Topic 5: Continuity and Differentiability[1]

Keywords for Today

- Continuity
- Differentiability
- Derivatives
- Chain Rule
- Linearization
- Taylor's Theorem
- Mean Value Theorem
- L'Hôpital's Rule

Continuity

Definition 1 (Continuity at a Point)

A function $f : \mathbb{R} \to \mathbb{R}$ is continuous at a point $c \in \mathbb{R}$ if:

$$\lim_{x \to c} f(x) = f(c)$$

Formally, this means: $\forall \epsilon > 0$, $\exists \delta > 0$ such that $|f(x) - f(c)| < \epsilon$ whenever $0 < |x - c| < \delta$.

[1] Instructors: Camilo Abbate and Sofia Olguin. This note was prepared for the 2025 UCSB Math Camp for Ph.D. students in economics. It incorporates materials from previous instructors, including Shu-Chen Tsao, ChienHsun Lin, and Sarah Robinson.

Let us compare this definition to the epsilon-delta definition of a limit. The difference in the epsilon-delta definition is that we changed $|f(x) - L| < \epsilon$ to $|f(x) - f(c)| < \epsilon$.

That is, we say that function $f(x)$ is continuous at point c if $f(c)$ is exactly the L on the epsilon-delta definition of a limit. This means there are no jumps, breaks, or holes at that point c.

Definition 2 (Continuity on an Interval)

A function $f : \mathbb{R} \to \mathbb{R}$ is continuous on an interval $I \subseteq \mathbb{R}$ if it is continuous at every point $c \in I$.

Definition 3 (Continuity of a Function)

A function $f : \mathbb{R} \to \mathbb{R}$ is continuous on its entire domain if it is continuous at every point in its domain.

Exercise 1

Determine whether the following functions f are continuous on $[-1, 1]$.

(1) $f(x) = |x|$

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

(3) $f(x) = \begin{cases} -1 & \text{if } x < 0 \\ 0 & \text{if } x = 0 \\ 1 & \text{if } x > 0 \end{cases}$

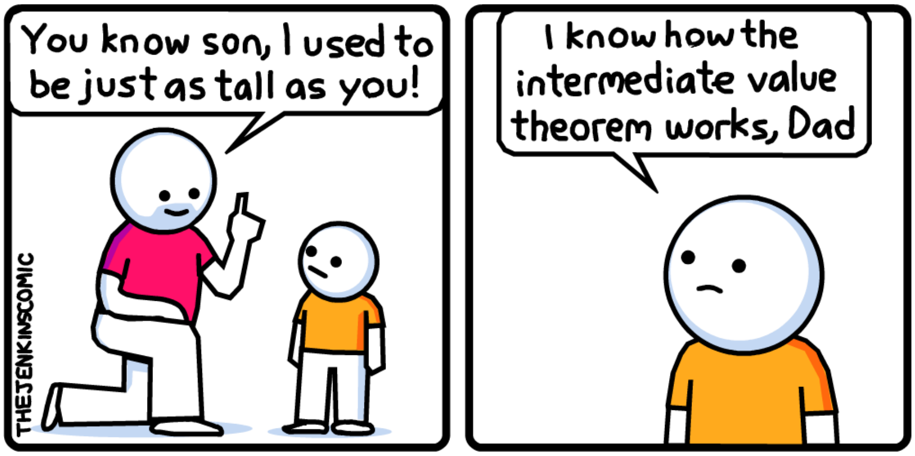
(4) $f(x) = \begin{cases} |x| & \text{if } x \text{ is rational} \\ -x^2 & \text{otherwise} \end{cases}$

There are three important properties of a continuous function.

Theorem 1: Intermediate Value Theorem

If $f : [a, b] \to \mathbb{R}$ is continuous on the closed interval $[a, b]$, and L is any number between $f(a)$ and $f(b)$ $(\min\{f(a), f(b)\} < L \leq \min\{f(a), f(b)\})$, then there exists at least one $c \in [a, b]$ such that:

$$f(c) = L$$



(Figure source: The Jenkins Comic).

Exercise 2

In fact, the information in the comic above is not sufficient to ensure the dad's "argument" to hold. There are some implicit conditions assumed.

Let $f(t)$ be the mapping from time to the dad's height. Let $f(a)$ be the dad's height at birth, and let $f(b)$ be the dad's height now. Let L be the child's height now.

We observe in the comic that $L < f(b)$. Which other conditions do we need to guarantee the statement that "The dad used to be just as tall as the son"?

The intermediate value theorem may sound trivial, but note that we don't need f to be differentiable or monotone.

Theorem 2: Extreme Value Theorem

If $f : [a, b] \to \mathbb{R}$ is continuous on the closed interval $[a, b]$, then f attains both a maximum and a minimum value on $[a, b]$.

In other words, there exist points $c, d \in [a, b]$ such that:

$$f(c) \leq f(x) \leq f(d) \quad \text{for all } x \in [a, b]$$

The extreme value theorem is fundamental in optimization.

Exercise 3

Consider $f(x) = (x-1)^2$.

(1) Does f attain both maximum and minimum values on $[0, 2]$?
(2) Does f attain both maximum and minimum values on $(0, 2)$? Which part of Theorem 2 is violated?

Exercise 4

Consider

$$f(x) = \begin{cases} x & \text{if } x \neq 1 \\ -1 & \text{if } x = 1 \end{cases}$$

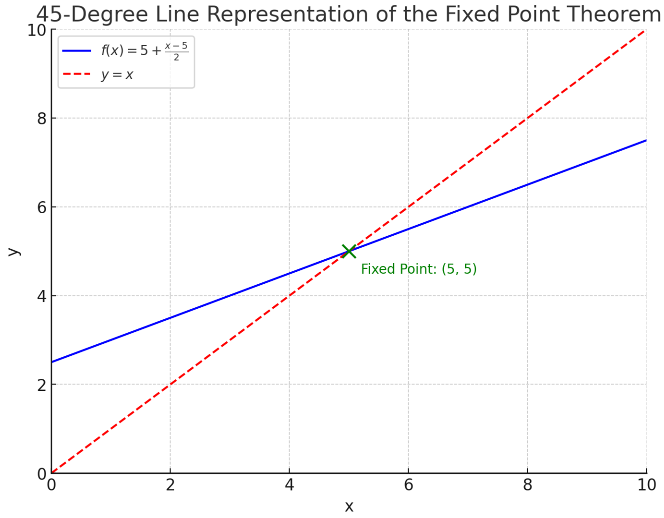
Does f attain both maximum and minimum values on $[0, 1]$? Which part of Theorem 2 is violated?

Theorem 3: Fixed Point Theorem

If $f : [a, b] \to [a, b]$ is continuous, then there exists at least one point $c \in [a, b]$ such that:

$$f(c) = c$$

The fixed-point theorem is fundamental in game theory. One of the most famous application of the fixed-point theorem is the Nash equilibrium. We also use the fixed-point theorem to derive the steady states in macroeconomics.



There is a corresponding fixed point theorem in $\mathbb{R}^N$, which is called the Brouwer's fixed point theorem. The intuition is as followed: Imagine compressing a sponge ball inward. Then, there will be at least one point inside the sponge ball that remains fixed.

The intermediate value theorem and extreme value theorem also hold in $\mathbb{R}^N$. Let's define the continuity in $\mathbb{R}^N$ formally.

Definition 4 (Continuity at a point in $\mathbb{R}^n$)

A function $f : \mathbb{R}^n \to \mathbb{R}^m$ is continuous at a point $\mathbf{c} = (c_1, c_2, \ldots, c_n) \in \mathbb{R}^n$ if:

$$\lim_{\mathbf{x} \to \mathbf{c}} f(\mathbf{x}) = f(\mathbf{c})$$

Formally, this means: $\forall \epsilon > 0$, $\exists \delta > 0$ such that $\|f(\mathbf{x}) - f(\mathbf{c})\| < \epsilon$ whenever $\|\mathbf{x} - \mathbf{c}\| < \delta$.

Here, $\|\cdot\|$ denotes the Euclidean norm, which for a vector $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ in $\mathbb{R}^n$ is given by:

$$\|\mathbf{x}\| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}.$$

Definition 5 (Continuity on a Set in $\mathbb{R}^n$)

A function $f : \mathbb{R}^n \to \mathbb{R}^m$ is continuous on a set $S \subseteq \mathbb{R}^n$ if it is continuous at every point $\mathbf{c} \in S$.

Differentiability

Definition 6 (Difference Quotient)

Let $f : \mathbb{R} \to \mathbb{R}$ be a function. We define the difference quotient of f as follows:

$$\frac{f(x') - f(x)}{x' - x} \text{ for any } x, x' \in X.$$

Intuitively, this quotient is the **slope** of f across the points x and x'.

If we make the two points extremely close, then the quotient becomes the slope at the point, or the derivative.

Definition 7 (Derivatives)

Let $f : \mathbb{R} \to \mathbb{R}$ be a function defined on some open interval containing a point a.

The derivative of f at the point a, denoted by $f'(a)$, is defined as the limit (if it exists):

$$f'(a) = \lim_{h \to 0} \frac{f(a+h) - f(a)}{h}$$

Definition 8 (Differentiability)

A function f is said to be differentiable at a point a if the derivative $f'(a)$ exists.

If f is differentiable at every point in an interval, we say that f is differentiable on that interval.

Theorem 4: Differentiability implies Continuity

If a function f is differentiable at a point a, then f is continuous at a.

However, note that the continuity **does not** imply differentiability.

Exercise 5

Determine if $f(x) = |x|$ is continuous at $x = 0$. Determine if $f(x) = |x|$ is differentiable at $x = 0$.

Here are some derivatives for common functions.

(1) Constants: for any $r \in \mathbb{R}$,
$$f(x) = r \;\Rightarrow\; f'(x) = 0.$$

(2) Powers of x: for any $k \in \mathbb{R} \setminus \{0\}$,
$$f(x) = x^k \;\Rightarrow\; f'(x) = kx^{k-1}.$$

(3) Polynomial functions
$$f(x) = a_n x^n + \cdots + a_1 x + a_0 = \sum_{k=0}^{n} a_k x^k$$
$$\Rightarrow f'(x) = n a_n x^{n-1} + \cdots + a_1 = \sum_{k=1}^{n} k a_k x^{k-1}.$$

(4) The exponential function:
$$f(x) = \exp(x) \;\Rightarrow\; f'(x) = \exp(x).$$

(5) The natural logarithm function:
$$f(x) = \log(x) \;\Rightarrow\; f'(x) = \frac{1}{x}.$$

Theorem 5

Let f, g are defined on $[a, b]$ and are differentiable at $x \in [a, b]$. Then the functions $f + g$, fg, and f/g are differentiable, and

(1) $(f + g)'(x) = f'(x) + g'(x)$;

(2) $(fg)'(x) = f'(x)g(x) + f(x)g'(x)$;

(3) $\left(\frac{f}{g}\right)'(x) = \frac{f'(x)g(x) - f(x)g'(x)}{g^2(x)}$.

We sometimes use the differential operator $\frac{d}{dx}$ to express the operation "taking derivatives."

Hence, when f is differentiable,

$$\frac{d}{dx}f(x) \equiv \frac{df}{dx}(x) \equiv f'(x).$$

If we define Δ as:

$$\Delta x = x' - x,\ \Delta f(x) = f(x') - f(x)$$

then the difference quotient can then be expressed as $\frac{\Delta f(x)}{\Delta x}$. When we take the limit to the difference quotient, it becomes the notation that we are familiar with, $\frac{df(x)}{dx}$.

When you take derivatives of the composite functions, you will need to apply the chain rule.

Theorem 6

Suppose $f : X \to Y$ is continuous on $[a, b]$, $f'(x)$ exists at some point $x \in [a, b]$, g is defined on an interval $I \in Y$, and g is differentiable at $f(x)$.

Let $h(x) = g \circ f(x) = g(f(x))$, then

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

.

Exercise 6

Find the derivative of the following functions.

(1) $f(x) = \sqrt{x-2}$
(2) $f(x) = \log(x^2)$
(3) $f(x) = (\log x)^2$

There is a strong connection between derivatives and monotonic behavior.

Theorem 7

Suppose f is differentiable in the open interval (a, b). Then

(1) f is monotonically increasing if $f'(x) \geq 0$;

(2) f is monotonically decreasing if $f'(x) \leq 0$;

(3) f is a constant if $f'(x) = 0$.

Linearization

An application of derivatives is to approximate the function values around a specific point (say, a). Let's review the definition of derivatives of f at a. Linearization is used in macroeconomics where the functions of interest might not have analytical representation.

$$f'(a) = \lim_{h \to 0} \frac{f(a+h) - f(a)}{h}$$

By rearranging the terms,

$$\lim_{h \to 0} \{ \frac{f(a+h) - f(a)}{h} - f'(a) \} = 0$$

Define $\varepsilon(h) = \frac{f(a+h)-f(a)}{h} - f'(a)$, and note that $\lim_{h \to 0} \varepsilon(h) = 0$. We can rearrange the terms in this definition of $\varepsilon(h)$ to have:

$$f(a+h) = f(a) + f'(a)h + \varepsilon(h)h$$

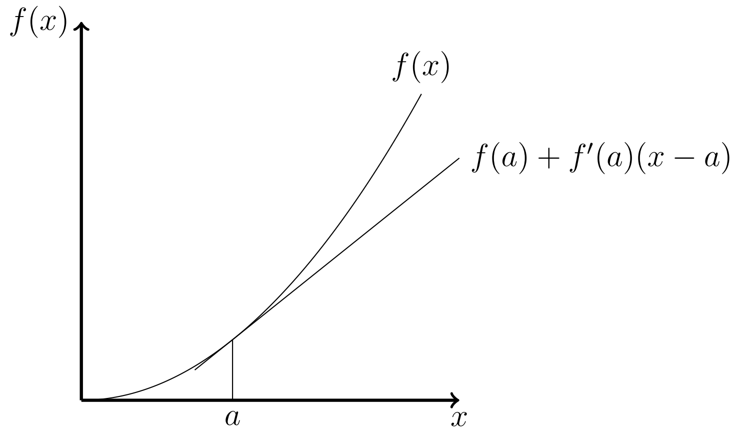
Define $x = a + h$, then we have:

$$f(x) = f(a) + f'(a)(x-a) + \varepsilon(x-a)(x-a)$$

Hence we can loosely write

$$f(x) \approx f(a) + f'(a)(x-a).$$

With this process, we approximate $f(x)$ with a linear function of the intercept of $f(a)$ and a slope of $f'(a)$. This is called a **linear approximation** of f at a.



Higher Order Derivatives

Let $g = \frac{df}{dx}$. If g is also differentiable, we can denote

$$g' \equiv \frac{d}{dx}\frac{df}{dx} \equiv \frac{d^2 f}{(dx)^2} \equiv f''.$$

This function f'' is called the **second order derivative** of f. The interpretation of f'' is *the change in slopes*. We can further define the higher order derivatives with the similar manner, and we denote the k-th derivative of f with $f^{(k)}$.

We say f is in the family of functions $\mathcal{C}^k$ if the derivatives f', f'',$\ldots$,$f^{(k)}$ exists and are continuous. The function f is **smooth** or **infinitely differentiable** if for any $k \in \mathbb{N}$, $f^{(k)}$ exists, and we denote $f \in \mathcal{C}^\infty$.

Taylor's theorem

Taylor's theorem provides a way to approximate the function values around a specific point.

Theorem 8

Let $f : \mathbb{R} \to \mathbb{R}$ be a function that is $(n+1)$-times continuously differentiable on an open interval containing the point a.

Then, $\forall x$ in this interval, the function $f(x)$ can be expressed as the **Taylor polynomial**:

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n + R_n$$

where the remainder term $R_n(x)$ is given by the **Lagrange form of the remainder**:

$$R_n = \frac{f^{(n+1)}(c)}{(n+1)!}(x-a)^{n+1}$$

for some c between a and x.

Exercise 7

Write out the Taylor polynomial of $f(x)$ around point a, with $n = 0$.
Write out the Taylor polynomial of $f(x)$ around point a, with $n = 1$. Compare it with the linearization.

Suppose $f(x) = x^2$ and $a = 1$. Draw the above two Taylor polynomials.

When $n = 0$, the Taylor theorem is called the mean value theorem.

Theorem 9: The mean value theorem

Let $f : [a, b] \to \mathbb{R}$ be continuous on $[a, b]$ and differentiable on (a, b).

Then, there exists $c \in (a, b)$ such that

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

Here is an exercise for writing down a model and derive results using math tools.

Exercise 8

The distance between downtown SB and UCSB is 10 miles. Suppose your friend tells you it took them 7.9 minutes to drive from downtown SB to UCSB.

Define speeding as there exists at least one point in time such that the speed, which exists in $\mathbb{R}$, is greater than 75 miles per hour. Use the mean value theorem to prove that your friend was speeding according to this definition.[a]

[a]Extension: What if your friend's vehicle can teleport? Can we still prove that they are speeding according to this definition? If not, which condition for the mean value theorem is violated?

L'Hôpital's rule

L'Hôpital's rule is used to find the limit of a quotient of two functions when the limit results in an indeterminate form, such as $\frac{0}{0}$ or $\frac{\infty}{\infty}$.

Theorem 10: L'Hôpital's rule

Suppose functions f and g are both differentiable on an open interval I containing c, except possibly at c itself.

If:

- $\lim_{x\to c} f(x) = 0$ and $\lim_{x\to c} g(x) = 0$, or
- $\lim_{x\to c} f(x) = \pm\infty$ and $\lim_{x\to c} g(x) = \pm\infty$,

and the limit of their derivatives exists:

$$\lim_{x\to c} \frac{f'(x)}{g'(x)} = L,$$

Then,

$$\lim_{x\to c} \frac{f(x)}{g(x)} = \lim_{x\to c} \frac{f'(x)}{g'(x)} = L.$$

Exercise 9

1. Evaluate the following limit: $\lim_{x\to 0} \frac{e^x-1}{x}$. Find the limit if it exists.

2. Evaluate the following limit: $\lim_{x\to 0} \frac{|x|}{x}$. Find the limit if it exists.

Topic 6: Topology on Metric Space[1]

Keywords for Today

Euclidean space, l^p space, open and closed ball, open and closed set, **compact** set, Heine-Borel Theorem, convext set, convex and concave function, quasiconvex and quasiconcave function, Jensen's inequality, Cauchy-Schwarz inequality, Hölder's inequality, Minkowski inequality

Metric Space

We studied the metric function last week, which measures the "distance" between elements in a set. Let's review the definition of a metric function:

Definition 1 (Metric Function and Metric Space)

Let X be a set. If there exists a **metric function** $d : X \times X \to \mathbb{R}^+$ satisfying the following conditions:

(1) $d(\mathbf{x}, \mathbf{x}) = 0$ for every $\mathbf{x} \in X$

(2) $d(\mathbf{x}, \mathbf{y}) > 0$ for every $\mathbf{x} \neq \mathbf{y} \in X$

(3) $d(\mathbf{x}, \mathbf{y}) = d(\mathbf{x}, \mathbf{y})$ for every $\mathbf{x}, \mathbf{y} \in X$

(4) $d(\mathbf{x}, \mathbf{y}) \leq d(\mathbf{x}, \mathbf{z}) + d(\mathbf{y}, \mathbf{z})$ for every $\mathbf{x}, \mathbf{y}, \mathbf{z} \in X$

then (X, d) is called a **metric space**. We say that d is a **metric** on X.

The last property is also called **triangle inequality**.

For example, the set X can be $X = \mathbb{R}^n$. Then, there are several different metrics d defined on $\mathbb{R}^n$ to form a metric space $(\mathbb{R}^n, d)$. Among the metrics defined on $\mathbb{R}^n$, the

[1] Instructors: Camilo Abbate and Sofia Olguin. This note was prepared for the 2025 UCSB Math Camp for Ph.D. students in economics. It incorporates materials from previous instructors, including Shu-Chen Tsao, ChienHsun Lin, and Sarah Robinson.

most commonly used on is the **Euclidean distance**.

Definition 2 (Euclidean Metric)

Let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. The **Euclidean distance** of $\mathbf{x}$ and $\mathbf{y}$ is defined as

$$d(\mathbf{x}, \mathbf{y}) = ||\mathbf{x} - \mathbf{y}||_2 = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \cdots + (x_n - y_n)^2}.$$

When $X = \mathbb{R}^n$ and d is the Euclidean metric, we often call this metric space an **Euclidean space**. However, Euclidean metric is not the only metric that can be defined on $\mathbb{R}^n$.

Let's see an example to illustrate the motivation of having metrics other than the Euclidean metric. Let's consider a network of individuals, where the "distances" between individuals are all identical. This modeling method motivates the following metric space.

Exercise 1

Let X be any set, and $\mathbf{x}, \mathbf{y} \in X$. Define

$$d(\mathbf{x}, \mathbf{y}) = \begin{cases} 1 & if \quad x \neq y \\ 0 & if \quad x = y \end{cases}$$

Verify that (X, d) is a metric space.

l^p SPACE

In $\mathbb{R}^n$, each element is $\mathbf{x} = (x_1, x_2, ..., x_n)$, where each $x_i \in \mathbb{R}$. What if we are interested in the metric for elements look like $\mathbf{x} = (\underbrace{x_1, x_2, ...}_{\text{infinite of them}})$?

When $\mathbf{x} = (x_1, x_2, ..., x_n)$, $\sum_{i=1}^{n} x_i^2$ is finite. However, when $\mathbf{x} = (x_1, x_2, ...)$, $\sum_{i=1}^{\infty} x_i^2$ is not necessarily finite. We focus on the cases where $\sum_{i=1}^{\infty} x_i^2$ is finite.

Definition 3 (l^2 Space)

Define

$$l^2 = \left\{ (x_1, x_2, ...) \mid \sum_{i=1}^{\infty} x_i^2 < \infty \right\}$$

And define the metric between $\mathbf{x}, \mathbf{y} \in l^2$:

$$d(\mathbf{x}, \mathbf{y}) = ||\mathbf{x} - \mathbf{y}||_2 = \left(\sum_{i=1}^{\infty} |x_i - y_i|^2 \right)^{1/2}$$

Then, (l^2, d) is a metric space. We call it a l^2 space.

The l^2 space can be generalized to a l^p space, where $p \geq 1$.

Definition 4 (l^p Space)

Define

$$l^p = \left\{ (x_1, x_2, ...) \mid \sum_{i=1}^{\infty} |x_i|^p < \infty \right\}$$

And define the metric between $\mathbf{x}, \mathbf{y} \in l^p$:

$$d(\mathbf{x}, \mathbf{y}) = ||\mathbf{x} - \mathbf{y}||_p = \left(\sum_{i=1}^{\infty} |x_i - y_i|^p \right)^{1/p}$$

Then, (l^p, d) is a metric space. We call it a l^p space.

Exercise 2

Verify that l^1 space (i.e., the l^p space with $p = 1$) is a metric space.

Open Ball and Closed Ball

Definition 5 (Open Ball)

For a metric space (X, d), an open ball centered at $x \in X$ with radius $r > 0$ is defined as:

$$B(x, r) = \{y \in X : d(x, y) < r\}$$

Definition 6 (Closed Ball)

For a metric space (X, d), an closed ball centered at $x \in X$ with radius $r > 0$ is defined as:

$$\overline{B}(x, r) = \{y \in X : d(x, y) \leq r\}$$

An open ball includes all points strictly within the radius but excludes the boundary points where the distance to x is exactly r. A closed ball, on the other hand, includes all points within the radius including those on the boundary.

Note that if we try to visualize an open ball (or a closed ball), it may not look like a "ball" if we are considering a metric space other than the Euclidean space.

Exercise 3

Let X be $\mathbb{R}$. For any $x, y \in X$, define

$$d(x, y) = |x - y|$$

Draw the open ball

$$B(x = 3, r = 2)$$

Exercise 4

Let X be $\mathbb{R}^2$. For any $\mathbf{x}, \mathbf{y} \in X$, define

$$d(\mathbf{x}, \mathbf{y}) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}$$

Draw the open ball

$$B(x = (0, 0), r = 2)$$

Exercise 5

Let X be $\mathbb{R}^2$. For any $\mathbf{x}, \mathbf{y} \in X$, define

$$d(\mathbf{x}, \mathbf{y}) = |x_1 - y_1| + |x_2 - y_2|$$

Draw the open ball

$$B(x = (0,0), r = 2)$$

Exercise 6

Let X be $\mathbb{R}^2$. For any $\mathbf{x}, \mathbf{y} \in X$, define

$$d(\mathbf{x}, \mathbf{y}) = \max\{|x_1 - y_1|, |x_2 - y_2|\}$$

Draw the open ball

$$B(x = (0,0), r = 2)$$

Exercise 7

Let X be any set. For any $\mathbf{x}, \mathbf{y} \in X$, define

$$d(\mathbf{x}, \mathbf{y}) = \begin{cases} 1 & if \quad x \neq y \\ 0 & if \quad x = y \end{cases}$$

Write down the sets

$$B(x, r = 1.1)$$

and

$$B(x, r = 0.9)$$

Open Set and Closed Set

After having the definition of open ball, we can define an "open set."

Definition 7 (Open Set)

Let (X, d) be a metric space. We call $V \subset X$ an **open set** if:

$$\forall x \in V, \quad \exists r > 0$$

such that

$$B(x, r) \subset V$$

In other words, for every point x in the set V, we draw an open ball centered at x, and the open ball in entirely inside V.

Definition 8 (Closed Set)

Let (X, d) be a metric space. We call $F \subset X$ a **closed set** if its complement $X \setminus F$ is an open set.

Theorem 1

Any open ball $B(X, r)$ is an open set, and any closed ball $\overline{B}(X, r)$ is a closed set.

The first is easy to see, but the second requires some proof. However, we will skip the proof in this note.

Exercise 8

Let X be $\mathbb{R}$. For any $\mathbf{x}, \mathbf{y} \in \mathbb{R}$, define $d(\mathbf{x}, \mathbf{y}) = |\mathbf{x} - \mathbf{y}|$. Check if the following sets are open or closed.

1. $(0, 1)$
2. $(0, \infty)$
3. $[0, \infty)$
4. $[0, 1]$

As the definition of closed set entails the complement set, one might think that a set is either open or closed (but not both). However, this guess is not true.

Exercise 9

Let X be $\mathbb{R}$. For any $\mathbf{x}, \mathbf{y} \in \mathbb{R}$, define $d(\mathbf{x}, \mathbf{y}) = |\mathbf{x} - \mathbf{y}|$. Verify that the following sets are both open and closed.

1. $\emptyset$
2. $\mathbb{R}$

Let (X, d) be any metric space. Verify that the following sets are both open and closed.

1. $\emptyset$
2. X

Theorem 2

The union of any collection of **open** sets is always an **open** set, and the intersection of any finite number of **open** sets is always an **open** set.

The union of any finite number of **closed** sets is always a **closed** set, and the intersection of any collection of **closed** sets is always a **closed** set.

Note that the intersection of an infinite collection of **open** sets may not be open. Besides, the union of an infinite collection of **closed** sets may not be closed.

Exercise 10

For the following two questions, let X be $\mathbb{R}$. For any $\mathbf{x}, \mathbf{y} \in \mathbb{R}$, define $d(\mathbf{x}, \mathbf{y}) = |\mathbf{x} - \mathbf{y}|$.

(1) Consider $I_n = (-\frac{1}{n}, \frac{1}{n}), n = 1, 2, ...$
For any n, is I_n an open set or a closed set (or both)?
Find the intersection of an infinite collection, that is, find $\cap_{n=1}^{\infty} I_n$. Is it an open set or a closed set (or both)?
(2) Consider $J_n = [\frac{1}{n}, 1 - \frac{1}{n}], n = 1, 2, ...$

For any n, is I_n an open set or a closed set (or both)?
Find the union of an infinite collection, that is, find $\cup_{n=1}^{\infty} J_n$. Is it an open set or a closed set (or both)?

Compact Set

For economists, the main goal of learning open ball and open set is to derive the concept of a "compact" set. A compact set is a set that is "small" in a certain way, and various theorems in optimization or statistics rely on assuming that we are analyzing a compact set.

Definition 9 (Open Cover)

A **cover** of a set A, denoted as $\mathcal{F}$, is a collection of sets whose union includes A. In math, it means $A \subset \mathcal{F}$. Or, equivalently,

$$\forall x \in A, \quad \exists V \in \mathcal{F} \text{ such that } x \in V$$

A cover of a set A, denoted as $\mathcal{F}$, is an **open cover** if it is a collection of open sets whose union includes A.

Exercise 11

Consider $\mathcal{F} = \{(\frac{1}{n}, \frac{2}{n}) | n = 1, 2, ...\})$.

(1) Is $\mathcal{F}$ a collection of open sets?
(2) Is $\mathcal{F}$ an open cover of $(0, 1)$?

Definition 10 (Subcover)

A **subcover** is a smaller collection of sets from the original cover that still covers the entire set A.

Definition 11 (Compact Set)

Let (X, d) be a metric space. $A \subset X$.

A set A is called a compact set if every open cover of A has a finite subcover. In other words,

$$\forall \text{ open cover } \mathcal{F}, \exists V_1, V_2, ..., V_n \in \mathcal{F} \text{ such that } A \subset \cup_{i=1}^{n} V_i$$

This means that no matter how you try to cover the set with open sets, you can always find a finite number of those sets that still cover all of A. In other words, compact sets don't have "holes" or "gaps" that require an infinite number of pieces to cover them completely.

Exercise 12

Use the definition of compact, prove that $(0, 1) \in \mathbb{R}^1$ is not compact. (Hint: consider the open cover $\mathcal{F} = \{(\frac{1}{n}, \frac{2}{n})|n = 1, 2, ...\}$)

One important property of a compact set is that it is a bounded set and a closed set.

Definition 12 (Bounded Set)

Let (X, d) be a metric space. A set $A \subset X$ is called **bounded** if:

$$\exists M > 0 \text{ and } x_0 \in X \text{ such that } d(x, x_0) \leq M \text{ for all } x \in A$$

Exercise 13

Prove that $(0, 1)$ is bounded.

A bounded set is one where the points don't stretch out infinitely far from each other. You can think of a bounded set as one that can be enclosed within a "box" or "ball" of a certain finite size.

Theorem 3

Let (X, d) be a metric space. If $A \subset X$ is compact, then A is bounded and closed.

Generally speaking, the reverse is not always true for all metric space. However, the reverse is true in an Euclidean space.

Theorem 4: Heine-Borel Theorem

Let (X, d) be a metric space, where $X = \mathbb{R}^n$ and d is the Euclidean metric.

$A \subset X$ is compact if and only if A is closed and bounded.

Heine-Borel Theorem is the most important one of our lecture today. This theorem is often how we verify if a set in $\mathbb{R}^n$ is compact. When we are focusing on an Euclidean space, we don't need to verify the compactness via constructing finite subcovering of any open cover. We only need to verify that the set is bounded and closed.

Exercise 14

Are the following sets bounded? Are they closed? Are they compact?

1. $[0, 1]$
2. $(0, 1)$
3. $[0, \infty)$
4. $(0, \infty)$

Convex Set

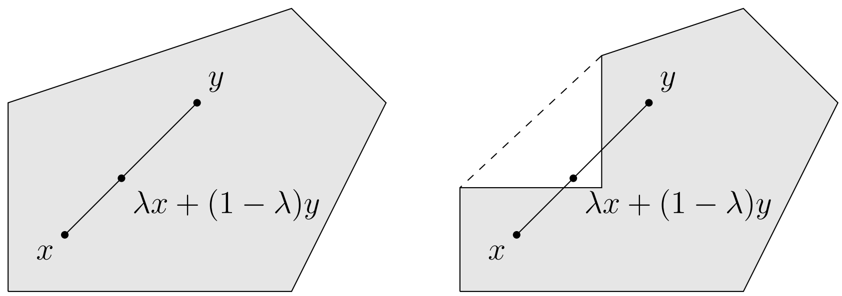
Next, we introduce another important concept widely used in economics and optimization: convex set.

Definition 13

A set $E \subset \mathbb{R}^n$ is a **convex set** if:

$$\forall \mathbf{x}, \mathbf{y} \in E \text{ and } \lambda \in [0,1] \text{ , the point } \lambda\mathbf{x} + (1-\lambda)\mathbf{y} \in E$$

Intuitively, the combination $\lambda\mathbf{x} + (1-\lambda)\mathbf{y}$ is a point on the section connecting $\mathbf{x}$ and $\mathbf{y}$. A convex set is a set that contains all sections between two points in the set. In other words, if you can connect two points in the set E, where some parts on the connecting section is not in the set E, then the E is not convex.



Consider any collection of points $p_1, \ldots, p_n$. The linear combination

$$\sum_{i=1}^{n} \alpha_i p_i = \alpha_1 p_1 + \cdots + \alpha_n p_n, \qquad \text{where } \sum_{i=1}^{n} \alpha_i = 1,\ \alpha_i \geq 0 \text{ for every } i$$

is called a **convex combination** of points $p_1, \ldots, p_n$.

Convex Function

Definition 14 (Convex/concave functions)

Let $f : X \to \mathbb{R}$. We say f is a **convex function** if for any $x_0, x_1 \in X$, $\lambda \in [0,1]$,

$$f(\lambda x_0 + (1-\lambda)x_1) \leq \lambda f(x_0) + (1-\lambda)f(x_1).$$

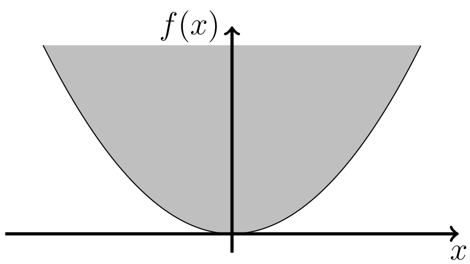
We say f is a **concave function** if for any $x, y \in X$, $\lambda \in [0,1]$,

$$f(\lambda x_0 + (1-\lambda)x_1) \geq \lambda f(x_0) + (1-\lambda)f(x_1).$$

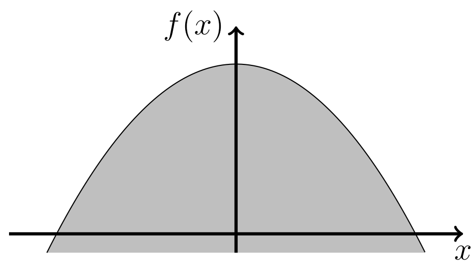
We say f is a **linear function** if f is convex and concave.

If f is a concave function, $-f$ is a convex function. This small trick is frequently used in programming.

Notice very carefully the difference between a convex set and a convex function. The level set (e.g., the contour line for $f(x,y)$) of a convex function is always a convex set.



convex function



concave function

Here is an useful theorem of using derivatives to verify the convexity of a function $f : \mathbb{R} \to \mathbb{R}$. We will discuss the multivariate case later.

Theorem 5

For a function $f : \mathbb{R} \to \mathbb{R}$ that is differentiable on an interval I, the the function

f is convex if its first derivative is increasing on I. That is,

$$f'(x_1) \leq f'(x_2) \text{ whenever } x_1 < x_2 \in I$$

For a function $f : \mathbb{R} \to \mathbb{R}$ that is twice differentiable on an interval I, the the function f is convex if its second derivative is positive on I. That is,

$$f''(x) > 0 \quad \forall x \in I$$

To verify the concavity of a function f, we can just verify the convexity of $-f$.

Definition 15 (Quasiconvex/quasiconcave functions)

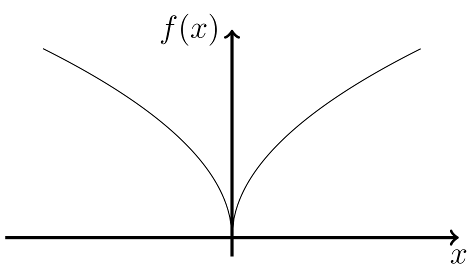
Let $f : X \to \mathbb{R}$. We say f is a **quasiconvex function** if for any $x_0, x_1 \in X$, $\lambda \in [0, 1]$,

$$f(\lambda x_0 + (1 - \lambda)x_1) \leq \max\{f(x_0), f(x_1)\}.$$

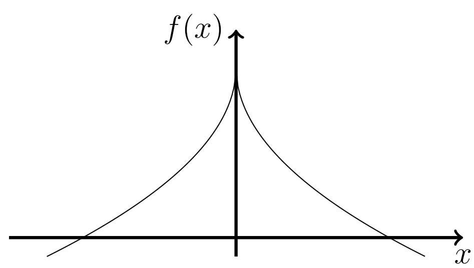
We say f is a **quasiconcave function** if for any $x, y \in X$, $\lambda \in [0, 1]$,

$$f(\lambda x_0 + (1 - \lambda)x_1) \geq \min\{f(x_0), f(x_1)\}.$$

We say f is a **quasilinear function** if f is quasiconvex and quasiconcave.



quasiconvex function



quasiconcave function

Theorem 6

If f is convex, then it is quasiconvex. If f is concave, then it is quasiconcave.

Some Common Inequalities

Now, we summarize some inequalities commonly used in economics and statistics.

Theorem 7: Jensen's Inequality in $\mathbb{R}^n$

Consider any $\mathbf{x}_1, ..., \mathbf{x}_n \in \mathbb{R}^n$, and any $\lambda_1, ..., \lambda_n \in [0, 1]$ such that $\sum_{i=1}^n \lambda_i = 1$. This forms a convex combination $\sum_{i=1}^n \lambda_i \mathbf{x}_i$. Suppose f is convex.

The Jensen's inequality states:

$$f\left(\sum_{i=1}^n \lambda_i \mathbf{x}_i\right) \leq \sum_{i=1}^n \lambda_i f(\mathbf{x}_i)$$

Jensen's inequality is basically restating the definition of convex function.

Theorem 8: Cauchy-Schwarz Inequality in $\mathbb{R}^n$

Consider any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, where $\mathbf{x} = (x_1, ..., x_n)$ and $\mathbf{y} = (y_1, ..., y_n)$. The Cauchy-Schwarz inequality states:

$$\left(\sum_{i=1}^n x_i y_i\right)^2 \leq \left(\sum_{i=1}^n x_i^2\right)\left(\sum_{i=1}^n y_i^2\right)$$

For example, in $\mathbb{R}^1$, the Cauchy-Schwarz inequality states $(xy)^2 \leq x^2 y^2$.

Theorem 9: Hölder's Inequality in $\mathbb{R}^n$

Consider any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, where $\mathbf{x} = (x_1, ..., x_n)$ and $\mathbf{y} = (y_1, ..., y_n)$. The Hölder's inequality states:

$$\sum_{i=1}^n |x_i y_i| \leq \left(\sum_{i=1}^n |x_i|^p\right)^{\frac{1}{p}} \left(\sum_{i=1}^n |y_i|^q\right)^{\frac{1}{q}}$$

for all $p, q \in [1, \infty)$ and $\frac{1}{p} + \frac{1}{q} = 1$.

Note that the Hölder's inequality is a generalization of Cauchy-Schwarz inequality.

Exercise 15

Show that when $p = q = 2$, the Hölder's inequality implies Cauchy-Schwarz inequality. (Hint: there is one more step after letting $p = q = 2$, which involves $|a + b| \leq |a| + |b|$.)

Theorem 10: Minkowski inequality in $\mathbb{R}^n$

Consider any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, where $\mathbf{x} = (x_1, ..., x_n)$ and $\mathbf{y} = (y_1, ..., y_n)$. The Minkowski inequality states:

$$\left(\sum_{i=1}^{n} |x_i + y_i|^p\right)^{\frac{1}{p}} \leq \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}} + \left(\sum_{i=1}^{n} |y_i|^p\right)^{\frac{1}{p}}$$

for all $p \in [1, \infty)$.

Exercise 16

Apply appropriate inequalities to the following questions.

(1) When we have the data $\mathbf{x} = (x_1, ..., x_n)$ and $\mathbf{y} = (y_1, ..., y_n)$, we often evaluate their relationship by calculating the correlation coefficient: $r = \frac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \overline{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \overline{y})^2}}$. Prove that $-1 \leq r \leq 1$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.

(2) Prove that $(\frac{a+b+2c+d}{5})^8 \leq \frac{a^8+b^8+2(c)^8+d^8}{5}$.

(3) In undergraduate statistics, we learned that the variance of X can be written as $Var(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2$. Prove that $\mathbb{E}[X^2] - \mathbb{E}[X]^2 \geq 0$ using algebra (i.e., not using the property that $Var(X) \geq 0$ directly). For simplicity, consider that X is a discrete random variable with some p.d.f. $P(X)$.

From (3), you might wonder if $\mathbb{E}[X^2] - \mathbb{E}[X]^2 \geq 0$ also work for a continuous random variable X. More generally, if the inequalities also work for integrals instead of summation. The answer is yes, and we will revisit these inequalities in the lecture for integrals.