Class #12

Friday, February 28, 2025 9:35 AM

A convex set implies that any convex combination of points in the set lie in the same set.

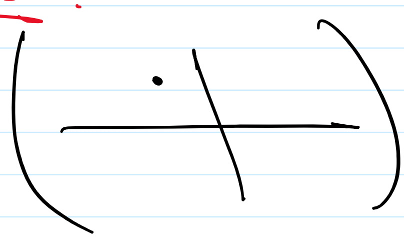
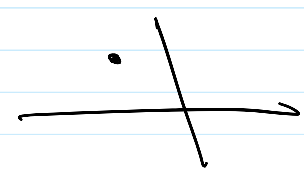
Convex hull

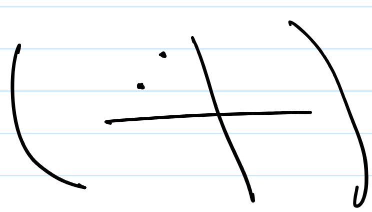
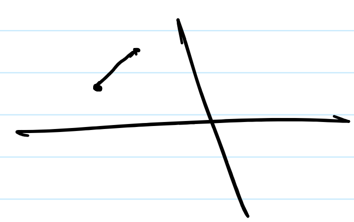
The convex hull of a set C is the set of all convex combinations of points in C

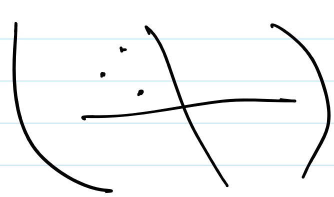
$$\text{conv } C = \{\theta_1 x_1 + \theta_2 x_2 + \cdots + \theta_k x_k : x_i \in C;\ \theta_i \geq 0;\ \sum_{i=1}^{k} \theta_i = 1\}$$

Convex hull of C is convex $\iff$ It is the smallest convex set that contains C.

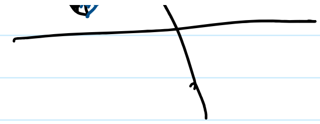
Examples:

Conv  = 

Conv  = 

Conv  = 

Conv  = 

Conv  = 

Convex combination $\Rightarrow$

$$\left.\begin{array}{l}\theta_i \geq 0, \; i = 1, \dots, k \\ \sum_{i=1}^{k} \theta_i = 1\end{array}\right\}$$ Probability mass function

$$\left.\begin{array}{l}p_X(x) \geq 0 \\ \int p_X(x)\,dx = 1\end{array}\right\}$$ Probability density functions

$$E[X] = \sum_{i=1}^{k} \underbrace{p_X(x_i)}_{\geq 0} x_i$$

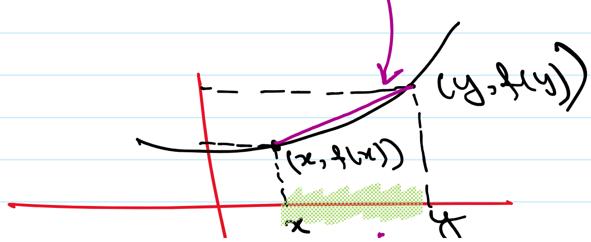
<u>Reading:</u> 2.1.4 (BV)

Convex functions

Let $f: \mathbb{R}^n \to \mathbb{R}$ with dom f being convex. Then f is termed a convex function if $\forall\ x, y \in \text{dom} f$ and $\theta \in [0, 1]$

$$f(\underbrace{\theta x + (1-\theta) y}_{\in\ \text{dom}\, f}) \leq \theta f(x) + (1-\theta) f(y)$$

Jensen's Inequality



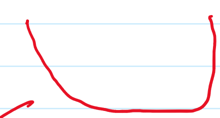
inequality

The chord (line segment) connecting $(x, f(x))$ and $(y, f(y))$ should lie above the function between x and y.

Strictly convex function

If $f(\theta x + (1-\theta)y) < \theta f(x) + (1-\theta) f(y)$ $\forall\, x, y \in \text{dom} f$ then f is called a strictly convex function.

Ex: A linear function is convex but it is not strictly convex.



not strictly convex but it is convex.

Concave functions: If $-f$ is convex then f is called concave (similarly strictly concave).

What if $\text{dom} f \neq \mathbb{R}^n$?

Extensions of convex functions on all $\mathbb{R}^n$

Let $f: \mathbb{R}^n \to \mathbb{R}$ be convex on $\text{dom} f$.

The convex extension of f on $\mathbb{R}^n$ is defined

The convex extension of f on $\mathbb{R}^n$ is defined as:

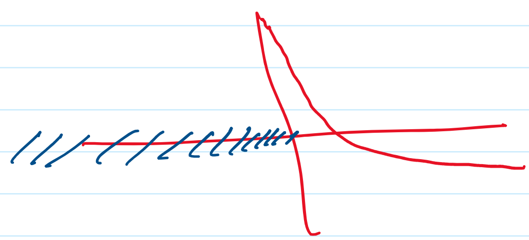
$$\tilde{f}(x) = \begin{cases} f(x), & x \in \text{dom} f \\ \infty, & x \notin \text{dom} f \end{cases}$$

↳ using this trick, we can ignore that $\text{dom} f \neq \mathbb{R}^n$

$$\text{dom} f = \{x : \tilde{f}(x) < \infty\}$$

Ex: $f(x) = -\log x$

$\text{dom} f = (0, \infty)$



Constrained Optimization trick

Let C be a convex set

Suppose we need to solve

$$\min_{x \in C} f(x) \quad ; \quad f : \mathbb{R}^n \to \mathbb{R}$$

$$\tilde{I}_C(x) = \begin{cases} 0, & x \in C \\ \infty, & x \notin C \end{cases} \Rightarrow \text{Convex function}$$

$$\min_{x \in \mathbb{R}^n} \left[f(x) + \tilde{I}_C(x) \right]$$

Reading: 3.1.1 and 3.1.2 (BV)

Convex function, convex combination, and Probability

Convex function, convex Combination, and Probability

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

$$\Rightarrow f(\underbrace{\boxed{\theta_1 x_1 + \cdots + \theta_k x_k}}_{\hookrightarrow \mathbb{E}[X]}) \leq \underbrace{\theta_1 f(x_1) + \cdots + \theta_k f(x_k)}_{\mathbb{E}[f(X)]}$$

$\theta_i \geq 0, \sum_{i=1}^{k} \theta_i = 1$

Let x_i's be the values that a random variable X takes and θ_i are the probabilities

$$P(X = x_i) = \theta_i$$

If f is convex $\Rightarrow f(\mathbb{E}[X]) \leq \mathbb{E}[f(X)]$

$\hookrightarrow$ Jensen's Inequality

$\hookrightarrow$ This holds even when X is continuous random variable.

Reading: 3.1.8 (BV)

Equivalent Characterizations of Convex Functions

① A function $f: \mathbb{R}^n \to \mathbb{R}$ is convex if and only if dom f is convex and its restriction to any line in $\mathbb{R}^n$ is convex:

Define: $g(t): \mathbb{R} \to \mathbb{R}$

$$g(t) = f(x + tv) \quad \forall\, x + tv \in \text{dom} f$$

Then f is convex $\Longleftrightarrow$ $g(t)$ is convex for every x and v.

② First-order condition of convexity

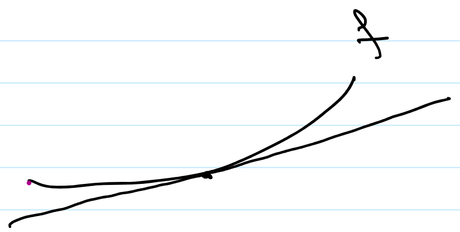
Let $f: \mathbb{R}^n \to \mathbb{R}$ be differentiable on $\operatorname{dom} f$ and $\operatorname{dom} f$ is open.

$f: \mathbb{R}^n \to \mathbb{R}$ is convex if and only if $\operatorname{dom} f$ is convex and

$$f(y) \geq \underbrace{f(x) + \nabla f(x)^T (y-x)}_{} \qquad (*)$$

$\forall\, x, y \in \operatorname{dom} f$

first-order approximation (linear approximation) of f around



The first-order approximation must be a uniform underestimator of f.

Global optimality condition for convex functions

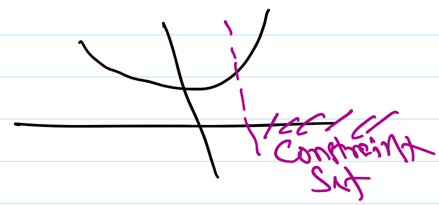
Let $x_0 \in \operatorname{dom} f$ be such that $\nabla f(x_0) = 0$ then x_0 is a global minimizer of f.

i.e. $\quad f(x_0) \leq f(y) \quad \forall\, y \in \operatorname{dom} f$

Proof: Take $x = x_0$ in $(*)$ (first-order convexity condition)

$$f(y) \geq f(x_0) + 0$$

$$\Rightarrow f(x_0) \leq f(y) \quad \forall\, y \in \operatorname{dom} f \qquad ///$$



Unconstrained optimization

x_0 is a global minimizer of $f \in C^1$ if and only if

$$\nabla f(x_0) = 0$$

Strictly Convex functions

$$f(y) > f(x) + \nabla f(x)^T (y - x)$$

$\Rightarrow$ A strictly convex function can only have a unique minimizer.

Indeed: Let x_1 and x_2 be two global minimizers with $\nabla f(x_1) = 0 = \nabla f(x_2)$

$$f(x_1) < f(y) \quad \forall\, y \in \operatorname{dom} f$$

$$f(x_1) < f(x_2)$$

$$\Rightarrow \text{Contradiction} \qquad ///$$

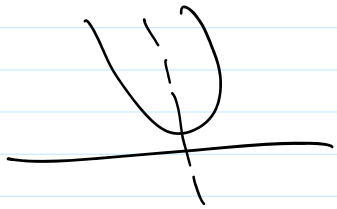
③ Monotonicity of gradients

Suppose $f: \mathbb{R}^n \to \mathbb{R}$ is differentiable with $\operatorname{dom} f$ being convex. Then f is convex if and only if convex. Then f is convex if and only if

$$(\nabla f(x) - \nabla f(y))^T (y-x) \geq 0 \quad \forall\ x, y \in \text{dom} f$$

This generalizes the concept of monotonicity of functions to $\mathbb{R}^n$

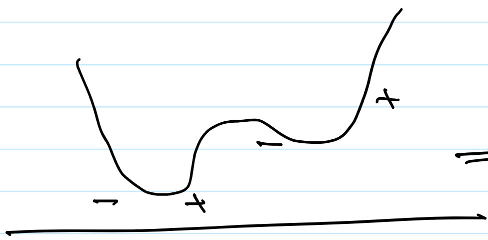
E.g.:



$f(x) = x^2$

$\Rightarrow f'(x) = 2x$

$\Rightarrow$ Increasing $\Rightarrow$ convex



$\Rightarrow$ not convex

(4) Second-order condition of convexity

Suppose $f: \mathbb{R}^n \to \mathbb{R}$ is twice differentiable with $\text{dom} f$ being open.

f is convex if and only if $\text{dom} f$ is convex and

$$\underbrace{\nabla^2 f(x) \succcurlyeq 0}_{\text{Positive semi definite.}} \quad \forall\ x \in \text{dom} f$$

Basically, the function at every point x has non-negative curvature.

Concave : $\nabla^2 f(x) \preceq 0$

Strictly convex functions $\not\Rightarrow \nabla^2 f(x) \succ 0$

Ex. $f(x) = x^4 \Rightarrow$ strictly convex.
but $f''(x) = 0$ at $x = 0$