RECAP

- Convex sets
  - Affine sets — Affine combinations, Affine hull
  - Cones — Conic — " — Conic — " —
  - Convex sets — Convex — " — Convex — " —
- Hyperplanes & halfspaces
  - Separating hyperplane
  - Supporting hyperplane
- Proper cone
  - generalized inequality
  - Dual

Convex Functions

Definition: $f: \mathbb{R}^n \to \mathbb{R}$ is convex if $\text{dom}(f)$ is convex

$$f(\theta \underline{x}_1 + (1-\theta)\underline{x}_2) \leq \theta f(\underline{x}_1) + (1-\theta) f(\underline{x}_2)$$

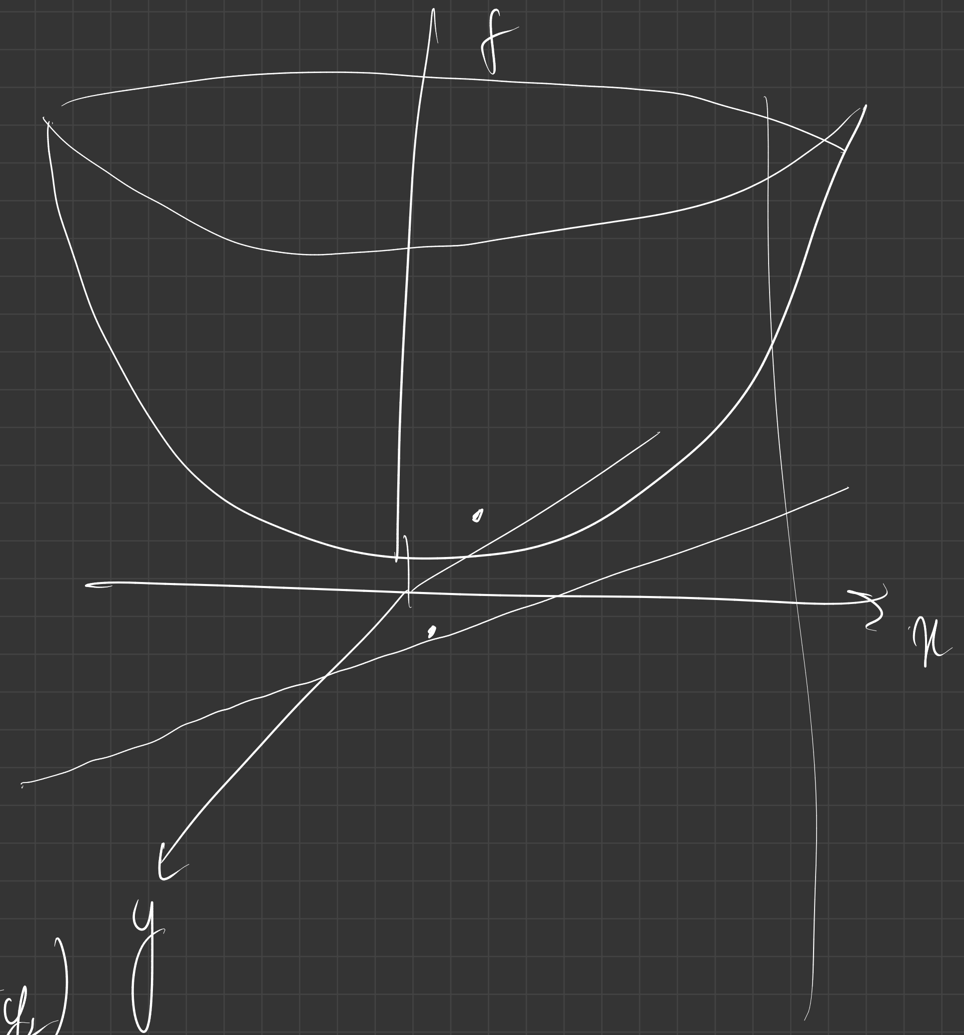
$$\forall \underline{x}_1, \underline{x}_2 \in \text{dom}(f)$$

$$0 \leq \theta \leq 1$$

Define $g: \mathbb{R} \to \mathbb{R}$

$$g(t) = f(\underline{x}_1 + t\underline{x}_2)$$ for some $\underline{x}_1, \underline{x}_2 \in \text{dom}(f)$

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



$$g(t) = f(x_1 + t x_2,\ y_1 + t y_2)$$

$$= (x_1 + t x_2)^2 + (y_1 + t y_2)^2$$

$$g''(t) = 2(x_2^2 + y_2^2) \geq 0$$

Lemma: f is convex iff

$$g(t) = f(\underline{x}_1 + t\,\underline{x}_2) \text{ is convex}$$

for all $\underline{x}_1, \underline{x}_2$.

Proof: Suppose f is convex.

Consider t_1, t_2, $0 \le \theta \le 1$

$$g(\theta t_1 + (1-\theta) t_2)$$

$$= f\left(\underline{x}_1 + (\theta t_1 + (1-\theta) t_2)\, \underline{x}_2 \right)$$

$$= f\left(\theta x_1 + (1-\theta) x_1 + \theta t_1 x_2 + (1-\theta) t_2 x_2 \right)$$

$$= f\left(\theta (x_1 + t_1 x_2) + (1-\theta)(x_1 + t_2 x_2) \right)$$

$$\leq \theta f(x_1 + t_1 x_2) + (1-\theta) f(x_1 + t_2 x_2)$$

$$= \theta g(t_1) + (1-\theta) g(t_2)$$

ST if $g(t)$ is convex for all x_1, x_2, then

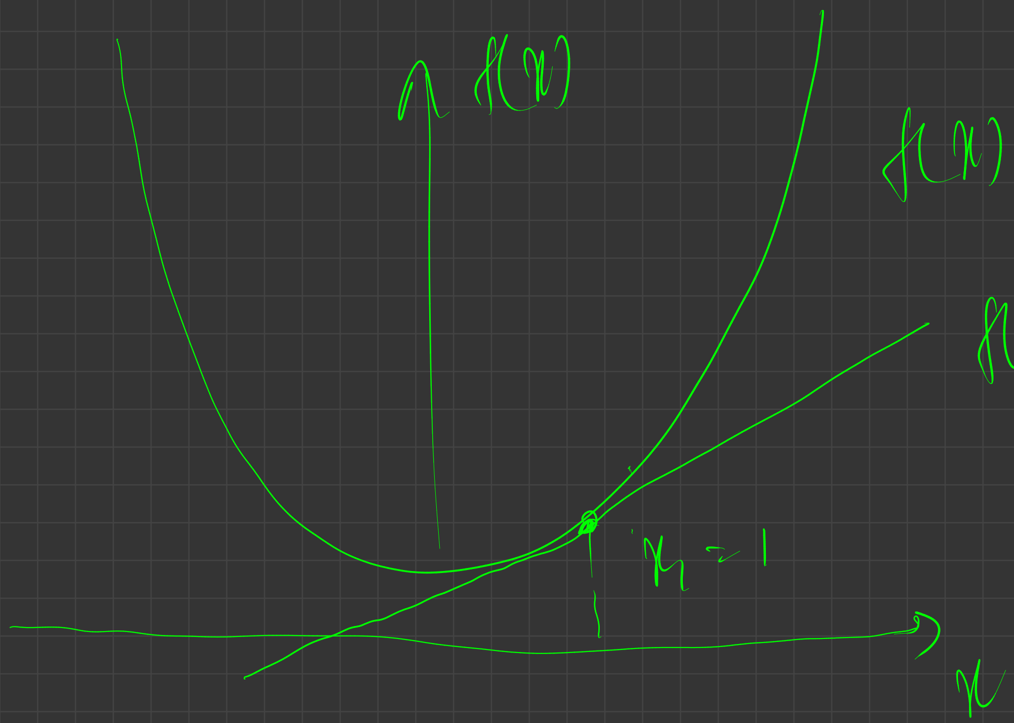
f is convex.

First derivative

Suppose ∇f exists.

f is convex iff

$$f(x_2) \geq \underbrace{f(x_1) + \nabla f(x_1)^T (x_2 - x_1)}_{\text{Tangent}}$$



$f(x) = x^2$

$f(x_1) + \nabla f(x_1)^T (x_2 - x_1)$

$1 + 2(x - 1)$

$= 2x - 1$

$$f(\theta \underline{x_2} + (1-\theta)\underline{x_1}) \leq \theta f(\underline{x_2}) + (1-\theta) f(x_1)$$

$$f(\underline{x_2}) \geq \frac{-(1-\theta) f(x_1) + f(\theta x_2 + (1-\theta)x_1)}{\theta}$$

$$= f(\underline{x_1}) + \frac{f(\theta x_2 + (1-\theta)x_1) - f(x_1)}{\theta}$$

$$= f(x_1) + \frac{f(x_1 + \theta(x_2 - x_1)) - f(x_1)}{\theta}$$

$$\xrightarrow[\theta \to 0]{} f(\underline{x_1}) + (\nabla f(x_1))^T (\underline{x_2} - \underline{x_1})$$

Suppose $f(x_2) \geq f(x_1) + \nabla f(x_1)^T (x_2 - x_1)$

$\forall\, x_1, x_2$

$z = \theta x_1 + (1-\theta) x_2 \quad , \quad x_1, x_2$

$f(x_1) \geq f(z) + \nabla f(z)^T (z - x_1) \quad \times \theta$

$f(x_2) \geq f(z) + \nabla f(z)^T (z - x_2) \quad \times (1-\theta)$

$$\theta f(x_1) + (1-\theta) f(x_2) \geq \theta f(z) + (1-\theta) f(z)$$

$$+ \underbrace{\nabla f(z)^T \Big[\overbrace{\theta z + (1-\theta) z}^{z} \; \overbrace{- \theta x_1 - (1-\theta) x_2}^{z} \Big]}_{0}$$

$\therefore \quad \theta f(x_1) + (1-\theta) f(x_2) \geq f(z)$

$$= f(\theta x_1 + (1-\theta) x_2)$$

Second derivative test

Suppose $\nabla^2 f$ exists.

f is convex iff $\nabla^2 f$ is PSD $\forall \underline{x}$.

Suppose $g(t) = f(\underline{x}_1 + t\,\underline{x}_2)$

$$g'(t) = \lim_{\delta \to 0} \frac{g(t+\delta) - g(t)}{\delta}$$

$$= \lim_{\delta \to 0} \frac{f(x_1 + (t+\delta)x_2) - f(x_1 + t x_2)}{\delta}$$

$$= \lim_{\delta \to 0} \frac{f(x_1 + t x_2 + \delta x_2) - f(x_1 + t x_2)}{\delta}$$

$$= \left(\nabla f(x_1 + t x_2)\right)^T \underline{x}_2 = \underline{x}_2^T \nabla f(x_1 + t x_2)$$

$$g''(t) = \lim_{\delta \to 0} \frac{g'(t+\delta) - g'(t)}{\delta}$$

$$= \lim_{\delta \to 0} \underline{x}_2^T \left[\frac{\nabla f(x_1 + (t+\delta)x_2) - \nabla f(x_1 + tx_2)}{\delta} \right]$$

$$= \lim_{\delta \to 0} \underline{x}_2^T \left[\frac{\nabla f(\underline{x}_1 + t\underline{x}_2 + \delta x_2) - \nabla f(x_1 + t\underline{x}_2)}{\delta} \right]$$

$$= \underline{x}_2^T \left[\nabla^2 f(x_1 + t x_2)^T \underline{x}_2 \right]$$

$$g''(t) = \underline{x}_2^T \nabla^2 f(x_1 + t x_2) \underline{x}_2$$

If $\nabla^2 f$ is PSD, then

$$x_2^T \nabla^2 f(x_1 + t x_2)\, x_2 \geq 0$$

$$\forall\, x_1, x_2, t$$

$\Rightarrow\ g''(t) \geq 0 \quad \forall\ t, x_1, x_2$

$\Rightarrow\ g$ is convex $\forall\ x_1, x_2$

$\Rightarrow\ f$ is convex

If f is convex, then g is convex $\forall\ x_1, x_2$

$\Rightarrow\ g''(t) \geq 0 \quad \forall\ t, x_1, x_2$

$\Rightarrow\ \nabla^2 f$ is PSD.

- Linear functions:

$$f(\underline{x}) = \underline{a}^T \underline{x} + \underline{b}$$

$$\nabla^2 f(\underline{x}) = 0_{n \times n} \qquad \text{PSD}$$

- $$f(\underline{x}) = \underline{x}^T A \underline{x} + \underline{b}$$

$$\nabla f(\underline{x}) = (A + A^T)\underline{x}$$

$$\nabla^2 f(\underline{x}) = A + A^T$$

$\longrightarrow$ When is this PSD?

$$\underline{x}^T (A + A^T) \underline{x} = 2\underline{x}^T A \underline{x}$$

$A + A^T$ is PSD iff A is PSD.

(b) $f(\underline{x}) = e^{\underline{a}^T \underline{x}} = e^{\sum_{j=1}^{n} a_j x_j} = \prod_{j=1}^{n} e^{a_j x_j}$

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = f(\underline{x})\, a_i a_j$$

$$\frac{\partial f}{\partial x_i} = \left(\prod_{j \neq i} e^{a_j x_j}\right) \frac{\partial}{\partial x_i}\left(e^{a_i x_i}\right)$$

$$= \left(\prod_{j \neq i} e^{a_j x_j}\right) a_i e^{a_i x_i}$$

$$= a_i f(\underline{x})$$

$$\left(\nabla^2 f\right)_{i,j} = a_i a_j f(\underline{x})$$

$$\nabla^2 f(\underline{x}) = \underbrace{\underline{a}\,\underline{a}^T f(\underline{x})}$$

Is this PSD?

$$y^T \left(a\, a^T f(x)\right) y$$

$$(y^T a)^2 \; f(x)$$

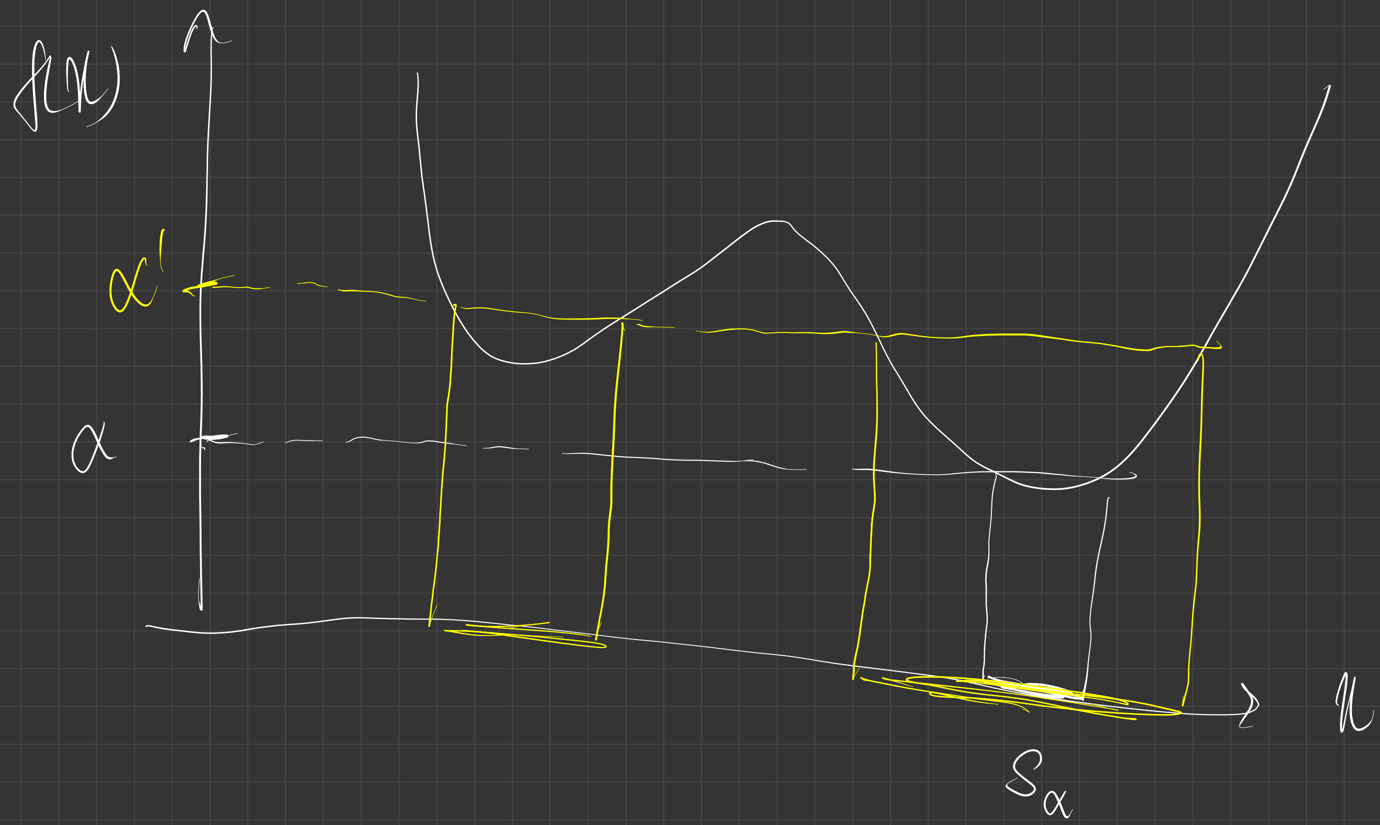
$$\geq 0 \qquad \geq 0$$

⑥ $f(\underline{x}) = \|\underline{x}\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p}$ — HW

Level set

$$S_\alpha = \{ x \in \mathbb{R}^n : f(x) \le \alpha \}$$

α-level set of f

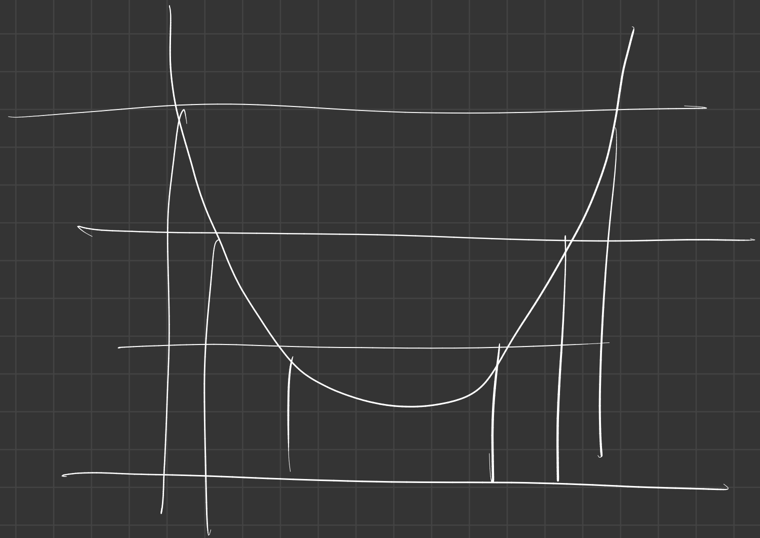


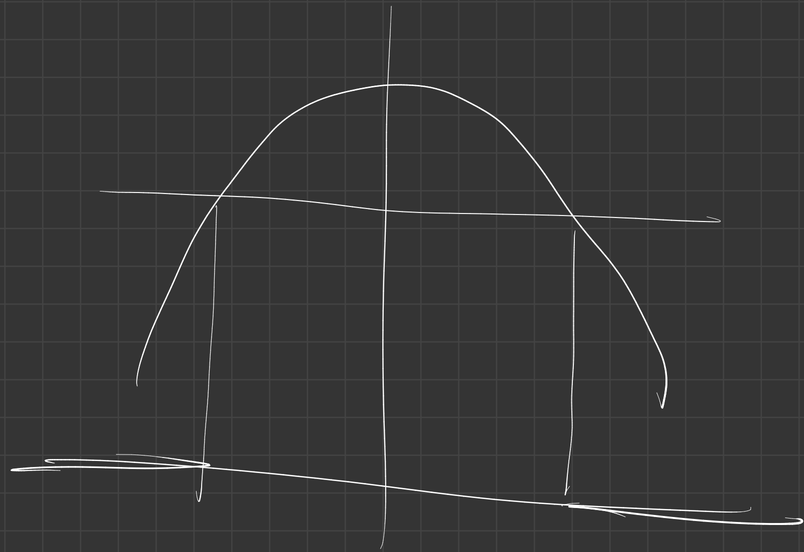
<u>Claim</u> : If f is convex, then the level sets are convex for all α

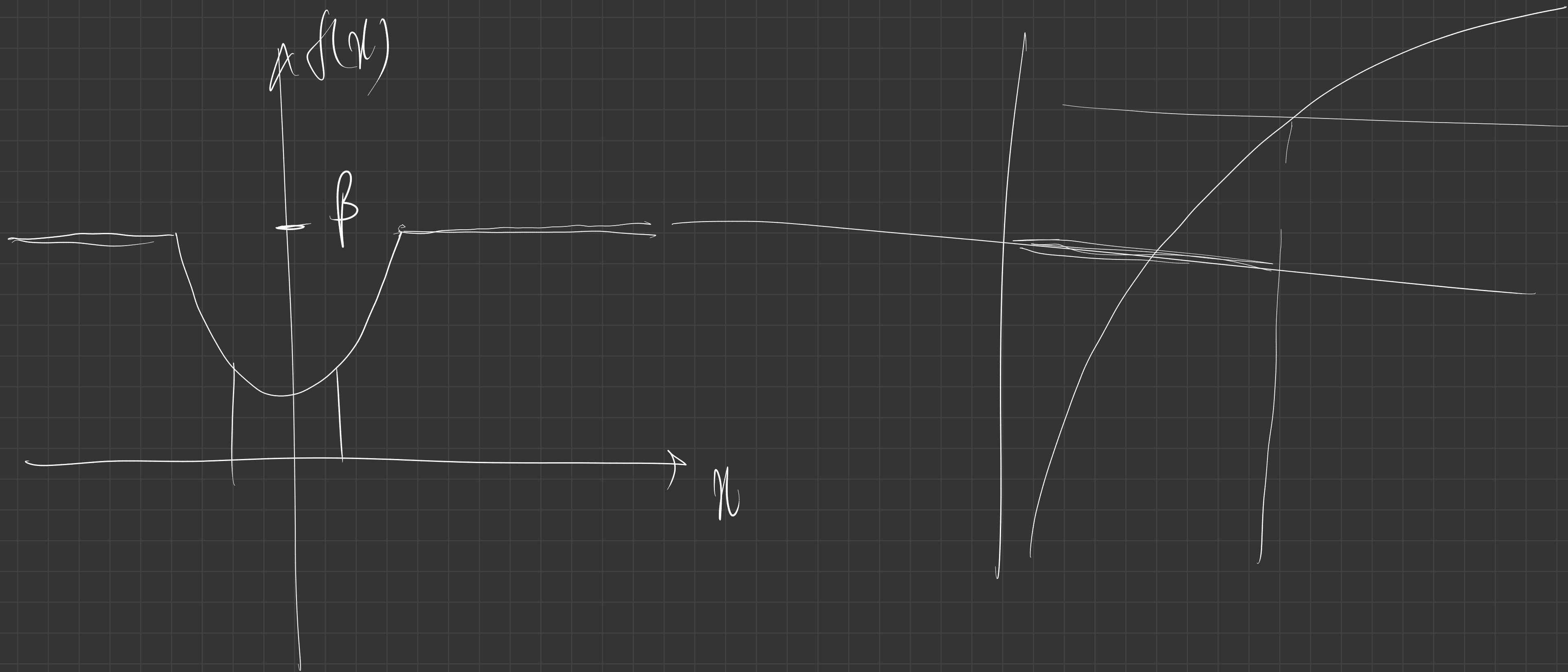
<u>Q</u> : If S_α is convex for all $\alpha \in \mathbb{R}$, then does it imply that f is convex?

$$S_\alpha = \{ \underline{x} : f(\underline{x}) \leq \alpha \}$$

<u>NO</u>






f(x)
β
x

If f is convex, then level sets are convex

$$S_\alpha = \{ \underline{x} : f(\underline{x}) \le \alpha \}$$

$\underline{x}_1, \underline{x}_2$

$$\left.\begin{aligned} f(\underline{x}_1) &\le \alpha \\ f(\underline{x}_2) &\le \alpha \end{aligned}\right\} \text{ since } \underline{x}_1, \underline{x}_2 \in S_\alpha$$

$$f(\theta \underline{x}_1 + (1-\theta) \underline{x}_2)$$

$$\le \theta f(\underline{x}_1) + (1-\theta) f(\underline{x}_2) \qquad \text{since } f \text{ is convex}$$

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

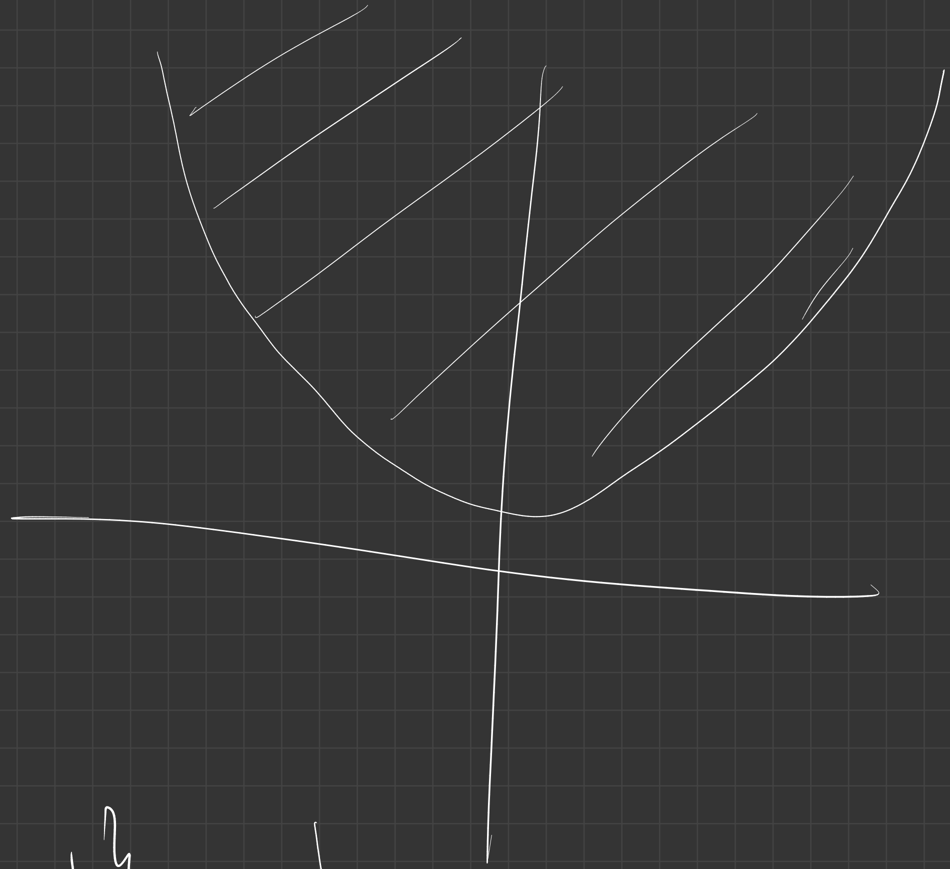
$$= \alpha \implies \theta \underline{x}_1 + (1-\theta) \underline{x}_2 \in S_\alpha$$

$$C = \left\{ \underset{\sim}{x} : \begin{array}{l} f_1(\underset{\sim}{x}) \le \alpha_1 \\ f_2(\underset{\sim}{x}) \le \alpha_2 \\ \vdots \\ f_k(\underset{\sim}{x}) \le \alpha_k \end{array} \right\}$$

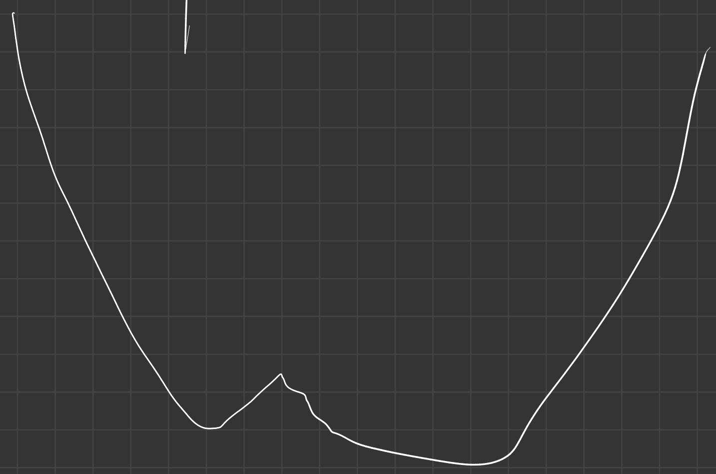
If $f_1 \cdots f_k$ are convex, then so is C.

Epigraph of f

Given f,



$$\mathrm{epi}(f) = \{ (\underline{x}, t) : t \geq f(\underline{x}) \}$$



Claim: epi(f) is convex iff f is convex

Proof: Suppose f is convex

(x_1, t_1), $(x_2, t_2) \in \mathrm{epi}(f)$

$\theta(x_1, t_1) + (1-\theta)(x_2, t_2)$

"

$(\theta x_1 + (1-\theta x_2), \theta t_1 + (1-\theta) t_2) \in \mathrm{epi}(f)$

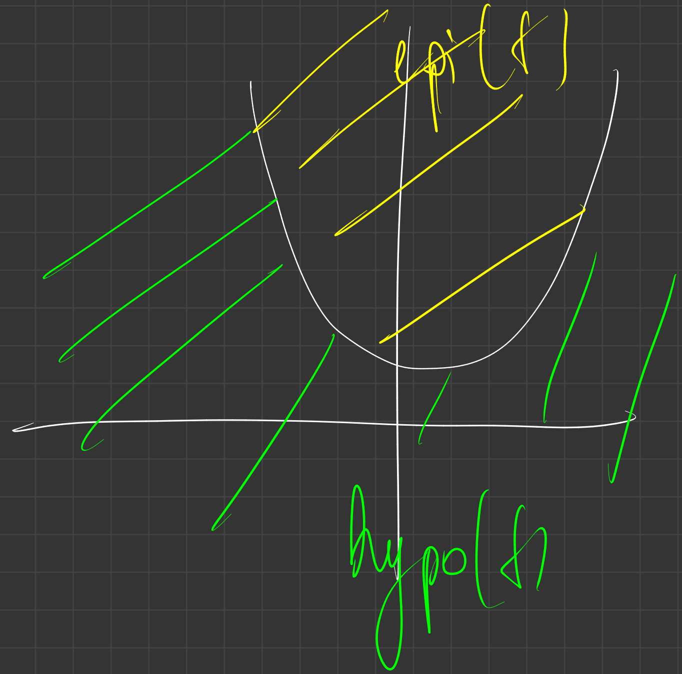
$$f(\theta x_1 + (1-\theta) x_2) \le \theta f(x_1) + (1-\theta) f(x_2)$$

$$\le \theta t_1 + (1-\theta) t_2$$

Hypograph

$$\text{hypo}(f) = \{(\underline{x}, t) : t \leq f(\underline{x})\}$$

f is concave iff $\text{hypo}(f)$ is convex.



$$f(\theta x_1 + (1-\theta)x_2) \le \theta f(x_1) + (1-\theta) f(x_2)$$

<u>Claim</u>: $$f(\overbrace{\theta_1 x_1 + \theta_2 x_2 - + \theta_k x_k}^{\mathbb{E}X}) \le \underbrace{\sum_{i=1}^{k} \theta_i f(x_i)}_{\mathbb{E}f(X)}$$

$$\theta_1, \theta_2 \sim \theta_k \geq 0$$

$$\sum_{i=1}^{k} \theta_i = 1$$

<u>Proof</u>: Induction. True for $k=2$

Assume true for $k-1$

Given $f(\theta_1 x_1 + \theta_2 x_2 + \cdots + \theta_{k-1} x_{k-1}) \le \theta_1 f(x_1) + \cdots + \theta_{k-1} f(x_{k-1})$ — (*)

$$f(\overbrace{\theta_1 x_1 + \theta_2 x_2 + \cdots + \theta_{k-1} x_{k-1}} + \theta_k x_k)$$

$$= f\left(\underbrace{\left(\frac{\theta_1 x_1 + \theta_2 x_2 \cdots + \theta_{k-1} x_{k-1}}{1-\theta_k}\right)}_{x_1'}(1-\theta_k) + \theta_k \underbrace{x_k}_{x_2'}\right)$$

$$\le (1-\theta_k) f\left(\frac{\theta_1}{1-\theta_k} x_1 + \frac{\theta_2}{1-\theta_k} x_2 + \cdots + \frac{\theta_{k-1}}{1-\theta_k} x_{k-1}\right) + \theta_k f(x_k)$$

Use $(*)$

$$\leq (1-\theta_k)\left(\frac{\theta_1}{1-\theta_k} f(x_1) + \frac{\theta_2}{1-\theta_k} f(x_2) + \cdots + \frac{\theta_{k-1}}{1-\theta_k} f(x_{k-1})\right)$$

$$+ \theta_k f(x_k)$$

$$= \sum_{i=1}^{k} \theta_k f(x_k)$$

Jensen's inequality: If f is convex, then

$$f(\mathbb{E}X) \leq \mathbb{E}f(X)$$ for all distribution on $\mathbb{R}$

Eg 1

$a, b \in \mathbb{R}_{\geq 0}$

$0 \leq \theta \leq 1$

$$\sqrt{ab} \leq \frac{a+b}{2} \qquad \Big| \qquad a^{\theta} b^{(1-\theta)} \leq \theta a + (1-\theta) b$$

$$\Updownarrow$$

$$\frac{1}{2}\log(ab) \leq \log\left(\frac{a+b}{2}\right)$$

$$-\log\left(\frac{a+b}{2}\right) \leq \frac{-\log(a)}{2} + \left(\frac{-\log b}{2}\right)$$

$$f\left(\frac{a+b}{2}\right) \leq \frac{1}{2} f(a) + \frac{1}{2} f(b)$$

Hölder's inequality: $\underline{x}, y \in \mathbb{R}^n$

$p, q > 0$, $\frac{1}{p} + \frac{1}{q} = 1$

$p > 1$

Then,
$$\sum_{i=1}^{n} |x_i y_i| \le \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p} \left(\sum_{i=1}^{n} |y_i|^q \right)^{1/q}$$

$$a^{1/p} b^{1/q} \le \frac{1}{p} a + \frac{1}{q} b$$

$$\sum_{i=1}^{n} |x_i y_i| \leq \sqrt{\sum_{i=1}^{n} x_i^2} \sqrt{\sum_{i=1}^{n} y_i^2}$$

$$a = \frac{|x_i|^2}{\sum_{i=1}^{n} x_i^2} \qquad b = \frac{|y_i|^2}{\sum_{i=1}^{n} y_i^2}$$

$$= \frac{x_i^2}{\|x\|^2} \qquad = \frac{y_i^2}{\|y\|^2}$$

Use AM-GM inequality

$$\frac{|x_i||y_i|}{\|x\| \, \|y\|} \leq \frac{1}{2}\left(\frac{x_i^2}{\|x\|^2} + \frac{y_i^2}{\|y\|^2}\right)$$

$$\sum_{i=1}^{n} \frac{|x_i y_i|}{\|x\| \, \|y\|} \le \frac{1}{2} \sum_{i=1}^{n} \left(\frac{x_i^2}{\|x\|^2} + \frac{y_i^2}{\|y\|^2} \right)$$

$$= 1$$

$$\sum_{i=1}^{n} |x_i y_i| \le \|x\| \, \|y\|$$

<u>HW</u> : Prove Hölder's inequality

CONVEX OPTIMIZATION PROBLEMS

Problem :

$$\text{Minimize} \quad f(\underline{x})$$

$$\text{ST} \quad f_i(\underline{x}) \leq 0 \qquad i = 1, 2, \dots m$$

$$h_i(\underline{x}) = 0 \qquad i = 1, 2, \dots, p$$

$$f, f_i, h_i : \mathbb{R}^n \to \mathbb{R}$$

Constraint set

or Feasible set

$$C = \left\{ \underline{x} \in \mathbb{R}^n : \begin{array}{ll} f_i(\underline{x}) \leq 0 & ; i = 1, \dots m \\ \& \; h_i(\underline{x}) = 0 & i = 1, \dots p \end{array} \right\}$$

$$F^* = \inf_{x \in C} f(x) \longrightarrow \text{Optimal value}$$

If $C = \phi$, then we define $F^* = \infty$.

Define: $\tilde{f}(x) = \begin{cases} f(x) & \text{if } x \in C \\ \infty & \text{if } x \notin C \end{cases}$

Claim: If f is convex, & C is convex, then $\tilde{f}$ is convex.

$x_1, x_2 \in \mathbb{R}^n \qquad 0 \le \theta \le 1$

Suppose $x_1 \notin C$ or $x_2 \notin C$

$$\tilde{f}(\theta x_1 + (1-\theta) x_2) \le \underbrace{\theta \tilde{f}(x_1) + (1-\theta) \tilde{f}(x_2)}_{}$$

$$\overset{\infty}{\shortparallel}$$

$$F^* = \inf_{x \in C} f(x)$$

$$X_{opt} = \text{set of all optimal points}$$

$$= \{ x \in C : f(x) = F^* \}$$

If $X_{opt} \neq \emptyset$, then we say that the problem is solvable.

A point $x \in C$ is ϵ-suboptimal if

$$f(x) - F^* \leq \epsilon .$$

① $f(x) = -x$

② $f(x) = e^{-x}$; $x \geq 0$

$f^* = 0 \longrightarrow$ No $x \in \mathbb{R}$ st $f(x) = f^*$

But can find ε-suboptimal pts.

* ε-suboptimal set: st of all ε-suboptimal pts.

Local optimality: We say that x^* is a local optimum of f if $\exists\ \varepsilon > 0$ st

$$f(x^*) \leq f(x) \quad \forall x \text{ st } \|x - x^*\| \leq \varepsilon$$

Constraints : $f_i(x) \le 0$

This constraint is active at x if $f_i(x) = 0$

inactive if $f_i(x) < 0$

This constraint is said to be redundant if removing it does not change C.

Problem

Minimize $f(x)$

st $f_i(x) \le 0$

$h_i(x) = 0$

$f(x) = 0 \quad \forall x$

Feasibility problem

$\equiv$ finding if the constraints are consistent

Statistical Estimation

$$X \longrightarrow \boxed{p_{Y|X}} \longrightarrow Y$$

parameter to estimate

observable

Assumption 1: We know $p_{Y|X}$

Assumption 2: X is random & $X \sim p_X$

X, Y discrete

Design $\hat{X}(Y)$ s.t. $\Pr[\hat{X} \neq X]$ is minimized.

MAP estimate: $\hat{x}(y) = \arg\max_x \; p(y|x)\, p(x)$

(Maximum A-posteriori) $= \arg\max_x \; p(x|y)$

$$\Pr[\hat{X} \neq X] = 1 - \Pr[\hat{X} = X]$$

$$= 1 - \sum_x \sum_y p(x,y)\, \mathbb{1}\{x = \hat{x}(y)\}$$

$$= 1 - \sum_y p(\hat{x}(y), y)$$

$$= 1 - \underbrace{\sum_{y} p(y)\, p(\hat{x}(y) \mid y)}$$

Maximized when

$$\hat{x}(y) = \arg\max_{x} p(x \mid y)$$

MAP estimation problem:

Maximize over x $p(x, y)$

$p(y|x)$ — likelihood function

$\log p(y|x)$ — log likelihood function

MAP estimation: Maximizing $\log p(x, y)$

$$= \log p(y|x) + \log p(x)$$

A distribution $p(x)$ is log-concave if $\log p(x)$ is concave

- $p(x) = \dfrac{e^{-x^2/2}}{\sqrt{2\pi}}$

$\log p(x) = -\dfrac{x^2}{2} - \log\sqrt{2\pi}$

- $p(x) = \dfrac{e^{-|x|/a}}{2a}$ log concave

- $p(x) = \begin{cases} \frac{1}{2a} & \text{if } x \in [-a, a] \\ 0 & \text{else} \end{cases}$

log concave

Maximum likelihood estimator

Given y,

$$x_{ML} = \arg\max_{x \in \mathcal{X}} p(y|x)$$

$$= \arg\max_{x \in \mathcal{X}} \log p(y|x)$$

Linear measurements & iid noise

Goal: Estimate $\underline{x}$ from

Measurements
$$\begin{cases} y_1 = \underline{a}_1^T \underline{x} + z_1 \\ y_2 = \underline{a}_2^T \underline{x} + z_2 \\ \vdots \\ y_m = \underline{a}_m^T \underline{x} + z_m \end{cases}$$

$z_1\ z_2 \dots z_m$ are iid

$\underline{a}_1 \dots \underline{a}_m$ are given

$$y = A\underline{x} + z$$

$\underline{x} \in \mathbb{R}^n$, A: $m \times n$ known matrix

z: random vector with iid components

Θ z: iid $\mathcal{N}(0, \sigma^2)$

$$\underline{x}_{ML} = \arg\max_{\underline{x}} \log p(y \mid \underline{x})$$

$$= \arg\max_{\underline{x}} \log \left[\frac{1}{(2\pi\sigma^2)^{n/2}} e^{-\|y - A\underline{x}\|^2 / 2\sigma^2} \right]$$

$$= \arg\min_{\underline{x}} \|y - A\underline{x}\|^2$$

② $z_i \sim \frac{1}{2\alpha} e^{-|z_i|/\alpha}$

$$\underline{x}_{ML} = \arg\min_{\underline{x}} \|\underline{y} - A\underline{x}\|_1$$

$$p(\underline{y}|\underline{x}) = \frac{1}{(2\alpha)^n} e^{-\sum_{i=1}^{n} |y_i - \underline{a}_i^T \underline{x}|/\alpha}$$

$$= \frac{1}{(2\alpha)^n} e^{-\frac{\|\underline{y} - A\underline{x}\|_1}{\alpha}}$$

③ $z_i \sim \text{Unif}(-\alpha, \alpha)$

$$p(\underline{y}|\underline{x}) = \begin{cases} \left(\frac{1}{2\alpha}\right)^n, & \text{if } y_i \in [\underline{a}_i^T\underline{x} - \alpha, \ \underline{a}_i^T\underline{x} + \alpha] \\ 0, & \text{else} \end{cases}$$

$$p(y|x) = \begin{cases} \dfrac{1}{(2a)^n} & \text{if } \underbrace{\max_{1\le i\le n} |y_i - a_i^T x| \le \alpha}_{\|y - Ax\|_\infty \le \alpha} \\ 0 & \text{else} \end{cases}$$

$$x_{ML} = \operatorname*{argmin}_{x \in \mathbb{R}^n} \|y - Ax\|_\infty$$

$$x_{ML} = \operatorname*{argmin}_{x : \|y - Ax\|_\infty \le \alpha} 1 \longrightarrow \text{feasibility problem}$$

MAP estimation :

$$y = A\underline{x} + z \quad : \quad z \sim \text{iid } \mathcal{N}(0, \sigma^2)$$

$$\underline{x} \sim \text{iid } \mathcal{N}(\underline{\mu}, \rho I)$$

$$\underline{x}_{MAP} = \arg\max_{\underline{x}} \log p(x, y)$$

$$= \arg\max_{\underline{x}} \left(\log p(y|x) + \log p(x) \right)$$

Sum of convex functions

If $f_1(x)$ & $f_2(x)$ are convex, then

$\alpha_1 f_1(x) + \alpha_2 f_2(x)$ is convex if

$\alpha_1 \geq 0$ & $\alpha_2 \geq 0$

$$\alpha_1 f_1(\theta x_1 + (1-\theta)x_2) + \alpha_2 f_2(\theta x_1 + (1-\theta)x_2)$$

$$\leq \alpha_1 \left(\theta f_1(x_1) + (1-\theta) f_1(x_2)\right) + \alpha_2 \left(\theta f_2(x_1) + (1-\theta) f_2(x_2)\right)$$

$$= \theta \left(\alpha_1 f_1(x_1) + \alpha_2 f_2(x_1)\right) + (1-\theta)\left(\alpha_1 f_1(x_2) + \alpha_2 f_2(x_2)\right)$$

Suppose $f(\underline{x}, \underline{y})$ where $f(\underline{x}, \underline{y})$ is convex in $\underline{x}$

for every $\underline{y}$

Then $\int_{y} f(\underline{x}, \underline{y}) \, dy$ is also convex

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

A: $n \times m$ matrix, $b \in \mathbb{R}^n$

$$g(x) = f(Ax + b) \qquad g: \mathbb{R}^m \to \mathbb{R}.$$

$$\begin{aligned}
g(\theta x_1 + (1-\theta) x_2) &= f(A(\theta x_1 + (1-\theta) x_2) + b) \\
&= f(\theta(Ax_1 + b) + (1-\theta)(Ax_2 + b)) \\
&\leq \theta f(Ax_1 + b) + (1-\theta) f(Ax_2 + b) \\
&= \theta g(x_1) + (1-\theta) g(x_2)
\end{aligned}$$

Suppose $f(\underline{x}) = g(\underline{h}(\underline{x}))$ $\quad \underline{h} : \mathbb{R}^n \to \mathbb{R}^m$

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

When is f convex?

Simpler: $g : \mathbb{R} \to \mathbb{R}$

$h : \mathbb{R} \to \mathbb{R}$.

$$f''(x) = \underbrace{g''(h(x))}_{\geq 0} \underbrace{(h'(x))^2}_{\geq 0} + \underbrace{g'(h(x))}_{\geq 0} \underbrace{h''(x)}_{\geq 0}$$

If g, h convex

If g, h concave $\quad \leq 0 \quad \geq 0$

g is non decreasing $\quad \geq 0 \quad \leq 0$

g	h	$\Rightarrow$	$f = g(h(x))$
Convex nondecreasing	Convex	$\Rightarrow$	Convex
Convex nonincreasing	Concave	$\Rightarrow$	Convex
Concave nondecreasing	Concave	$\Rightarrow$	Concave
Concave nonincreasing	Convex	$\Rightarrow$	Concave

Give an example of g, h convex but $g(h(x))$ is not.

$$h(x) = -\log x$$

$$g(x) = -x$$

$$g(h(x)) = \log x \rightarrow \text{concave}$$

Suppose $g : \mathbb{R}^m \to \mathbb{R}$, $\underline{h} : \mathbb{R}^n \to \mathbb{R}^m$

g : convex & nondecreasing

h_i : convex

$$\underline{h} = \begin{pmatrix} h_1 \\ h_2 \\ \vdots \\ h_m \end{pmatrix}$$

$$\underline{h}(\underline{x}) = \begin{pmatrix} h_1(\underline{x}) \\ h_2(\underline{x}) \\ \vdots \\ h_m(\underline{x}) \end{pmatrix}$$

g is said to be nondecreasing if

$g(\underline{x}_1) \leq g(\underline{x}_2)$ as long as $\underline{x}_1 \preceq \underline{x}_2$ (componentwise)

Then $f(\underline{x}) = g(\underline{h}(\underline{x}))$ is convex

HW

Check convexity properties

$$\underline{h}(\underline{x}) = \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} \qquad h_i(\underline{x}) = \frac{e^{x_i}}{\sum^{n} e^{x_j}}$$

$\int^{x_1}$

Proof:

$$f(\theta \underline{x_1} + (1-\theta)\underline{x_2})$$

$$= g(\underline{h}(\theta x_1 + (1-\theta)\underline{x_2}))$$

$$\leq g(\theta \underline{h}(\underline{x_1}) + (1-\theta)\underline{h}(\underline{x_2}))$$

$$\leq \theta g(\underline{h}(\underline{x_1})) + (1-\theta) g(h(\underline{x_2}))$$

for each i,

$$h_i(\theta \underline{x_1} + (1-\theta)\underline{x_2})$$

$$\leq \theta h_i(\underline{x_1}) + (1-\theta) h_i(\underline{x_2})$$

Examply : ① $g(x)$ is convex

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

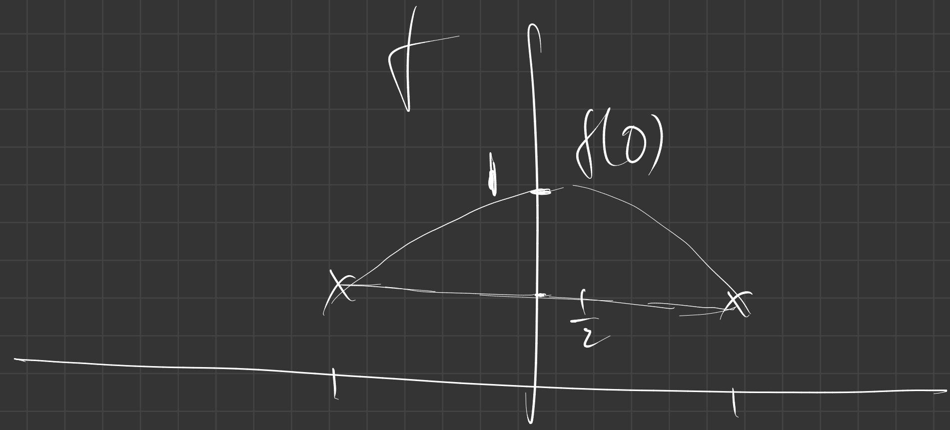
② $g_1 \; g_2 \cdots g_m$ all convex

$\Rightarrow$ $e^{\sum_{i=1}^{m} g_i(x)}$ is convex

③ $g(x)$ is positive & convex

$\frac{1}{g(x)}$ is convex ?

$f(x) = \frac{1}{x^2+1}$ $\qquad f''(x) =$



$f(1)\,,\quad f(-1)$

Exercise $\quad f(x) = \left(\sum_{i=1}^{n} |g_i(x)|^p \right)^{1/p} \qquad 0 < p < \infty$

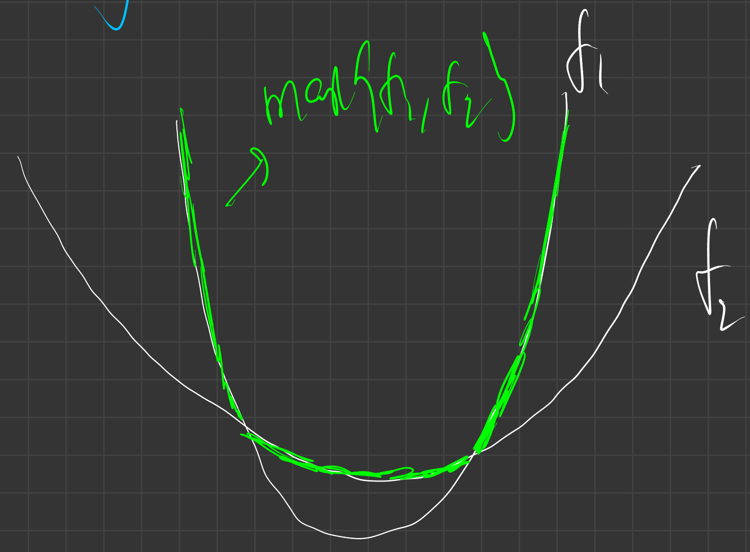
$$\left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p}$$

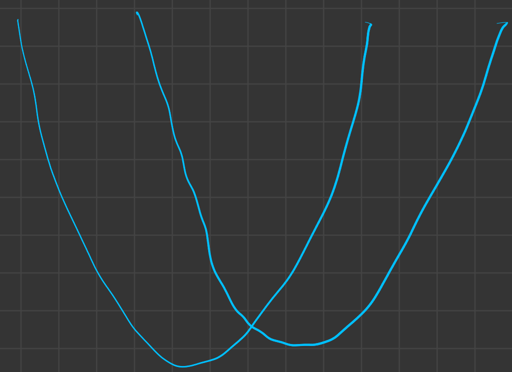
When is f convex/concave?

Suppose $f_1(x)$ & $f_2(x)$ are convex.

$$f(x) = \max_{x} \{ f_1(x), f_2(x) \}$$

is convex





Proof:
$$f(\theta x_1 + (1-\theta) x_2)$$

$$= \max \{ f_1(\theta x_1 + (1-\theta) x_2), \; f_2(\theta x_1 + (1-\theta) x_2) \}$$

$$\le \max \{ \theta f_1(x_1) + (1-\theta) f_1(x_2), \; \theta f_2(x_1) + (1-\theta) f_2(x_2) \}$$

$$\leq \max\{\theta f_1(x_1), \theta f_2(x_1)\} + \max\{(1-\theta) f_1(x_2), (1-\theta) f_2(x_2)\}$$

$$= \theta \max\{f_1(x_1), f_2(x_1)\} + (1-\theta)\max\{f_1(x_2), f_2(x_2)\}$$

$$= \theta f(x_1) + (1-\theta) f(x_2)$$

More generally if $f(\underline{x}, y)$ is convex in x for each y, then

$$\bar{f}(\underline{x}) = \sup_{y \in A} f(\underline{x}, y) \quad \text{is convex}$$

Example

$f(A) = \lambda_{max}(A)$ $\quad$ $\mathrm{dom}(f) = S^n_+$

λ_{max} → largest eigenvalue of A

$$\lambda_{max}(A) = \sup_{x \in \mathbb{R}^n} \frac{x^T A x}{\|x\|^2}$$

$$\lambda_{max}(A) = \sup_{\substack{x \in \mathbb{R}^n \\ \|x\| = 1}} \underbrace{x^T A x}$$

linear (hence convex) function of A

→ convex.

Let $(\underline{u}_1, \ldots, \underline{u}_n)$ be the eigenbasis corresp to A.

$$\underline{x} = \alpha_1 \underline{u}_1 + \alpha_2 \underline{u}_2 + \cdots + \alpha_n \underline{u}_n$$

$$\downarrow$$

$$\lambda_{max}$$

$$\underline{x}^T A \underline{x} = \alpha_1^2 \lambda_1 + \alpha_2^2 \lambda_2 + \cdots + \alpha_n^2 \lambda_n$$

If $\underline{x}$ is a unit vector,

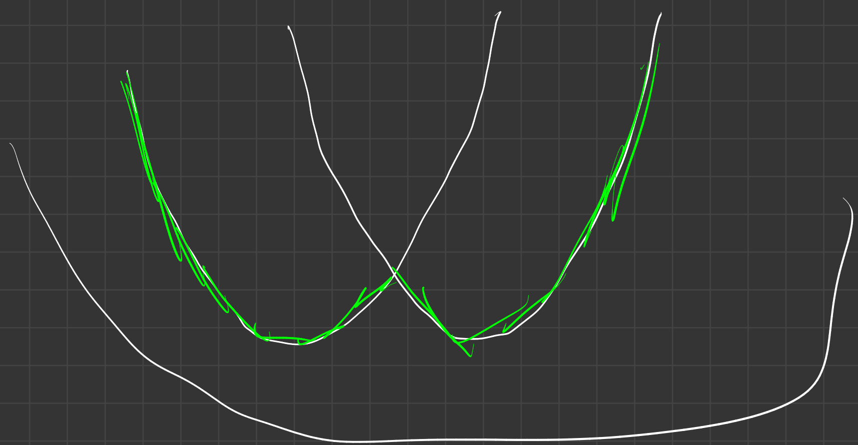
$$\alpha_1^2 + \alpha_2^2 + \cdots + \alpha_n^2 = 1$$

$$\begin{array}{ccc} \alpha_1^2 & \alpha_2^2 & \alpha_n^2 \\ \parallel & \parallel & \parallel \\ \beta_1 & \beta_2 & \beta_n \end{array}$$

Suppose f_1, f_2, f_3 are convex function

Let $sl(x)$ = second larger of $f_1(x), f_2(x), f_3(x)$

$(x-1)^2, \quad (x+1)^2$



Not convex.

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

$g(x) = \min\limits_{y \in \{+1,-1\}} f(x,y)$

(3) $f(\underline{x})$ = sum of the $\underline{k}$ largest components of $\underline{x}$

$$f_{i_1 i_2}(\underline{x}) = x_{i_1} + x_{i_2}$$

$$f(\underline{x}) = \max_{\substack{i_1, i_2 \\ i_1 \neq i_2}} f_{i_1 i_2}(\underline{x})$$

$\rightarrow$ convex for every i_1, i_2

$\Rightarrow$ f is convex.

Minimum/Infimum of convex function

If, $f(x, y)$ is convex in (x, y),

then $\inf\limits_{y \in C} f(x, y)$ is convex

$C \to$ convex

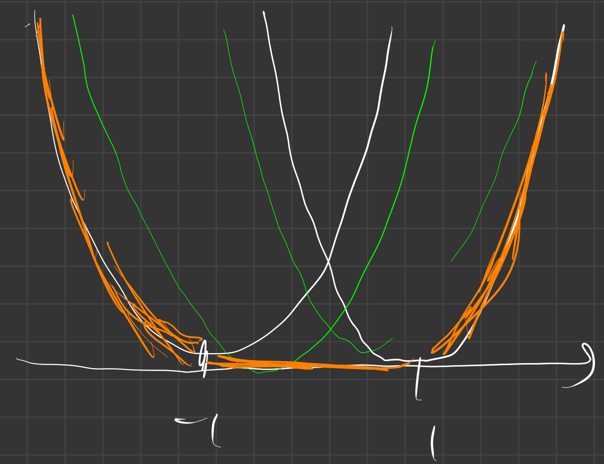
$x \in \mathbb{R}^n$, $y \in \mathbb{R}^m$

$f : \mathbb{R}^{n+m} \to \mathbb{R}$

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

$$g(x) = \inf_{y \in (-1,1)} f(x,y)$$

$$= \begin{cases} (x+1)^2 & \text{for } x \leq -1 \\ (x-1)^2 & \text{for } x \geq 1 \\ 0 & \text{for } x \in (-1,1) \end{cases}$$



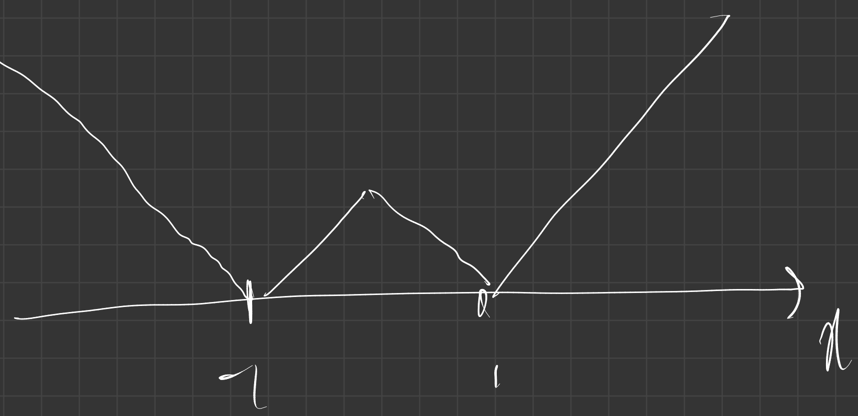
② C is a convex set.

$$f(\underline{x}) = \text{distance b/w } \underline{x} \text{ \& } C$$

is convex

$$= \inf_{y \in C} \|\underline{x} - y\|$$

$C = \{1, -1\}$ $\qquad x \in \mathbb{R}$



$$f_{12}(x) = x_1 + x_2$$

$$f_{23}(x) = x_2 + x_3$$

$$f_{13}(x) = x_1 + x_3$$

Conjugate of a function

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

$$f^*(y) = \sup_{\underline{x} \in \mathrm{dom}(f)} \underbrace{\left(\underline{x}^T y - f(\underline{x})\right)}_{g(\underline{x},\, y)}$$

This is also called the Legendre-Fenchel transform.

- For each $\underline{x}$, $g(\underline{x}, y)$ is a linear fn of y (hence convex in y)

$\Rightarrow$ $f^*(y)$ is convex irrespective of f.

Examples

① $f(\underline{x}) = \underline{a}^T \underline{x} + b$

$$f^*(\underline{y}) = \sup_{\underline{x}} \left(\underline{x}^T \underline{y} - \underline{a}^T \underline{x} - b \right)$$

$$= \begin{cases} -b & \text{if } \underline{y} = \underline{a} \\ \infty & \text{else.} \end{cases}$$

$$f^{**}(\underline{x}) = \sup_{\underline{y}} \Big(\underbrace{\underline{x}^T \underline{y} - f^*(\underline{y})}_{g(x,y)} \Big)$$

$$g(x,y) = \begin{cases} -\infty & \text{if } \underline{y} \neq \underline{a} \\ \underline{a}^T \underline{x} + b & \text{if } \underline{y} = \underline{a} \end{cases}$$

$$\Rightarrow \quad f^{**}(\underline{x}) = a^T x + b = f(x)$$

② $f(x) = x^2 \qquad f: \mathbb{R} \to \mathbb{R}$

$$f^*(y) = \sup_x \Big(\underbrace{xy - x^2}_{\text{concave}} \Big)$$

$$\frac{\partial g}{\partial x} = 0 \Rightarrow y - 2x = 0 \Rightarrow x = y/2$$

$$f^*(y) = y^2/4$$

$$f^{**}(x) = \sup_y \big(\underbrace{xy - f^*(y)}_{\text{concave}} \big)$$

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

Q: Is $f^{**}(x) = f(x)$ for all f?

A: No, since f^{**} is always convex

NOTE: $f^{**} = f$ iff f is convex & epigraph of f is closed.

(3) $f(x) = -\ln x \qquad f: \mathbb{R}_{++} \to \mathbb{R}$

$$g(x,y) = xy + \ln x$$

$$f^*(y) = \sup_{x \in \mathbb{R}_{++}} g(x,y) = \begin{cases} \infty & \text{if } y \geq 0 \\ -1 - \ln(-y) & \text{if } y < 0 \end{cases}$$

(for $y<0$, set $\frac{\partial g}{\partial x} = 0$ & substitute)

(4) $f(\underline{x}) = \|\underline{x}\|_2$

$$f^*(y) = \sup_{\underline{x}} \left(\underline{x}^T y - \|x\|_2 \right)$$

$$= \begin{cases} \infty & \text{if } \|y\| > 1 \\ 0 & \text{if } \|y\| \leq 1 \end{cases}$$

Case 1: $\|y\| > 1$.

Take $\underline{x} = \alpha y$

$$g(x, y) = \alpha \|y\|^2 - \alpha \|y\| = \alpha \underbrace{\|y\| \left(\|y\| - 1 \right)}_{> 0}$$

$\rightarrow \infty$ as $\alpha \rightarrow \infty$

$$\Rightarrow \sup_{x} g(x, y) = \infty$$

Case 2 : $\|y\| \le 1$

By Cauchy - Schwartz

$$g(x,y) = x^T y - \|x\|_2 \le \|x\|\,\|y\| - \|x\|$$

$$= \|x\| \underbrace{\left(\|y\| - 1\right)}_{\le 0}$$

$\Rightarrow \quad g(x,y) \le 0 \quad \forall x \in \mathbb{R}^n$

$\sup_{x} g(x,y) = 0 \qquad$ (Take $x = 0$)

HW : Find f^* & f^{**} for $f(x) = x \ln x$

$f : \mathbb{R}_{++} \to \mathbb{R}$.