Week 3 # CONVEX FUNCTIONS

Let us look at the conditions / tests for convexity of functions.

A function f is convex function when these two conditions holds

Zeroth Order Condition

(a) dom(f) is convex set

(b) for $x, y \in$ dom(f)

$$f(\theta x + (1-\theta)y) \le \theta f(x) + (1-\theta) f(y), \text{ for } \theta \in [0,1].$$

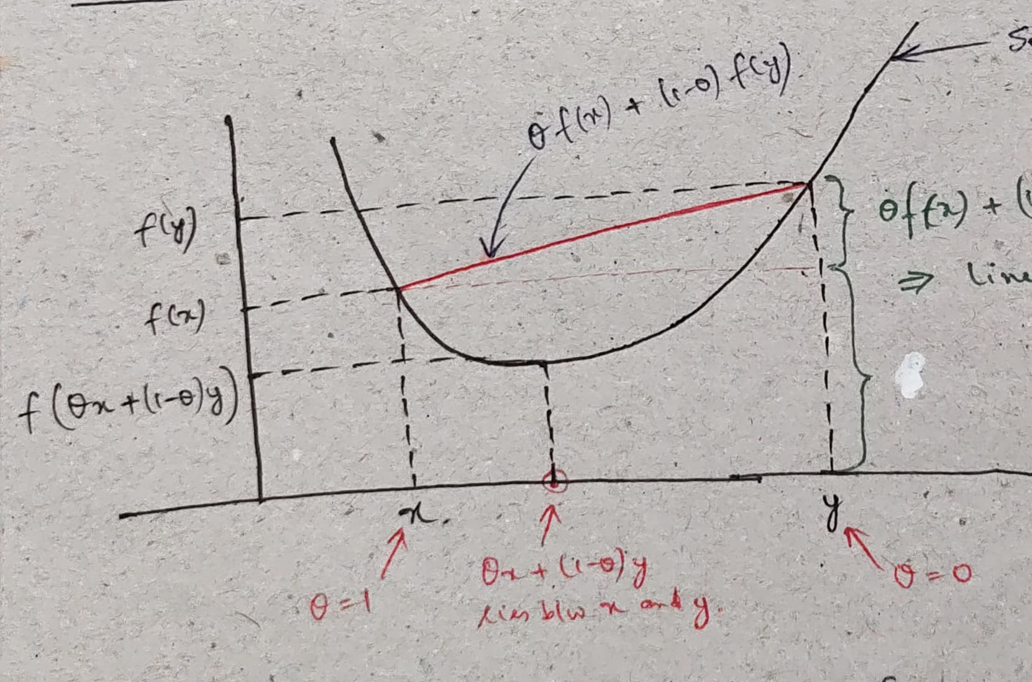
- Note that, $x, y \in$ dom(f). (ie) dom(f) is convex.
  Then only $\theta x + (1-\theta)y \in$ dom(f).
- Condition (a) has to be TRUE, then only condition (b) could be TRUE.

Example. $f(x) = \frac{1}{x}$, $x \neq 0$ is not convex.

since dom(f) $= \mathbb{R} \setminus \{0\}$ is not a convex set (Union of disjoint sets).

Because, the domain is not a convex set, the function $f(x)$ is not a convex function.

Geometric Intuition for condition of convexity



$$\theta f(x) + (1-\theta) f(y) \ge f(\theta x + (1-\theta)y)$$

$\Rightarrow$ line $\ge$ function

$\forall\, x, y \in$ dom(f)

Function f is concave $\Longleftrightarrow$ $-f$ is convex.

(a) dom(f) is convex set

(b) $f(\theta x + (1-\theta)y) \ge \theta f(x) + (1-\theta) f(y)$ This line will be below the function curve.

Example. $f(x) = \|x\|$. Is this a convex function?

(a) $\text{dom}(f) = \mathbb{R}^n$.

(i) Domain of Norm is $\mathbb{R}^n$. $\Rightarrow$ Convex Set.

(b) $f(\theta x + (1-\theta) y) = \| \theta x + (1-\theta) y \|$

$\leq \|\theta x\| + \|(1-\theta) y\|$ Triangle inequality.

$= |\theta| \|x\| + |1-\theta| \|y\|$ Homogeneity

$= \theta \|x\| + (1-\theta) \|y\|$ $0 \leq \theta \leq 1$

$= \theta f(x) + (1-\theta) f(y)$.

$\Rightarrow \|x\|$ is a convex function

$\Rightarrow f$ is convex.

Let us look into more complicated example.

$$f(x) = \max\{x_1, x_2, \ldots, x_n\} = \max_{1 \leq i \leq n} \{x_i\}.$$

Check whether $f(x)$ is convex or not.

$$\underline{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n, \quad \text{dom}(f) = \mathbb{R}^n$$

Consider two points $\underline{x}, \underline{y} \in \mathbb{R}^n$.

Consider another point. $z = \theta \underline{x} + (1-\theta) \underline{y}$.

<u>To show</u>: $f(\theta x + (1-\theta) y) \leq \theta f(x) + (1-\theta) f(y)$.

$$\Rightarrow \max_i \{\theta x_i + (1-\theta) y_i\} \leq \theta \max_i \{x_i\} + (1-\theta) \max_i \{y_i\}$$

<u>Proof</u>. Suppose $j_* = \arg\max_i \{\theta x_i + (1-\theta) y_i\}$

(or) $\max_i \{\theta x_i + (1-\theta) y_i\} = \theta x_{j_*} + (1-\theta) y_{j_*}$

$\leq \max_i \{\theta x_i\} + \max_i \{\theta y_i\}$

$= \theta \max_i \{x_i\} + (1-\theta) \max_i \{y_i\}$

$$\boxed{\begin{aligned} x_{j_*} &\leq \max_i \{x_i\} \\ y_{j_*} &\leq \max_i \{y_i\} \end{aligned}}$$

$f(\Sigma \theta_i x_i) \leq \Sigma \theta_i f(x_i)$. This is similar to Jensen's inequality.

$$\boxed{\text{Jensen's inequality}: f(EX) \leq E f(x)}$$

First and Second order Conditions for checking convexity.

① First order condition. (Gradient - based).

f is Convex $\Rightarrow f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle$

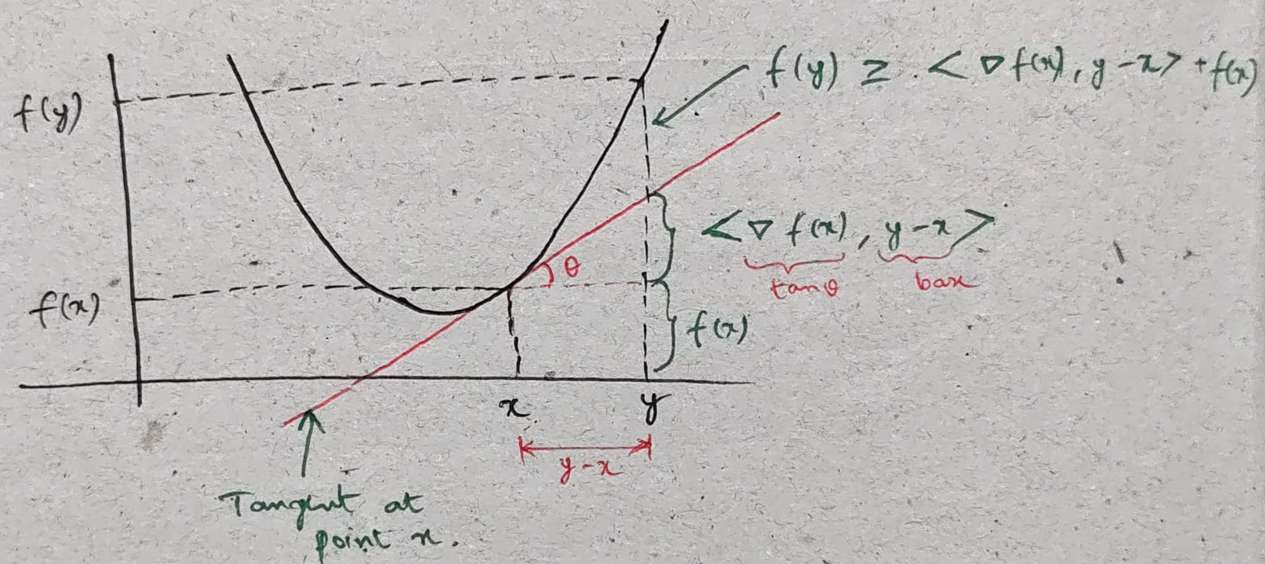
$\forall \; x, y \in \text{dom}(f)$.

Note : $[\nabla f(x)]_i = \underbrace{\dfrac{\partial f}{\partial x_i}}$; $x \in \mathbb{R}^n$.

Partial derivative of f w.r.t x_i. $(i = 1, 2, \dots, n)$.

Example

$n = 1$, $\nabla f(x) = \dfrac{df}{dx} \leftarrow$ Slope of the function at point x.



Note. bound holds $\forall \; y \in \text{dom}(f) \rightarrow$ (Global behaviour of the function f)

but depends only on $\nabla f(x) \rightarrow$ (Local property).

Example. Suppose there exists a point x_0 such that $\nabla f(x_0) = 0$

I st order : $f(y) \geq f(x_0) + \langle \cancel{\nabla f(x_0)}^{\,0}, y - x_0 \rangle$

$\Rightarrow f(y) \geq f(x_0) \; \forall \; y \in \text{dom}(f)$.

$\Rightarrow x_0$ is global minimizer of f.

$\Rightarrow x_0 = \arg\min\limits_{x} f(x)$.

(or) $f(x_0) = \min\limits_{x} f(x)$.

② Second order condition

Hessian Matrix : $[\nabla^2 f(x)]_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}$, $i = 1, \ldots, n$, $j = 1, \ldots, n$.

Example $f(x) = a^T x + b \Rightarrow \nabla^2 f = 0_{n \times n}$

$\Delta \rightarrow$ Delta
$\nabla \rightarrow$ Nabla.

Second order condition for convexity

$$\nabla^2 f(x) \succeq 0 \quad \forall x \in dom(f)$$

Hessian of $f(x)$ has to be Positive Semi definite (PSD).

Second Order Examples

Let's look at several examples of verifying convexity using the Second order condition. (ie) Hessian Matrix has to be Positive Semi Definite (PSD).

(i) Simple Example

$$f(x) = \underbrace{\frac{1}{2} x^T P x + q^T x + r}_{\text{Quadratic form.}} , \text{ where } P \in S^n.$$

$$= \frac{1}{2} \sum_{i,j} P_{ij} x_i x_j + \sum_i q_i x_i + r$$

$$\frac{\partial f}{\partial x_i \partial x_j} = \begin{cases} \frac{P_{ij}}{2} + \frac{P_{ji}}{2} = P_{ij} , & \text{if } i \neq j \\ P_{ii} , & \text{if } i = j \end{cases}$$

$$\Rightarrow \nabla^2 f(x) = P \quad (\text{constant})$$

f is convex $\Leftrightarrow P \succeq 0$
(ie) P is Positive Semi Definite.

(ii) Scalar Example

$f(x) = x \log x$, $x > 0$ (ie) $dom(f) : R_{++}$

$\frac{df}{dx} = \log x + 1$, $\frac{d^2 f}{dx^2} = \frac{1}{x} > 0 \quad \forall x > 0$.

$\Rightarrow f$ is convex.

(iii) Entropy function Example

$$H(x) = -x\log x - (1-x)\log(1-x) \quad , \quad x \in (0,1).$$

$$\frac{d^2 H}{dx^2} = -\frac{1}{x} - \frac{1}{1-x} < 0$$

$\Rightarrow H$ is concave / $-H$ is convex.

(iv) Complicated Example

$$f(x) = \log\left(\sum_{i=1}^{n} e^{x_i}\right).$$

This function is called log-sum-exp function, which is useful in Geometric Programs.

$$z_i = e^{x_i} \quad , \quad f(x) = \log\left(\sum z_i\right)$$

$$\frac{\partial f}{\partial x_i} = \frac{\partial f}{\partial z_i}\frac{\partial z_i}{\partial x_i} = \frac{z_i}{\sum_i z_i}$$

$$e^{x_i} = z_i$$

$$\frac{\partial^2 f}{\partial x_i^2} = \frac{\partial}{\partial z_i}\left(\frac{\partial f}{\partial x_i}\right)\frac{\partial z_i}{\partial x_i}$$

$$= -\frac{z_i^2}{\left(\sum_{k=1}^{n} z_k\right)^2} + \frac{z_i}{\left(\sum_{k=1}^{n} z_k\right)} \quad , \quad i=j$$

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial f}{\partial z_j}\left(\frac{\partial f}{\partial x_i}\right)\left(\frac{\partial z_j}{\partial x_j}\right) = \frac{-z_i z_j}{\left(\sum z_k\right)^2} \quad , \quad i \neq j.$$

How to verify $\nabla^2 f(x) \succeq 0 \;\; \forall x$?

- Cannot calculate Eigen values.

Recall : A is PSD $\Leftrightarrow u^T A u \geq 0 \;\; \forall u \in \mathbb{R}^n$

$$u^T \nabla^2 f(x) u = \sum_{i,j} \left[\nabla^2 f(x)\right]_{ij} u_i u_j$$

$$= \underbrace{\sum_{i \neq j}\left(\frac{\partial^2 f}{\partial x_i \partial x_j}\right) u_i u_j}_{\text{off diagonal}} + \underbrace{\sum_{i=1}^{n} u_i^2 \frac{\partial^2 f}{\partial x_i^2}}_{\text{diagonal}}$$

$$= -\sum_{i \neq j} \frac{u_i u_j z_i z_j}{\left(\sum_{k=1}^{n} z_k\right)^2} + \sum_{i=1}^{n} \frac{u_i^2 z_i}{\left(\sum_{k=1}^{n} z_k\right)} - \sum_{i=1}^{n} \frac{u_i^2 z_i^2}{\left(\sum_{k=1}^{n} z_k\right)^2}$$

$$= -\frac{1}{\left(\sum_{k=1}^{n} z_k\right)^2} \left[\sum_{i,j} u_i u_j z_i z_j - \left(\sum_{k=1}^{n} z_k\right)\left(\sum_{i=1}^{n} u_i^2 z_i\right) \right].$$

Now, we use Cauchy-Schwarz Inequality.

$$(\underline{a}^T \underline{b})^2 \leq (\underline{a}^T \underline{a})(\underline{b}^T \underline{b})$$

$$(a^T b)^2 = \left(\sum_i a_i b_i\right)^2 = \sum_{i,j} a_i a_j b_i b_j = \left(\sum_{i=1}^{n} u_i z_i\right)^2$$

$$\Rightarrow \underbrace{a_i b_i = u_i z_i}$$

$a_i^2 = z_i$ | $b_i^2 = u_i^2 z_i$

$a_i = \sqrt{z_i}$ | $b_i = u_i \sqrt{z_i}$

$\Rightarrow$ Cauchy-Schwarz can be applied.

$\Rightarrow \nabla^2 f(x) \succeq 0 \Rightarrow$ PSD $\Rightarrow f$ is convex.

$\Rightarrow u^T \nabla^2 f(x) u \geq 0 \quad \forall u$

Key points.
(a) Calculate $u^T \nabla^2 f(x) u$
(b) Use Cauchy Schwarz inequality.

Exercise : Show that $\underbrace{g(x) = \left(\prod_{i=1}^{n} x_i\right)^{\frac{1}{n}}}_{\text{Geometric Mean}}$ is concave.

Convex Sets and Functions

Lets look at the relationship b/w Convex Sets and Functions.

(i) We'll see two notions of convexity. One pertaining to Sets, another one pertaining to functions. And we'll see how they're closely related to each other.

Let us consider a function, $f : \mathbb{R}^n \rightarrow R$.

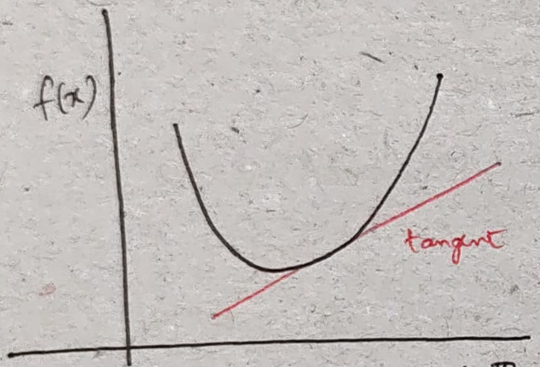
We define Epi graph of the function f.

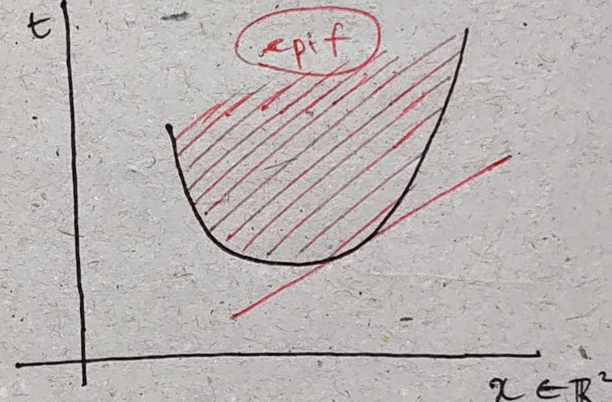
$$\text{epi } f = \left\{ \begin{bmatrix} x \\ t \end{bmatrix} \in \mathbb{R}^{n+1} \;\middle|\; f(x) \leq t \right\} \quad \forall \; x \in \text{dom}(f).$$

epi ≡ above

Example: $f(x) = \|x\|$, then

$$\text{epi } f : \left\{ \begin{bmatrix} x \\ t \end{bmatrix} \in \mathbb{R}^{n+1} \;\middle|\; \|x\| \leq t \right\}$$





Result :

f is Convex function $\Longleftrightarrow$ epi f is convex set.

Let us compare the First order condition for convexity (vs) supporting hyperplane theorem.

- I^{st} order condition.

  The Tangent is always below the function

- Supporting hyperplane.

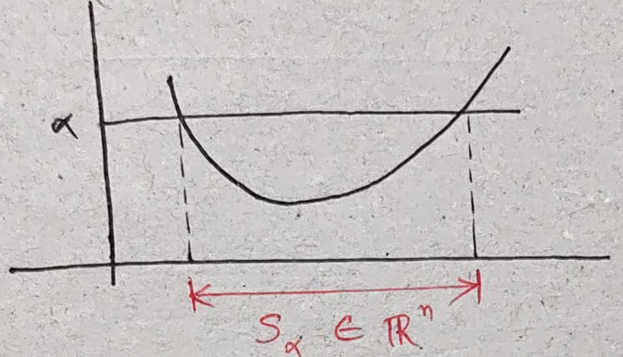
  There exists hyperplane such that it divides the space into two half spaces. one of those Half space contains the entire convex set.

Quasi-Convex functions

Quasi-convex functions are generalization of convex functions, in the sense that they're defined using convex sets. And they include more functions than convex functions.

Given α : define $S_\alpha = \{\underline{x} \mid f(\underline{x}) \leq \alpha\} \subseteq \mathbb{R}^n$

In contrast with epi f, this is a sublevel set.



Result : $f(\underline{x})$ is convex $\Rightarrow$ S_α is convex, when the set is non-empty.

<u>Proof</u> : S_α : $x, y \in S_\alpha$

$x \in S_\alpha \Rightarrow f(x) \leq \alpha$

$y \in S_\alpha \Rightarrow f(y) \leq \alpha$

$f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta) f(y)$ Zeroth-order

$\leq \theta\alpha + (1-\theta)\alpha$

$= \alpha$

$\Rightarrow \theta x + (1-\theta)y \in S_\alpha$

$\Rightarrow S_\alpha$ is convex.

Converse is not true.

Possible that f is non-convex but S_α is still convex set.

<u>Example</u>. $f(x) = \log(x)$

f is concave, not convex function.

$$\frac{d^2 f}{dx^2} = -\frac{1}{x^2} < 0, \quad x > 0.$$

$$S_\alpha = \{x \mid \log(x) \leq \alpha\} = \{x > 0 \mid x < e^\alpha\}$$

(or) $\{0 < x \leq e^\alpha\}$

This is an interval in one dimension. So, it is obviously convex. The original function $\log(x)$ is not convex.

This is an example of non-convex function for which the sublevel set is convex. This kind of functions is called as Quasi-Convex functions.

<u>Quasi-Convex functions</u> : This is a function f such that S_α (sub-level set) is convex set. Obviously it includes convex functions.

<u>Some definitions</u>

$$\underbrace{S^\alpha}_{\text{Super level set.}} = \{x \mid f(x) \geq \alpha\}$$

Function f is Quasi concave, such that S^α is convex set

<u>Generalize convexity idea</u>

Monotonic functions are both quasi-convex and quasi-concave.

<u>Operations that preserve Convexity</u>

Let us look at various operations that preserve convexity in the context of convex function.

1. f_i is convex $\forall\ i = 1, 2, \ldots, n$

   $\Rightarrow \sum w_i f_i(x)$ also convex function, for $w_i \geq 0$.

   $\Rightarrow$ Can be verified using Hessian.

2. f is convex

   $\Rightarrow f(Ax + b)$ is convex when $Ax + b \in \operatorname{dom}(f)$

   $\Rightarrow$ Can be verified using Zeroth / First order.

3. Pointwise Maximum.

   Let $\{f_i(\underline{x})\}_{i=1}^{n}$ is convex,

   $\Rightarrow g(x) = \max\limits_{1 \leq i \leq n} \{f_i(\underline{x})\}$ is also convex function.

<u>Proof</u> : Zeroth order :

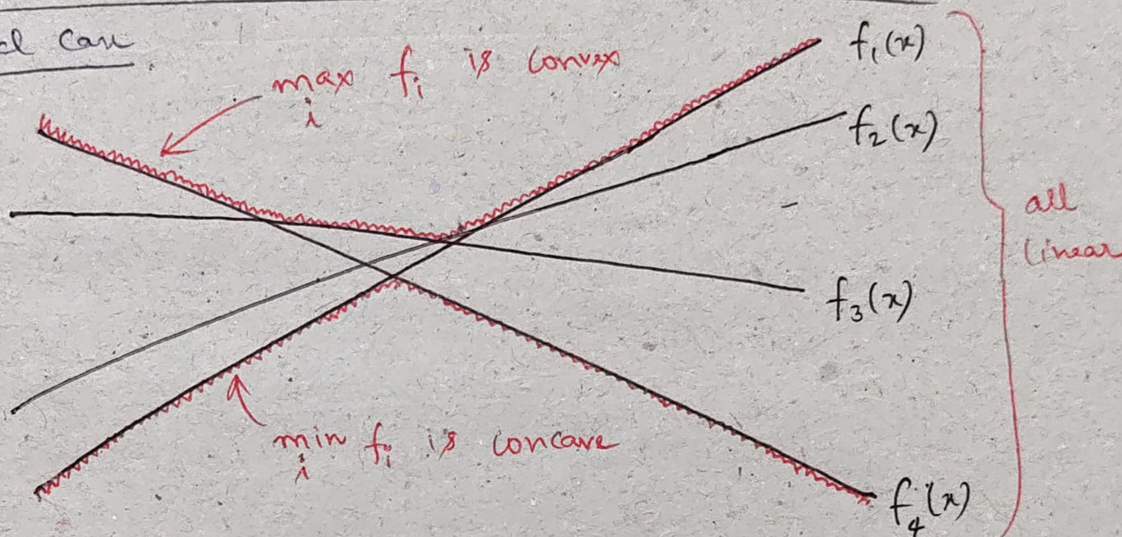
$$\underline{x}, \underline{y} \in \operatorname{dom}(g) = \bigcap_{i=1}^{n} \operatorname{dom}(f_i)$$

$$g(\theta x + (1-\theta) y) = \max_i f_i(\theta x + (1-\theta) y)$$
$$\leq \max_i \theta f_i(x) + (1-\theta) f_i(y)$$
$$\leq \theta \underbrace{\max_i f(x)}_{g(x)} + (1-\theta) \underbrace{\max_i f_i(y)}_{g(y)}$$

$$\boxed{\text{Recall } \max_i \{a_i + b_i\} \leq \max_i \{a_i\} + \max_i \{b_i\}.}$$

$$= \theta g(x) + (1-\theta) g(y).$$

Special case



Since f_i are linear / affine,

$$f_i(x) = a_i^T x + b_i.$$

$\max_i \{a_i^T x + b_i\}$ is convex

$\min_i \{a_i^T x + b_i\}$ is concave

Support of $S_C(x) = \max_{y \in C} \{x^T y\}$ ← Maximum of set C can be arbitrary. Need not be convex

Eg. $C = \{a \mid \|a\| \leq 1\}$

Example.

$S_{B(0,1)}(x) = \max_{\|y\| \leq 1} x^T y$ is also convex.

Composition Rules

Let us look at the composition Rules, which are useful for determining whether a function is convex or not.

Let us first look at the Scalar case.

Scalar: $f(x) = \underbrace{h(g(x))}$ (or) $\underbrace{h \circ g(x)}$

Composition of two functions h and g.

$$f : \mathbb{R} \to \mathbb{R}$$

How to find whether f is convex or not?
By taking Partial Derivative...

$$\frac{d^2 f}{dx^2} = \underbrace{h''(g(x))\,(g'(x))^2}_{\text{First term}} + \underbrace{g''(x)\, h'(g(x))}_{\text{Second term}} \leftarrow \text{By applying Chain Rule.}$$

Finding sufficient condition: When both terms are ≥ 0,

① $h''(\cdot) \geq 0 \Rightarrow h$ is convex.

② $g''(\cdot) \geq 0$, $h'(\cdot) \geq 0 \Rightarrow g$ is convex, h is non-decreasing

(or)

$g''(\cdot) \leq 0$, $h'(\cdot) \leq 0 \Rightarrow g$ is concave, h is non-increasing.

Summary:

$$\left.\begin{array}{l} \text{(i) } g \text{ is convex, } h \text{ is convex, non-decreasing.} \\ \quad\text{(or)} \\ \text{(ii) } g \text{ is concave, } h \text{ is convex, non-increasing} \end{array}\right\} h \circ g(x) \text{ is convex.}$$

Result is also true when f is non-differentiable.

Examples. ① When $e^{g(x)}$ is convex?

$\underbrace{h(y) = e^y}$ ← convex, non-decreasing

$\Rightarrow g(x)$ need this to be convex.

$\Rightarrow g(x)$ is convex

$\Rightarrow e^{g(x)}$ is also convex.

② When $-\log(g(x))$ is convex ?

$$\underbrace{h(y) = -\log(g(x))}$$ convex, non-increasing

$\Rightarrow g(x)$ is concave

$\Rightarrow -\log(g(x))$ is convex.

Fail case :

$h(x) = x$, $dom(h) = [1, 2]$ ← Restricted

$g(x) = x^2$, $dom(g) = \mathbb{R}$.

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

$dom(f) : \{x \mid g(x) \in [1,2]\}$, $x \in \mathbb{R}$

$x^2 \in [1,2]$.

Rule : h is convex, non-increasing in $[1, 2]$,

g is convex

but,

$$dom(f) = \underbrace{[-\sqrt{2}, -1]} \cup \underbrace{[1, \sqrt{2}]}$$

Disjoint intervals

$\Rightarrow dom(f)$ is not convex

$\Rightarrow f$ is definitely not convex.

Composition Rule says that f is convex, but the contradiction is due to domain restrictions.

Fix : Allow $f(x) \in \underbrace{\mathbb{R} \cup \{-\infty, \infty\}}_{\text{Extended Real Line}}$,

where,

$\mathbb{R}$: Real Line. $\{-\infty < x < \infty\}$

$\mathbb{R} \cup \{-\infty, \infty\}$: Extended Real Line $\{-\infty \le x \le \infty\}$

Extended function :

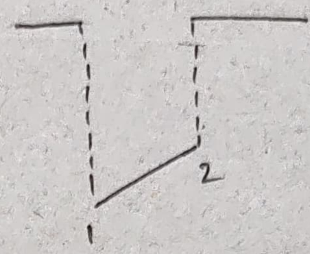
Convex : $\tilde{f}(x) = \begin{cases} f(x), & x \in dom(f) \\ \infty, & x \notin dom(f) \end{cases}$

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

Revised Rule : (Replacing h, g with $\tilde{h}, \tilde{g}$)

$$\left.\begin{array}{l}\tilde{g} \text{ is convex}, \tilde{h} \text{ is convex, non-decreasing} \\ \tilde{g} \text{ is concave}, \tilde{h} \text{ is convex, non-increasing}\end{array}\right\} \Rightarrow f \text{ is convex.}$$

$$\Rightarrow \tilde{h}(x) = \begin{cases} x, & x \in [1,2] \\ \infty, & x \notin [1,2] \end{cases}$$



$\Rightarrow \tilde{h}$ is not a non-decreasing function.

Vector Composition Rule

$$f(x) = h\big(g_1(x), g_2(x), \ldots, g_m(x)\big),$$

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

$g_i : \mathbb{R}^n \to \mathbb{R}$

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

$h(\underline{g}(x)) \to$ Vector valued function.

Composition Rule :

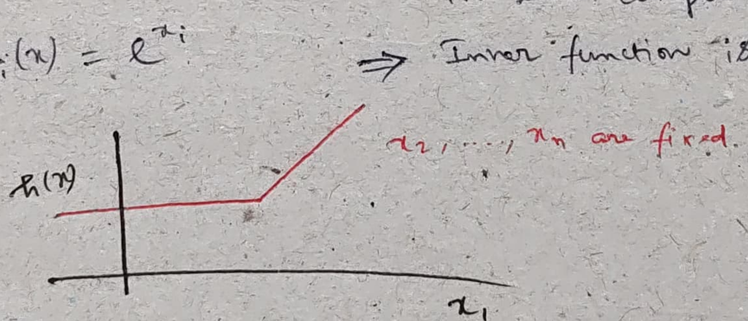
$$\left.\begin{array}{l}\tilde{g}_i \text{ is convex}, \tilde{h} \text{ is convex, non-decreasing in each component} \\ \tilde{g}_i \text{ is concave}, \tilde{h} \text{ is convex, non-increasing in each component}\end{array}\right\} =$$

$= \Rightarrow f$ is convex function.

Example.

$h(x) = \max\{x_i\} \Rightarrow$ Outer function is convex, non-decreasing in each component.

$g_i(x) = e^{x_i} \Rightarrow$ Inner function is convex.



$\Rightarrow h(x)$ is constant, increasing and non decreasing.

$\Rightarrow \max_i e^{x_i}$ is convex.

Perspective Function and Minimization

Perspective Function Transform is given by

$$g\left(\begin{bmatrix} x \\ t \end{bmatrix}\right) = t\, f(x/t), \quad t > 0.$$

Result is that,

$$f \text{ is convex} \Rightarrow g \text{ is convex.}$$

Note that,

$$f : \mathbb{R}^n \to \mathbb{R}, \quad g : \mathbb{R}^{n+1} \to \mathbb{R}.$$

Examples.

① $f(x) = x^T x$

Hessian: $\nabla^2 f(x) = 2I \geq 0 \Rightarrow$ Positive Definite

$$g(x) = t\left(\frac{x}{t}\right)^T\left(\frac{x}{t}\right) = \frac{x^T x}{t}, \quad t > 0 \Rightarrow \text{Convex.}$$

---

② $f(x) = -\log(x)$.

which is obviously convex, coz log is concave.

$$g(x,t) = -t \log(x/t)$$

$$= -t \log x + t \log t$$

This is also convex.

Minimization.

Consider a function, $f : \underbrace{\mathbb{R}^n}_{x} \times \underbrace{\mathbb{R}^m}_{y} \to \mathbb{R}$ which takes two inputs and outputs a Real value.

Requirements.

① $f(x, y)$ is convex w.r.t $\begin{bmatrix} x \\ y \end{bmatrix}$

(or) $f(x, y)$ is convex jointly in (x, y).

② C is a convex set $\subseteq \mathbb{R}^m$,

$$g(x) = \min_{y \in C} f(x, y) \text{ is convex.}$$

This is in contrast with pointwise maximization result.

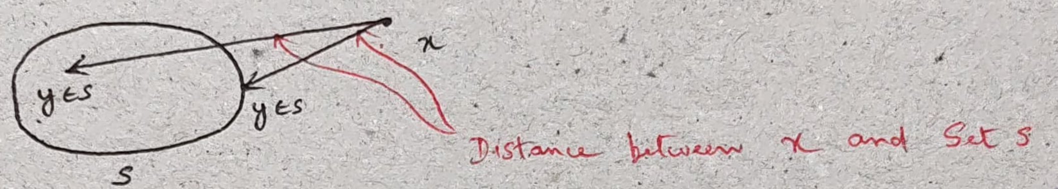
Example: $f(\underline{x}, \underline{y}) = \|\underline{x} - \underline{y}\|$, $x, y \in \mathbb{R}^n$

$$= \left\| \begin{bmatrix} I & -I \end{bmatrix} \begin{bmatrix} \underline{x} \\ \underline{y} \end{bmatrix} \right\|$$

$\Rightarrow$ This is convex.

$$g(\underline{x}) = \min_{y \in S} \|\underline{x} - y\| = \text{dist}(x, S).$$

where $S \rightarrow$ convex set.



Find the value of y such that this distance is minimum

$\therefore$ The distance function is a convex function.
