondence, $x \in A$.

We say that F has the closed graph property at x if for any converging sequence $(x_n, y_n)_n$ of element in the graph F, its limit belongs to the graph of F.
We say that F has the closed graph property if it has the property above for all $x \in A$.

Example 1.7.18 (Closed graph property)

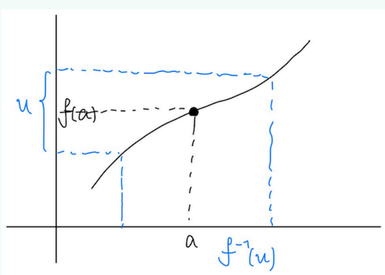


F does not have the closed graph property because the graph of F is not closed.

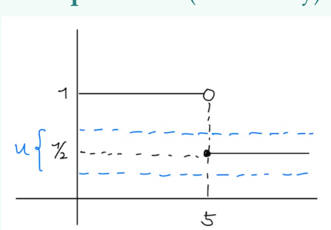
Proposition: $F : A \rightrightarrows B$ correspondence. If graph F is **compact**, then F has the closed graph property.

$f : \mathbb{R} \to \mathbb{R}, x \to f(x)$. f is continuous at $x = a \in \mathbb{R}$, if and only if $\lim_{x\to a} f(x) = f(a)$. Equivalently, f is continuous at $x = a$ if and only if for any open set u containing $f(a)$, $f^{-1}(u)$ is an open set containg a.

Example 1.7.19 (Continuity)



Example 1.7.20 (Continuity)



u open set containing $f(5) = \frac{1}{2}$, $f^{-1}(u) = [5, \infty)$ is not an open set containing 5.

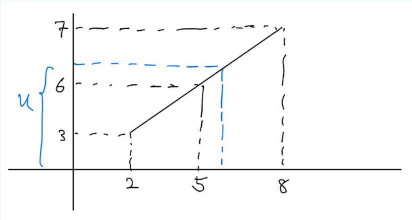
We are now going to generalize the notion of continuity for correspondence.

Definition 1.7.6: Hemi-continuity

$A, B \subseteq \mathbb{R}, a \in A$
$F : A \rightrightarrows B$ is a correspondence

We say that F is upper hemi-continuous at $x = a$ if for all open set u containing $f(a)$, its pre-image $F^{-1}(u)$ is an open set containing a. The correspondence F is upper hemi-continuous if the above property holds for every $a \in A$.

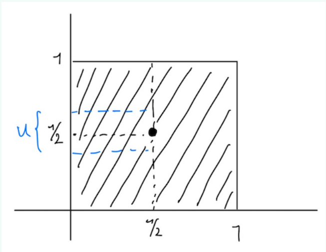
Example 1.7.21 (Hemi-continuity)



$f(\{5\}) = [0, 6]$
f is upper hemi-continuous.

Example 1.7.22 (Hemi-continuity)
$f : [0, 1] \rightrightarrows [0, 1]$

$$x \longrightarrow \begin{cases} [0,1] & x \neq \frac{1}{2} \\ \frac{1}{2} & x = \frac{1}{2} \end{cases}$$



u open set containing $f(\frac{1}{2}) = \frac{1}{2}$
$f^{-1}(u) = \{\frac{1}{2}\}$ is not an open set. $\Rightarrow f$ is not hemi-continuous.

Theorem 1.7.5 Kakutani fixed point theorem

- $A, B \subseteq \mathbb{R}^n$ convex and compact
- $F : A \rightrightarrows B$

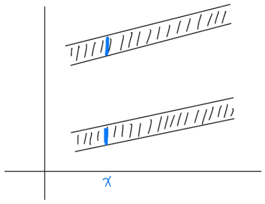
If

1. F is upper hemi-continuous
2. $F(x)$ is convex, $\forall x \in A$

then F has at least one fixed point.

Remark:

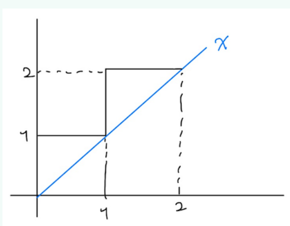
- This theorem is a generalization of the Brouwer fixed point theorem.
- The second property of Kakutani fixed point theorem means that correpondence cannot have the follwing behavior



Example 1.7.23 (Kakutani)

$F : [0, 2] \rightrightarrows [0, 2]$

$$x \longrightarrow \begin{cases} \{1\} & 0 \leq x < 1 \\ [1, 2] & x = 1 \\ \{2\} & 1 < x \leq 2 \end{cases}$$

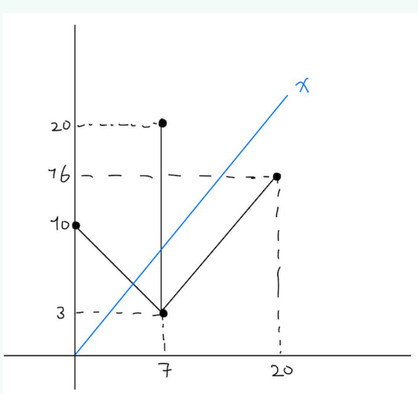


Fix $F = \{1, 2\}$

Example 1.7.24 (Kakutani)

$F : [0, 20] \rightrightarrows [0, 20]$

$$x \longrightarrow \begin{cases} 10 - x & 0 \leq x < 7 \\ [3, 20] & x = 7 \\ x - 4 & 7 < x \leq 20 \end{cases}$$



$10 - x = x \leftrightarrow x = 5$
F is a fixed point because $\{F\} \subset F(\{7\})$
Fix $F = \{5, 7\}$

Hemi-continuous is hard to prove for the Kakutani theorem, thus we have the following **proposition**: $F : A \rightrightarrows B$ correspondence. If F has the closed graph property, then F is upper hemi-continuous. The scheme is as the following: Graph F is compact $\Rightarrow$ Graph F has the closed graph property $\Rightarrow$ F is upper hemi-continous.

1.8 What is necessary to know in this chapter?

1. Functions, sequences, cardinality, continuity
2. Typology in $\mathbb{R}^n$
3. Hyperplane separation theorem
4. Weierstrass theorem
5. Intermediate value theorem
6. Banach fixed point theorem
7. Brouwer fixed point theorem
8. Kakutani fixed point theorem

Chapter 2

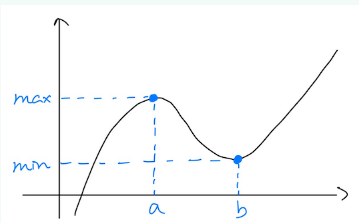
Optimization

General set up of the section:

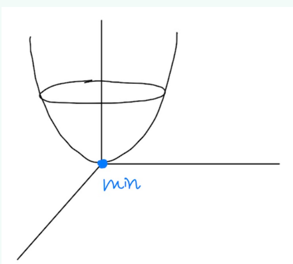
$$\begin{aligned} f: \quad & u \to \mathbb{R}, u \subseteq \mathbb{R}^n \\ & (x_1, \ldots, x_n) \to f(x_1, \ldots, x_n) \end{aligned}$$

Motivation: we are going to explore tools to compute maxima and minima of f.

Example 2.0.1 ($f : \mathbb{R} \to \mathbb{R}$)



Example 2.0.2 ($f : \mathbb{R}^2 \to \mathbb{R}$)

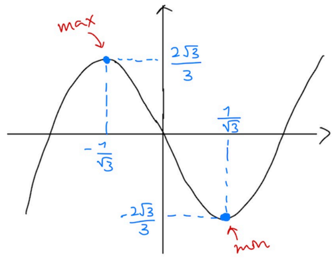


Previously,

$$\begin{aligned} f: \quad & \mathbb{R} \to \mathbb{R} \\ & x \to x^3 - x \end{aligned}$$

- Find the domain: $D_f = \mathbb{R}$
- $f'(x) = 3x^2 - 1$
- Zeros of the first derivative: $f'(x) = 0 \Leftrightarrow x = \pm\frac{1}{\sqrt{3}} \to$ critical points
- Sign of f' defines the monotony of f: $f' < 0 \to f$ is decreasing; $f' > 0 \to f$ is increasing
- $x = \frac{1}{\sqrt{3}} \to$ minimizer; $f(\frac{1}{\sqrt{3}}) = -\frac{2\sqrt{3}}{3} \to$ minimum

- $x = -\frac{1}{\sqrt{3}} \rightarrow$ maximizer; $f(-\frac{1}{\sqrt{3}}) = \frac{2\sqrt{3}}{3} \rightarrow$ maximum



If $f : u \rightarrow \mathbb{R}$ is a smooth map, it is not necessarily true that f has a maximum or minimum. However, if $f : K \rightarrow \mathbb{R}, K \subseteq \mathbb{R}^n$ where K is compact, then f has a maximum and minimum (**Weierstrass theorem**).

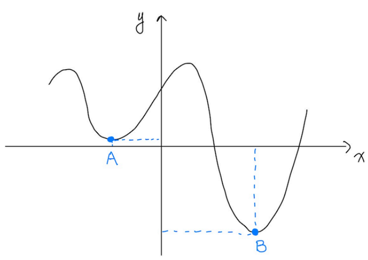
2.1 Formal definition

Definition 2.1.1: Minimizer / Maximinzer

If $f : u \rightarrow \mathbb{R}, u \subseteq \mathbb{R}^n, x_0 \in u$.

1. We say that x_o is a **global** minimizer of f if $f(x_o) \leq f(x), \forall x \in u$
2. We say that x_o is a **local** minimizer of f if there exists a neighborhood of v such that $\forall x \in v, f(x_0) \leq f(x)$

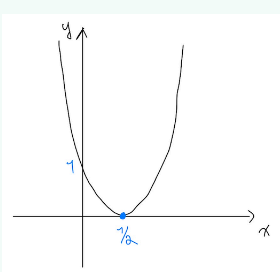
We can define analogously the global maximizer and local maximizer.



A is a local minimum and B is a global minimum.

Example 2.1.1 (Minimizer)

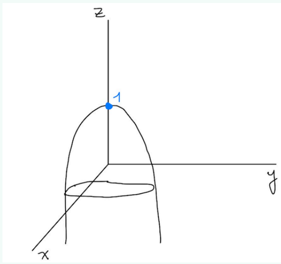
$$\begin{aligned} f : \quad & \mathbb{R} \rightarrow \mathbb{R} \\ & x \rightarrow 4x^2 - 4x + 1 \end{aligned}$$



$\frac{1}{2}$ is a global minimizer
$f(\frac{1}{2}) = 0$ is the minimum
f does not have a maximum

Example 2.1.2 (Maximizer)

$$\begin{aligned} f: \quad & \mathbb{R}^2 \to \mathbb{R} \\ & (x, y) \to 1 - x^2 - y^2 \end{aligned}$$



$(0,0)$ is the global maximizer
1 is the global maximum

2.2 Optimization: how to compute maximizer / minimizer?

$$\begin{aligned} f: \quad & u \to \mathbb{R}, u \subseteq \mathbb{R}^n \\ & (x_1, \ldots, x_n) \to f(x_1, \ldots, x_n) \end{aligned}$$

$$\underbrace{\mathrm{J}_f(x_1, \ldots, x_n)}_{\text{Jacobian Maxtrix}} = \left[\frac{\partial f}{\partial x_1}(x_1, \ldots, x_n) \cdots \frac{\partial f}{\partial x_n}(x_1, \ldots, x_n)\right]_{1 \times n}$$

Gradient of f is

$$\nabla f(x_1, \ldots, x_n) = \left(\frac{\partial f}{\partial x_1}(x_1, \ldots, x_n) \cdots \frac{\partial f}{\partial x_n}(x_1, \ldots, x_n)\right)$$

and the **critical points** of f are the zeros of $\nabla f(x_1, \ldots, x_n) \Leftrightarrow \nabla f(x_1, \ldots, x_n) = \vec{0}$

Example 2.2.1 (Critical point)

$$\begin{aligned} f: \quad & \mathbb{R}^2 \to \mathbb{R} \\ & (x, y) \to 1 - x^2 - y^2 \end{aligned}$$

Finding the gradient: $\nabla f(x_1, \ldots, x_n) = (-2x; -2y)$

Finding the zeros of the gradient: $\nabla f(x_1, \ldots, x_n) = (0,0) \Leftrightarrow \begin{cases} -2x = 0 \\ -2y = 0 \end{cases} \Leftrightarrow \begin{cases} x = 0 \\ y = 0 \end{cases}$

$\Rightarrow (0,0)$ is the unique crtical point of f

Proposition: Under the following conditions

- $f: D_f \to \mathbb{R}, D_f \subseteq \mathbb{R}^n$
- $x_0 \in \mathrm{int}(D_f)$

If x_0 is an extremum (maximum or a minimum), then x_0 is a critical point.

Wrong Concept 2.1: The reverse of the proposition

The reverse is not true. For example, $f(x) = x^3 \Rightarrow f'(x) = 3x^2$. $f'(x) = 0 \Leftrightarrow x = 0$. $\{0\}$ is not a critical point but rather a saddle point.

The previous result gives **candidates** for maximizer or minimizer. Now, another question occers: how do we check that the critical point is a maximizer or minimizer?

$$
\begin{aligned}
f: \quad & u \to \mathbb{R}, u \subseteq \mathbb{R}^n \\
\nabla f(x_1, \dots, x_n) &= \left(\frac{\partial f}{\partial x_1}(x_1, \dots, x_n) \cdots \frac{\partial f}{\partial x_n}(x_1, \dots, x_n)\right) \\
\underbrace{\mathrm{H}_f(x_1, \dots, x_n)}_{\text{Hessian Matrix}} &= \begin{bmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1} & \frac{\partial^2 f}{\partial x_2 \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_1} \\ \vdots & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \cdots & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_n} \end{bmatrix}_{n \times n}
\end{aligned}
$$

Example 2.2.2 (Hessian matrix)

$$
\begin{aligned}
f: \quad & \mathbb{R}^2 \to \mathbb{R} \\
& (x, y) \to 1 - x^2 - y^2
\end{aligned}
$$

$\nabla f(x, y) = (-2x, -2y)$

$\mathrm{H}_f(x, y) = \begin{bmatrix} -2 & 0 \\ 0 & -2 \end{bmatrix}$

Theorem 2.2.1 Schwarz theorem
If f is C^2, differentiable twice, then $\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}$.

In particular, the this result forces the Hessian matrix to be symmetric, $A^T = A$.

$$\mathrm{H}_f(x_1, \dots, x_n)^T = \mathrm{H}_f(x_1, \dots, x_n)$$

where H_f defines a quadratic form.

2.3 Quadratic forms in $\mathbb{R}^n$

Sum of monomials of degree 2 in $\mathbb{R}^n$.

Problem 2.1: Is it a quadratic form?

$\mathbb{R} : (x)$

- $P(x) = 7x^2 \to$ It is a quadratic form
- $Q(x) = 3 + 7x^2 \to$ Not a quadratic form because 3 is not of degree 2

$\mathbb{R}^2 : (x, y)$

- $P(x) = 7x^2 + 8xy \to$ It is a quadratic form
- $Q(x) = 3^2 x + y^2 \to$ Not a quadratic form because the first term is not of degree 2

$\mathbb{R}^3 : (x, y, z)$

- $P(x) = 7x^2 + 8xy + \sqrt{2}y^2 \to$ It is a quadratic form

To generalize, $\mathbb{R}^n(x_1, \dots, x_n) : P(x_1, \dots, x_n) = \sum_{i,j \in \{1,\dots,n\}} c_{ij} x_i x_j$ is a quadratic form.

Associated to the quadratic form Q in $\mathbb{R}^n$, we may define a matrix A such that

$$Q(x_1,\dots,x_n) = [x_1 \cdots x_n]A\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, A \in \mathrm{M}_{n\times n}(\mathbb{R}) \tag{2.1}$$

Example 2.3.1 (A matrix)
$P(x,y) = 3x^2 + 8xy + 5y^2$

$P(x,y) = [x \quad y]\begin{bmatrix} 3 & 5 \\ 3 & 5 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$

$P(x,y) = [x \quad y]\begin{bmatrix} 3 & 8 \\ 0 & 5 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$

There are infinitely many matrices A such that equation (2.1) holds. However, just one is symmetric. For the previous example

$$P(x,y) = [x \quad y]\begin{bmatrix} 3 & 4 \\ 4 & 5 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$$

Trick is to put the coefficient of the squared terms on the diagonal and the coefficient divided by two for the cross terms to fill in the rest of the matrix.

Problem 2.2: What is the symmatric matrix A?

$P(x,y,z) = x^2 + 8xy - y^2 - 3xz + 10z^2$

$$A = \begin{bmatrix} x^2 & xy & xz \\ yx & y^2 & yz \\ zx & zy & z^2 \end{bmatrix} = \begin{bmatrix} 1 & 4 & -\frac{3}{2} \\ 4 & -1 & 0 \\ -\frac{3}{2} & 0 & 10 \end{bmatrix}$$

2.3.1 Classification of quadratic forms in $\mathbb{R}^n$

Q : a quadratic form

Definition 2.3.1: Q classifications

1. Q is positively defined (P.D.) if $\forall x \in \mathbb{R}^n \backslash \{\vec{0}\} \quad Q(x) > 0$
2. Q is negatively defined (N.D.) if $\forall x \in \mathbb{R}^n \backslash \{\vec{0}\} \quad Q(x) < 0$
3. Q is semi-positively defined (S.P.D.) if $\forall x \in \mathbb{R}^n \quad Q(x) \geq 0 \quad \exists y \in \mathbb{R}^n \backslash \{\vec{0}\} : Q(y) = 0$
4. Q is semi-negatively defined (S.N.D) if $\forall x \in \mathbb{R}^n \quad Q(x) \leq 0 \quad \exists y \in \mathbb{R}^n \backslash \{\vec{0}\} : Q(y) = 0$
5. Q is undefined (UND.) if $\exists x, y \in \mathbb{R}^n : Q(x) = 0,\ Q(y) = 0$

Example 2.3.2 (Q classifications)

1. $Q(x,y) = x^2 + 3y^2 \longrightarrow$ is P.D.
2. $Q(x,y) = -3x^2 - 7y^2 \longrightarrow$ is N.D.
3. $Q(x,y) = \underbrace{(x-y)^2}_{\geq 0} \longrightarrow$ is S.P.D. since $Q(1,1) = 0$
4. $Q(x,y) = -(7x-y)^2 \longrightarrow$ is S.N.D.
5. $Q(x,y) = x^2 - y^2 \longrightarrow$ is UND. since $Q(1,0) = 1$ and $Q(0,1) = -1$

In general, just by observing the quadratic forms is difficult to classify. We are going to establish two criteria to help us.

Let A be the symmetric matrix associated to Q and let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues associated to A.

Theorem 2.3.1 Classification using eigenvalues

1. $\lambda_1 > 0, \ldots, \lambda_n > 0 \Rightarrow Q$ is P.D.
2. $\lambda_1 < 0, \ldots, \lambda_n < 0 \Rightarrow Q$ is N.D.
3. $\lambda_1 = 0, \lambda_2 > 0, \ldots, \lambda_n > 0 \Rightarrow Q$ is S.P.D
4. $\lambda_1 = 0, \lambda_2 < 0, \ldots, \lambda_n < 0 \Rightarrow Q$ is S.N.D
5. $\exists i, j : \lambda_i > 0, \lambda_j < 0 \Rightarrow Q$ is UND.

Remark: To find the eigenvalues

- $A \in M_{n\times n}(\mathbb{R})$
- $\exists v \neq \mathbb{R}^n : Av = \lambda v$
- $\vec{v} \longrightarrow$ eigenvector
- $\lambda \longrightarrow$ eigenvalues
- λ eigenvalues iff $P(\lambda) = 0 \longrightarrow \det(A - \lambda\, \mathrm{I}_n)$

Problem 2.3: What is the eigenvector?

$$A = \begin{bmatrix} 1 & 2 & 0 \\ 2 & 4 & 0 \\ 0 & 0 & 7 \end{bmatrix}$$

$$\begin{aligned} P(\lambda) &= \det(A - \lambda I_d) \\ &= \begin{vmatrix} 1-\lambda & 2 & 0 \\ 2 & 4-\lambda & 0 \\ 0 & 0 & 7-\lambda \end{vmatrix} \\ &= (7-\lambda)\begin{vmatrix} 1-\lambda & 2 \\ 2 & 4-\lambda \end{vmatrix} \\ &= (7-\lambda)[(1-\lambda)(4-\lambda) - 4] \\ &= (7-\lambda)(4 - \lambda - 4\lambda + \lambda^2 - 4) \\ &= (7-\lambda)\underbrace{(\lambda^2 - 5\lambda)}_{\lambda(\lambda-5)} \end{aligned}$$

$\Rightarrow \lambda = 7 \wedge \lambda = 0 \wedge \lambda = 5$

Choosing $\lambda = 7$

$$A\begin{bmatrix}x\\y\\z\end{bmatrix}=7\begin{bmatrix}x\\y\\z\end{bmatrix}\Leftrightarrow$$

$$(A-\lambda I_d)\begin{bmatrix}x\\y\\z\end{bmatrix}=\vec{0}\Leftrightarrow$$

$$\begin{bmatrix}-6&2&0\\2&-3&0\\0&0&0\end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix}\Leftrightarrow$$

$$\begin{cases}-6x+2y=0\\2x-3y=0\\z\in\mathbb{R}\end{cases}\Leftrightarrow\begin{cases}x=0\\y=0\\z\in\mathbb{R}\end{cases}$$

$\Rightarrow E_7=\langle(0,0,\mathbb{R})\rangle$

Example 2.3.3 (Classification using eigenvalues)

1. $Q(x,y)=x^2+3y^2$

   $$A=\begin{bmatrix}1&0\\0&3\end{bmatrix}\to\lambda_1=1,\ \lambda_2=3\Rightarrow \text{P.D.}$$

2. $Q(x,y)=-3x^2-7y^2$

   $$A=\begin{bmatrix}-3&0\\0&-7\end{bmatrix}\to\lambda_1=-3,\ \lambda_2=-7\Rightarrow \text{N.D.}$$

3. $Q(x,y)=(x-y)^2$

   $$A=\begin{bmatrix}1&-1\\-1&1\end{bmatrix}$$

   $$P(\lambda)=\begin{vmatrix}1-\lambda&-1\\-1&1-\lambda\end{vmatrix}=(1-\lambda)^2-1=1-2\lambda+\lambda^2-1=\lambda(\lambda-2)\to\lambda_1=0,\ \lambda_2=2\Rightarrow \text{S.P.D}$$

4. $Q(x,y)=-(7x-y)^2$

   $$A=\begin{bmatrix}-49&7\\7&-1\end{bmatrix}\to\lambda_1=0,\ \lambda_2<0\Rightarrow \text{S.N.D}$$

5. $Q(x,y)=x^2-y^2$

   $$A=\begin{bmatrix}1&0\\0&-1\end{bmatrix}\to\lambda_1=1,\ \lambda_2=-1\Rightarrow \text{UND.}$$

Associated to a quadratic form $Q:\mathbb{R}^n\to\mathbb{R}$, we may define a unique symmetric matrix A.

Theorem 2.3.2 Sylvestor's theorem: Leading Minors Method

If $\det A\neq 0$ and

- $\triangle_1>0,\triangle_2>0,\triangle_3>0,\triangle_4>0,\ldots\Rightarrow Q$ is P.D.

- $\triangle_1 < 0, \triangle_2 > 0, \triangle_3 < 0, \triangle_4 > 0, \ldots \Rightarrow Q$ is N.D.
- Q is undefined otherwise

The leading minors are the determinants of each $\triangle_i$

$$\Delta_1 \quad \Delta_2 \quad \Delta_3 \quad \cdots \quad \Delta_n$$

$$Q = \begin{pmatrix} q_{11} & q_{12} & q_{13} & \cdots & q_{1n} \\ q_{21} & q_{22} & q_{23} & & \\ q_{31} & q_{32} & q_{33} & & \vdots \\ \vdots & & & \ddots & \\ q_{n1} & & \cdots & & q_{nn} \end{pmatrix}$$

Example 2.3.4 (Classification of Q using all three methods)

1. $Q(x, y) = x^2 - y^2$
   - Definition
     $f(1, 0) = 1 > 0$
     $f(0, 1) = -1 < 0$
     $\Rightarrow$ UND.
   - Eigenvalues
     $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$
     $\lambda_1 = 1, \lambda_2 = -1 \Rightarrow$ UND.
   - Sylvester's
     $\triangle_1 = 1, \triangle_2 = -1 \Rightarrow$ UND.
2. $Q(x, y, z) = -x^2 - 2y^2 - 3z^2$
   - Definition
     All negative except at $\vec{0} \Rightarrow$ N.D.
   - Eigenvalues
     $A = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -3 \end{bmatrix}$
     $\lambda_1 = -1, \lambda_2 = -2, \lambda_3 = -3 \Rightarrow$ N.D.
   - Sylvestor's
     $\triangle_1 = -1, \triangle_2 = 2, \triangle_3 = -6 \Rightarrow$ N.D.

Theorem 2.3.3 Local minimizer / maximizer

- $f : u \to \mathbb{R}$
- $u \subseteq \mathbb{R}^n$ is open
- $\nabla f(x_0) = \vec{0},\ x_0 \in \mathbb{R}^n$

If H_f defines a P.D. quadratic form, then x_0 is a local minimizer.
If H_f defines a N.D. quandratic form, then x_0 is a local maximizer.

Definition 2.3.2: Saddle point

If x_0 is a critical point and x_0 is neither a maximizer nor a minimizer, then x_0 is a saddle point.

Example 2.3.5 (Local maximizer / minimizer)

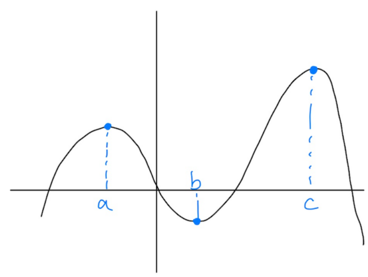
1. $f : \mathbb{R} \to \mathbb{R} \qquad x \to -(x+2)^2 + 3$

   $f'(x) = -2(x+2)$
   $f'(x) = 0 \Leftrightarrow x = -2 \Rightarrow -2$ is a critical point
   $f''(x) = -2 < 0 \Rightarrow$ is N.D.
   $\Rightarrow -2$ is a local maximizer.

2. $f : \mathbb{R}^2 \to \mathbb{R} \qquad f(x,y) = x^2 + 3y^2$

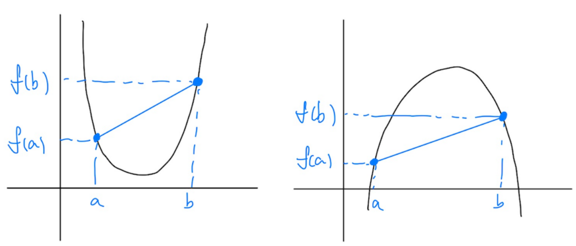
   $J_f = \begin{bmatrix} 2x & 6y \end{bmatrix} \to \nabla f(x,y) = (2x, 6y)$
   $\nabla f(x,y) = \vec{0} \Leftrightarrow \begin{cases} 2x = 0 \\ 6y = 0 \end{cases} \Leftrightarrow \begin{cases} x = 0 \\ y = 0 \end{cases}$
   $\therefore\ (0,0)$ is a critical point
   $H_f = \begin{bmatrix} 2 & 0 \\ 0 & 6 \end{bmatrix} \Rightarrow \lambda_1 = 2, \lambda_2 = 6 \Rightarrow$ defines a P.D. quadratic form
   $(0,0)$ is a local minimizer.

Now we are going to check if a local extremum could be seen as a global extremum. For example, $f : \mathbb{R} \to \mathbb{R}$



c is a global maximizer because $f(c)$ is the maximum of the map. $f(x) \le f(c)$, $\forall x \in \mathbb{R}$.

We first need to establish the concept of **convexity** of the graph of a map. The graph of f is convex if the line connecting any two points $A \hookrightarrow (a, f(a))$ and $B \hookrightarrow (b, f(b))$ is above the graph of f. The graph of f is concave if the line connecting any two points $A \hookrightarrow (a, f(a))$ and $B \hookrightarrow (b, f(b))$ is below the graph of f.



We can extend the definition for maps defined in $\mathbb{R}^n$, above $\to$ inside and below $\to$ outside.

Definition 2.3.3: Classification of H_f

If H_f defines a positvely defined quadratic form for all points $x \in \mathbb{R}^n$, then we say that H_f is positive. Analogous result for negative.

Theorem 2.3.4 Global minimizer / maximizer

$f : u \longrightarrow \mathbb{R}\,, u \subseteq \mathbb{R}^n\,, u$ is open

- H_f is positive $\Rightarrow$ graph of f is convex $\Rightarrow$ any local minimizer is a global minimizer
- H_f is negative $\Rightarrow$ graph of f is concave $\Rightarrow$ any local maximizer is a global maximizer

2.4 Optimization with restrictions given by equalities

Setting up the problem, we have a **map**

$$f : u \longrightarrow \mathbb{R}, u \subseteq \mathbb{R}^n \quad \text{differentiable}$$

and we have the **restrictions**

$$\begin{cases} g_1(x_1, \ldots, x_n) = 0 \\ g_2(x_1, \ldots, x_n) = 0 \\ \vdots \\ g_k(x_1, \ldots, x_n) = 0 \end{cases}$$

with the **hypothesis** that $\text{rank}[g_1 \cdots g_k] \neq 0$, the restrictions are not linearly dependent. For example, $x^2 + y^2 = 1, x^2 + y^2 = 4$ are linearly dependent and the hypothesis avoides this type of situation.

The technique is

$$\mathcal{L}(x_1, \ldots, x_n, \lambda_1, \ldots, \lambda_p) = f(x_1, \ldots, x_n) - \lambda_1 g_1(x_1, \ldots, x_n) - \cdots - \lambda_p g_p(x_1, \ldots, x_n)$$

where $\mathcal{L}$ is the Lagrangean map.

Theorem 2.4.1 Lagrangean

If x^* is a minimum/maximum of f with the restrictions $g_i(x) = 0, i = 1, \ldots, p$, then x_0 is a solution of

$$\nabla\mathcal{L}(x_1, \ldots, x_n, \lambda_1, \ldots, \lambda_p) = \vec{0}$$

where $\lambda_1, \ldots, \lambda_p$ are the Lagrange multipliers.

Do not forget that the problem has a solution if the hypothesis holds.

Example 2.4.1 (Lagrangean)

$$\begin{aligned} f:\ & \mathbb{R}^2 \longrightarrow \mathbb{R} \\ & (x, y) \longrightarrow xy \end{aligned}$$

with the restriction $x^2 + y^2 = 2$

We have one restriction: $g_1(x, y) = x^2 + y^2 - 2 = 0$

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

$$\begin{aligned} \nabla\mathcal{L}(x, y, \lambda) &= \left(\frac{\partial\mathcal{L}}{\partial x}, \frac{\partial\mathcal{L}}{\partial y}, \frac{\partial\mathcal{L}}{\partial\lambda}\right) \\ &= (y - 2\lambda x, x - 2\lambda y, -(x^2 + y^2 - 2)) \end{aligned}$$

$$\nabla\mathcal{L}(x, y, \lambda) = \vec{0} \Leftrightarrow \begin{cases} y - \lambda 2x = 0 \\ x - \lambda 2y = 0 \\ -(x^2 + y^2 - 2) = 0 \end{cases}$$

Now solving the system of equations.

From the first two equations, we have $y = 2\lambda x$ and $x = 2\lambda y$. Replacing y in the second equation

$$x = 2\lambda(2\lambda x) \Rightarrow x(1 - 4\lambda^2) = 0 \Rightarrow \begin{cases} x = 0 \\ \lambda = \pm\frac{1}{2} \end{cases}$$

- If $x = 0$, then then we have $y = 0$ from the first equation. But this contradicts the restriction $x^2 + y^2 = 2$. So, no critical point in this case.
- If $\lambda = \frac{1}{2}$, then from the first equation $y = x$. Replacing in the restriction $x^2 + x^2 = 2 \Rightarrow x = \pm 1 \Rightarrow y = \pm 1$. So, we have two critical points $(1, 1)$ and $(-1, -1)$.
- If $\lambda = -\frac{1}{2}$, then from the first equation $y = -x$. Replacing in the restriction $x^2 + (-x)^2 = 2 \Rightarrow x = \pm 1 \Rightarrow y = \mp 1$. So, we have two critical points $(1, -1)$ and $(-1, 1)$.

$\Rightarrow$ The critical points are $(1, 1), (-1, -1), (1, -1), (-1, 1)$

Now we have to classify the critical points. Evaluating f at each critical point,

- $f(1, 1) = 1 \longrightarrow$ maximum
- $f(-1, -1) = 1 \longrightarrow$ maximum
- $f(1, -1) = -1 \longrightarrow$ minimum
- $f(-1, 1) = -1 \longrightarrow$ minimum

Now consider the case where

$$f : u \longrightarrow \mathbb{R}, \quad u \subseteq \mathbb{R}^n \quad \text{open}$$

and restriction is given by an equality

$$g(x) = 0$$

We would use the Lagrangean technique to solve the problem. And at the end, we would get candidates for maximizers/minimizers. If $g(x) = 0$ is a compact set, then we can evaluate f at each candidate and choose the maximum/minimum value.

$$\begin{cases} f(x_1, y_1) = m_1 \\ f(x_2, y_2) = m_2 \\ \vdots \\ f(x_k, y_k) = m_k \end{cases}$$

The problem is when $g(x) = 0$ does not define a compact set.

Theorem 2.4.2

Suppose that (x^*, λ^*) is a solution of the Lagrangean system associated to f with the restriction $g(x) = 0$. If $\mathcal{L}(x, \lambda^*)$ is convex, then the critical point (x^*, λ^*) is a minimizer. If $\mathcal{L}(x, \lambda^*)$ is concave, then the critical point (x^*, λ^*) is a maximizer.

Example 2.4.2

$$f(x, y, z) = x + 2z$$

with restrictions: $\begin{cases} x + y + z = 1 \\ x^2 + y^2 + z = \frac{7}{4} \end{cases}$

Applying the Lagrangean technique,

$$g_1(x,y,z) = x + y + z - 1 = 0$$
$$g_2(x,y,z) = x^2 + y^2 + z - \frac{7}{4} = 0$$

Checking the hypothesis, we have

$$\operatorname{rank}\begin{bmatrix} g_1 & g_2 \end{bmatrix} = 2$$

which are linearly independent. So, we can apply the Lagrangean technique.

$$\mathcal{L}(x,y,z,\lambda_1,\lambda_2) = x + 2z - \lambda_1(x+y+z-1) - \lambda_2(x^2+y^2+z-\frac{7}{4})$$

$$\begin{aligned}\nabla\mathcal{L}(x,y,z,\lambda_1,\lambda_2) &= \left(\frac{\partial \mathcal{L}}{\partial x}, \frac{\partial \mathcal{L}}{\partial y}, \frac{\partial \mathcal{L}}{\partial z}, \frac{\partial \mathcal{L}}{\partial \lambda_1}, \frac{\partial \mathcal{L}}{\partial \lambda_2}\right) \\ &= (1-\lambda_1-2\lambda_2 x, -\lambda_1-2\lambda_2 y, 2-\lambda_1-\lambda_2, -(x+y+z-1), -(x^2+y^2+z-\frac{7}{4}))\end{aligned}$$

$$\nabla\mathcal{L}(x,y,z,\lambda_1,\lambda_2) = \vec{0} \Leftrightarrow \begin{cases} 1-\lambda_1-2\lambda_2 x = 0 \\ -\lambda_1-2\lambda_2 y = 0 \\ 2-\lambda_1-\lambda_2 = 0 \\ -(x+y+z-1) = 0 \\ -(x^2+y^2+z-\frac{7}{4}) = 0 \end{cases}$$
$$(\cdots)$$
$$\Leftrightarrow \begin{cases} \lambda_1 = 3 \\ \lambda_2 = -1 \\ x = 1 \\ y = \frac{3}{2} \\ z = -\frac{3}{2} \end{cases} \vee \begin{cases} \lambda_1 = 1 \\ \Lambda_2 = 1 \\ x = 0 \\ y = -\frac{1}{2} \\ z = \frac{3}{2} \end{cases}$$

Beginning with the first case,

$$H_{\mathcal{L}} = \begin{bmatrix} -2\lambda_2 & 0 & 0 \\ 0 & -2\lambda_2 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Since the eigenvalues of the Hessian are non-negative, it means that $\mathcal{L}$ is convex (but not strictly), so $(1, \frac{3}{2}, -\frac{3}{2})$ is a minimizer of the probelm.

In the second case,

$$H_{\mathcal{L}} = \begin{bmatrix} -2\lambda_2 & 0 & 0 \\ 0 & -2\lambda_2 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} -2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Since the eigenvalues of the Hessian are non-positive, it means that $\mathcal{L}$ is concave (but not strictly), so $(0, -\frac{1}{2}, \frac{3}{2})$ is a maximizer of the probelm.

2.5 Optimization with restrictions given by inequalities

The goal is to maximize/minimize the differentiable map

$$f : u \to \mathbb{R}, u \subseteq \mathbb{R}^n \quad \text{open}$$

subject to the restrictions

$$\begin{cases} l_1(x_1, \ldots, x_n) \leq 0 \\ l_2(x_1, \ldots, x_n) \leq 0 \\ \vdots \\ l_k(x_1, \ldots, x_n) \leq 0 \end{cases}$$

with $x \in \mathbb{R}^n$.

Theorem 2.5.1 Karush-Kuhn-Tucker

The solution x^* of the optimization problem with inequalities are solutions of the system:

$$\begin{cases} \nabla f(x^*) - \mu_1 \nabla l_1(x^*) - \mu_2 \nabla l_2(x^*) - \cdots - \mu_k \nabla l_k(x^*) = \vec{0} \\ \mu_1 l_1(x^*) = 0 \\ \mu_2 l_2(x^*) = 0 \\ \vdots \\ \mu_k l_k(x^*) = 0 \\ l_1(x^*) \leq 0 \\ l_2(x^*) \leq 0 \\ \vdots \\ l_k(x^*) \leq 0 \end{cases}$$

In this case, the critical points associated to $\mu_j < 0$, they are minimizers, and the critical points associated to $\mu_j > 0$, they are maximizers.

Idea for the proof: With just one restriction,

$$\mu = 0 \vee \underbrace{l_1(x^*) = 0}_{\text{on the boundary of the region}}$$

$$\mathcal{L}(x, y, \lambda) = f(x, y) - \mu l_1(x)$$

$$\nabla \mathcal{L}(x, y, \lambda) = \nabla f(x, y) - \mu \nabla f_1(x)$$

Example 2.5.1 (Karush-Kuhn-Tucker)

$$f(x, y) = x^2 - y$$

$$\text{restrictions: } x^2 + y^2 \leq 1$$

$$l_1(x, y) = x^2 + y^2 - 1 \leq 0$$

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

$$\begin{aligned} \nabla \mathcal{L}(x, y, \mu) &= \left(\frac{\partial \mathcal{L}}{\partial x}, \frac{\partial \mathcal{L}}{\partial y} \right) \\ &= (2x - \mu 2x, -1 - \mu 2y) \end{aligned}$$

$$\nabla \mathcal{L}(x, y, \mu) = \vec{0} \Leftrightarrow \begin{cases} 2x - \mu 2x = 0 \\ -1 - \mu 2y = 0 \\ \mu(x^2 + y^2 - 1) = 0 \\ x^2 + y^2 - 1 \leq 0 \end{cases}$$

From the first equation, we have $x(1-\mu) = 0 \Rightarrow \begin{cases} x = 0 \\ \mu = 1 \end{cases}$ and we would get $y = \pm 1$ from the second equation.But both points contradict the restriction. So, no critical point in this case.

But we have the second case where $\begin{cases} \mu = 1 \\ x^2 + y^2 - 1 = 0 \\ y = -\frac{1}{2} \end{cases}$, replacing in the restriction we have $x^2 + \left(-\frac{1}{2}\right)^2 - 1 = 0 \Rightarrow x = \pm\frac{\sqrt{3}}{2}$.

In addition, we have two other cases where $\begin{cases} x = 0 \\ y = 1 \\ \mu = -\frac{1}{2} \end{cases}$ and $\begin{cases} x = 0 \\ y = -1 \\ \mu = \frac{1}{2} \end{cases}$.

So, critical points are: $\underbrace{(0, 1, -\frac{1}{2})}_{\text{minimizer}}, \underbrace{(0, -1, \frac{1}{2}), (\frac{\sqrt{3}}{2}, -\frac{1}{2}, 1), (-\frac{\sqrt{3}}{2}, -\frac{1}{2}, 1)}_{\text{maximizers}}$