MATHEMATICS OF DATA SCIENCE

September 26, 2023

Program

Today : Convex functions Part 1/2
(Next Tuesday : Part 2/2)

Thursday : Tutorials 5.15pm - 6.45pm

Announcements :

- Next week : Tutorials
  Gr 02 C. ROYER Tuesday 5.15pm - 6.45pm
  Gr 01 J. LESCA Thursday 5.15pm - 6.45pm

- October 10 (Tuesday) : No class !
  ⇒ Moved to Tuesday December 12

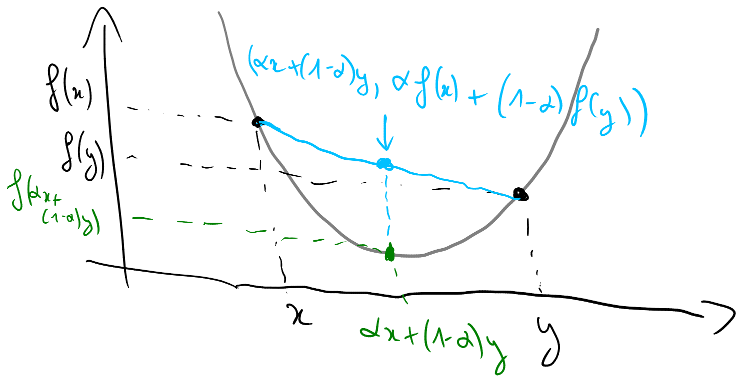
CONVEX FUNCTIONS

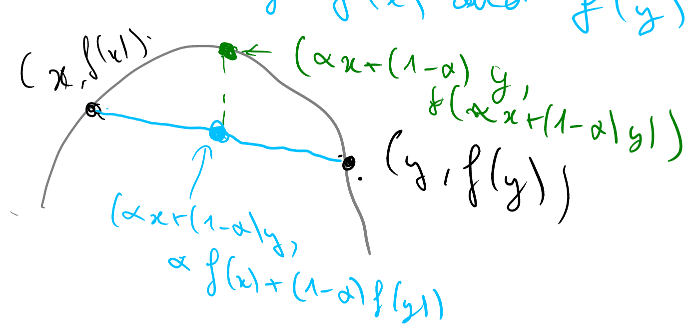
(Section 1.2 of the lecture notes)

① Definition and examples

Def: Let $\mathcal{X} \subseteq \mathbb{R}^n$ be a convex set.
and $f : \mathcal{X} \subseteq \mathbb{R}^n \to \mathbb{R}$.
The function f is called convex (on $\mathcal{X}$) if
$\forall (x,y) \in \mathcal{X}^2$, $\forall \alpha \in [0,1]$,

$$f(\underbrace{\alpha x + (1-\alpha) y}_{\in \mathcal{X}}) \leq \underbrace{\alpha f(x) + (1-\alpha) f(y)}_{\text{convex combination of } f(x) \text{ and } f(y)}$$





Additional definitions ($\mathcal{X} \subseteq \mathbb{R}^n$ convex set, $f : \mathcal{X} \to \mathbb{R}$)

- f is strictly convex if
$\forall (x,y) \in \mathcal{X}^2$ with $x \neq y$, $\forall \alpha \in (0,1)$
$$f(\alpha x + (1-\alpha) y) < \alpha f(x) + (1-\alpha) f(y)$$

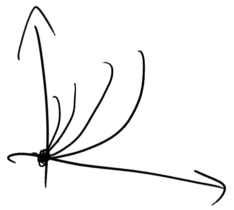
- f is concave (on $\mathcal{X}$) if
$\forall (x,y) \in \mathcal{X}^2$, $\forall \alpha \in [0,1]$
$$f(\alpha x + (1-\alpha) y) \geq \alpha f(x) + (1-\alpha) f(y)$$

Equivalently, f is concave if $-f$ is convex.

Examples (in 1 dimension $n=1$)

- $x \mapsto ax+b$ with $a \in \mathbb{R}, b \in \mathbb{R}$ is convex and concave on $\mathbb{R}$ (convex + concave $\Rightarrow$ linear)
- $x \mapsto x^a$, $a \in \mathbb{R}$

  Convex when $a > 1$ on $\mathbb{R}_+ = [0, \infty)$
  or when $a < 0$ on $\mathbb{R}_{++} = (0, \infty)$

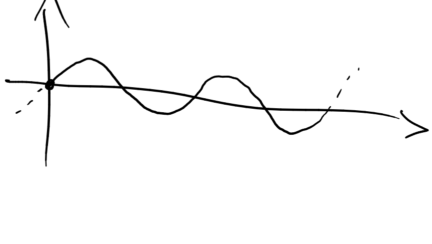


$\swarrow x \mapsto x^{-1}$

- $x \mapsto e^{-x}$ is strictly convex
- $x \mapsto \ln x$ is concave on $\mathbb{R}_{++}$

  $x \mapsto -\ln x$ is convex on $\mathbb{R}_{++}$

Counter-example: $x \mapsto \sin(x)$ is neither convex nor concave on $\mathbb{R}$



Examples in dimension $n \geq 1$

- $x \mapsto \frac{1}{2}\|x\|^2 = \frac{1}{2}\sum_{i=1}^{n} x_i^2$ is convex on $\mathbb{R}^n$
- $x \mapsto \|x\|_N$ where $\|.\|_N$ is a norm on $\mathbb{R}^n$ is a convex function on $\mathbb{R}^n$

(Any norm $\|.\|_N$ satisfies $\|x+y\|_N \leq \|x\|_N + \|y\|_N \quad \forall (x,y)$
and $\|\lambda x\|_N = |\lambda| \, \|x\|_N \quad \forall \lambda \in \mathbb{R}$

$\forall (x,y) \in (\mathbb{R}^n)^2, \quad \forall \alpha \in [0,1],$

$$\|\alpha x + (1-\alpha) y\|_N \leq \|\alpha x\|_N + \|(1-\alpha) y\|_N$$
$$= \alpha \|x\|_N + (1-\alpha) \|y\|_N \quad)$$

- $x \mapsto \ln\left(\sum_{i=1}^{n} e^{x_i}\right)$ is convex on $\mathbb{R}^n$
- $x \mapsto a^T x + b$ $\quad a \in \mathbb{R}^n, b \in \mathbb{R}$ (linear from $\mathbb{R}^n$ to $\mathbb{R}$) is both convex and concave

↳ The characterization of convex functions given in the definition can be generalized

Jensen's inequality

Let $X \subseteq \mathbb{R}^n$ be a convex set
Let $f : X \to \mathbb{R}$ be convex

Then, $\forall x_1, \ldots, x_k \in X$, $\forall \alpha_1, \ldots, \alpha_k \in \mathbb{R}$ such that $\alpha_i \geq 0 \; \forall i = 1..k$ and $\sum_{i=1}^{k} \alpha_i = 1$,

$$f\Big(\underbrace{\sum_{i=1}^{k} \alpha_i x_i}_{\in X}\Big) \leq \sum_{i=1}^{k} \alpha_i f(x_i)$$

Some operations that preserve convexity

- "Any conic combination of convex functions is convex"

$\forall f_1, \ldots, f_k : X \subseteq \mathbb{R}^n \to \mathbb{R}$ with X convex, f_i convex on X $\forall i$

$\forall \alpha_1, \dots, \alpha_k \in \mathbb{R}$ such that $\alpha_i \geq 0 \; \forall i = 1..k$

$f = \sum_{i=1}^{k} \alpha_i f_i$ is convex on $\mathcal{X}$

- If f: $\mathbb{R}^n \to \mathbb{R}$ is convex on $\mathbb{R}^n$.

  and $g : \mathbb{R}^m \to \mathbb{R}$

  $g(y) = f(Ay + b)$ where $A \in \mathbb{R}^{n \times m}$ and $b \in \mathbb{R}^n$

  Then g is convex on $\mathbb{R}^m$

- Let $f_1, \dots, f_m$ be m functions from $\mathbb{R}^n$ to $\mathbb{R}$

  and Suppose that $f_1, \dots, f_m$ are convex on $\mathbb{R}^n$

  let $g : \mathbb{R}^m \to \mathbb{R}$ be nondecreasing with respect to each of its arguments

  For the 1st argument $\left[\begin{array}{l} \forall (x,y) \in \mathbb{R}, \; \forall a_2, \dots, a_m \in \mathbb{R}, \\ x \leq y \Rightarrow g(x, a_2, \dots, a_m) \leq g(y, a_2, \dots, a_m) \end{array} \right.$

  Then, the composition function

  $h : \mathbb{R}^m \to \mathbb{R}$

  $x \mapsto g(f_1(x), f_2(x), \dots, f_m(x))$

  is convex

- $f_1, f_2 : \mathbb{R}^n \to \mathbb{R}$ convex

  Then $x \mapsto \max(f_1(x), f_2(x))$ is convex.

Also true for more than two functions

② Extended value functions

↳ Motivation : The definition of convexity that we gave is not general enough. In particular, it does not cover functions that take values $-\infty$ or $+\infty$

(Ex : $\mathbb{R}_x \to \mathbb{R}$
$$x \mapsto \begin{cases} 1/x & \text{if } x > 0 \\ +\infty & \text{if } x = 0 \end{cases}$$)

$\Rightarrow$ The convexity inequality
$$f(\alpha x + (1-\alpha) y) \leq \alpha f(x) + (1-\alpha) f(y)$$
is ambiguous when $f(x) = +\infty$ and $f(y) = -\infty$

$\Rightarrow$ We need another characterization

Def: Extended value function

- A function f with inputs in $\mathbb{R}^n$ is called extended value if $f(x) \in \overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\} = [-\infty, \infty]$

- The domain of $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is defined as
$$\mathrm{dom}(f) = \{x \in \mathbb{R}^n \mid f(x) < +\infty\}$$

Def: Epigraph

Let $f: \mathbb{R}^n \to \overline{\mathbb{R}}$ ($X \subseteq \mathbb{R}^n$ convex). The epigraph of f, $\mathrm{epi}(f)$ is defined as

$$\mathrm{epi}(f) = \{ (x,t) \in \mathbb{R}^{n+1} \mid x \in \overset{X}{\mathbb{R}^n} \text{ and } t \geq f(x) \}$$

Def: $f: \mathbb{R}^n \to \overline{\mathbb{R}}$ is called convex if its epigraph is a convex set.

__Remarks__:

- This definition holds when f is real-valued, that is when $f: \mathbb{R}^n \to \mathbb{R}$
- It illustrates a connection between convex functions and convex sets: can characterize convexity of a function by convexity of a set (the epigraph)

Q) Can we characterize convexity of a set through convexity of a function?

A) Yes!

Def: Let $X \subseteq \mathbb{R}^n$. The __indicator function__ of X is defined as $\delta_X : \mathbb{R}^n \to \mathbb{R}$

$$\forall x \in \mathbb{R}^n, \quad \delta_X(x) = \begin{cases} 0 & \text{if } x \in X \\ +\infty & \text{otherwise} \end{cases}$$

th) $X \subseteq \mathbb{R}^n$ is convex $\Leftrightarrow$ δ_X is convex
$\Leftrightarrow$ δ_X is convex on X

$$f(\alpha x + (1-\alpha) y) \leq \alpha f(x) + (1-\alpha) f(y) \ ?$$
$\in X$!
$\forall (x,y) \in X$? $\forall \alpha \in [0,1]$