Convex Functions

RECAP

Goal: Minimize $f_0(\underline{x})$

ST

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

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

① KKT conditions

(Sufficient if convex)

Convex Function

$f : V \longrightarrow \mathbb{R}$ is convex if $\forall\, \underline{x}_1, \underline{x}_2 \in \mathrm{Dom}(f)$

convex

$$f(\alpha \underline{x}_1 + (1-\alpha)\underline{x}_2) \le \alpha f(x_1) + (1-\alpha) f(x_2)$$

$$\forall \alpha \in [0, 1]$$

If $\mathrm{Dom}(f) \neq V$,

$$\tilde{f}(\underline{x}) = \begin{cases} f(\underline{x}) & \text{if } \underline{x} \in \mathrm{dom}(f) \\ \infty & \text{else} \end{cases}$$

Suppose $\underline{x}_1 \notin \mathrm{Dom}(f)$ & $\alpha \in (0, 1)$

$$\text{RHS} = \underbrace{\alpha \tilde{f}(x_1)}_{= \infty} + (1-\alpha)\tilde{f}(x_2) = \infty$$

If x_1 & $x_2 \in \mathrm{Dom}(f)$, $\tilde{f}(x_1) = f(x_1)$

$\tilde{f}(x_2) = f(x_2)$

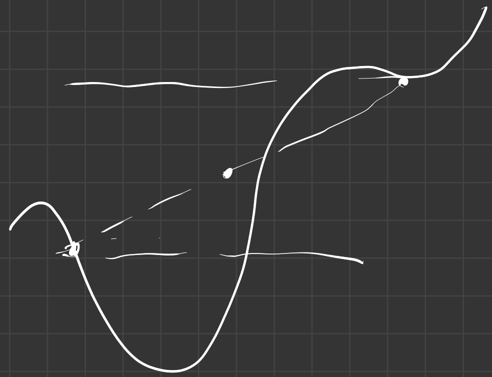
$$\tilde{f}(\alpha x_1 + (1-\alpha)x_2) \le \alpha f(x_1) + (1-\alpha) f(x_2)$$

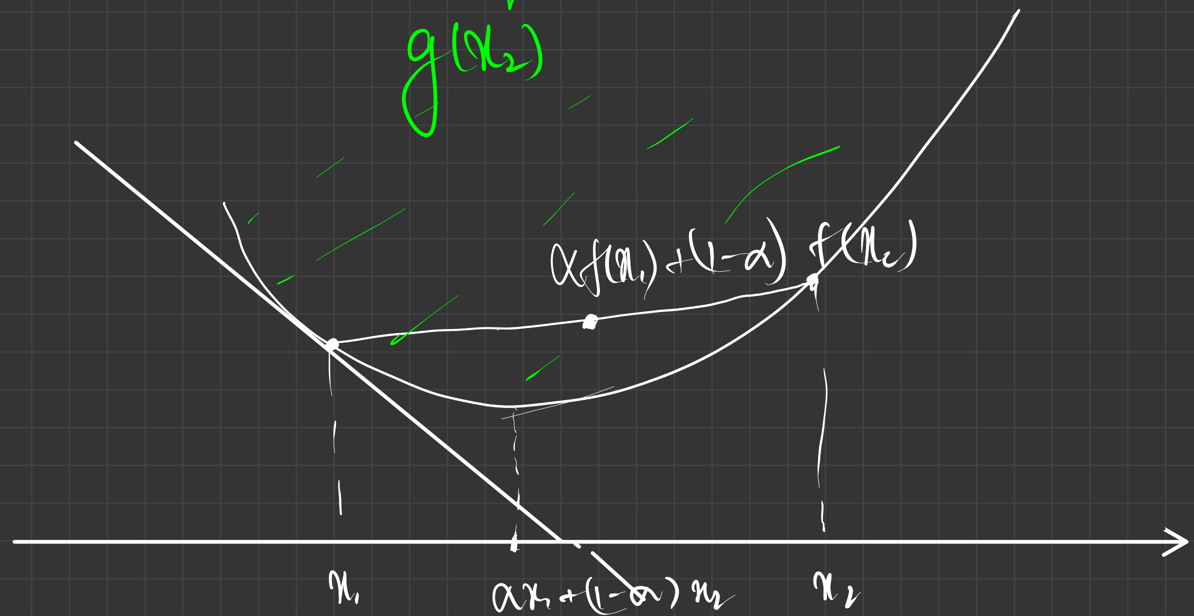
First order characterization

Suppose $f : \mathbb{V} \to \mathbb{R}$ & $\nabla f(\underline{x})$ exists $\forall \underline{x} \in \mathbb{V}$.

f is convex iff $\forall \underline{x}_1, \underline{x}_2$,

$$f(\underline{x}_2) \geq \underbrace{f(\underline{x}_1) + \left(\nabla f(\underline{x}_1)\right)^T (\underline{x}_2 - \underline{x}_1)}_{g(\underline{x}_2)}$$





f is convex iff $f(x)$ lies above the tangent to f at x_1
$\forall x_1$

Proof: Suppose f is convex & ∇f exists.

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

$$\underline{x} = (1-t)\underline{x}_1 + t\underline{x}_2 \qquad t \in [0,1]$$

$$= \underline{x}_1 + t(\underline{x}_2 - \underline{x}_1)$$

Since f is convex,

$$f(\underline{x}) \leq (1-t)f(x_1) + t f(x_2) = f(x_1) - t f(x_1) + t f(x_2)$$

$$t f(x_2) \geq t f(x_1) + f(x) - f(x_1)$$

$$f(x_2) \geq f(x_1) + \frac{f(x_1 + t(x_2 - x_1)) - f(x_1)}{t} \qquad -(1)$$

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

$$f(x_2) = f(x_1) + \frac{\left(f(x_1 + t(x_2 - x_1)) - f(x_1)\right)}{t(x_2 - x_1)} (x_2 - x_1)$$

$$= f(x_1) + \underbrace{\frac{f(x_1 + t(x_2 - x_1)) - f(x_1)}{(x_1 + t(x_2 - x_1)) - x_1}}_{\text{as } t \to 0 \;\longrightarrow\; f'(x_1)} \times (x_2 - x_1)$$

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

$$\lim_{y \to x} \frac{\| f(y) - f(x) - (\nabla f(x))^T (y - x) \|}{\| y - x \|} = 0$$

Take $y = x_1 + t(x_2 - x_1)$ & $x = x_1$

$y - x = t(x_2 - x_1)$

$$\lim_{t\to 0} \frac{\| f(y) - f(x) - (\nabla f(x))^T [t(x_2 - x_1)] \|}{t \|x_2 - x_1\|}$$

Define $e(t) = f(y) - f(x) - \nabla f(x)^T [t(x_2 - x_1)]$

$$\frac{f(y) - f(x)}{t} = \frac{e(t)}{t} + (\nabla f(x))^T (x_2 - x_1) \quad \text{— ②}$$

where $\lim_{t\to 0} \frac{\|e(t)\|}{t\|x_2 - x_1\|} = 0$.

Use ② in ①

$$f(x_2) \geq f(x_1) + \frac{e(t)}{t} + \nabla f(x_1)^T (x_2 - x_1)$$

As $t \to 0$, $\frac{e(t)}{t} \to 0$

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

Consider f st $\forall x_1, x_2,$

$$f(x_2) \geq f(x_1) + \nabla f(x_1)^T (x_2 - x_1) \quad \text{—} (3)$$

Fix x_1, x_2, $\alpha \in [0,1]$

$$\underline{x} = \alpha \underline{x_1} + (1-\alpha) \underline{x_2}$$

Using (3),

$$f(x_1) \geq f(x) + \nabla f(x)^T (x_1 - x) \quad \text{—} (4)$$

$$f(x_2) \geq f(x) + \nabla f(x)^T (x_2 - x) \quad \text{—} (5)$$

α (4) $+ (1-\alpha)$ (5)

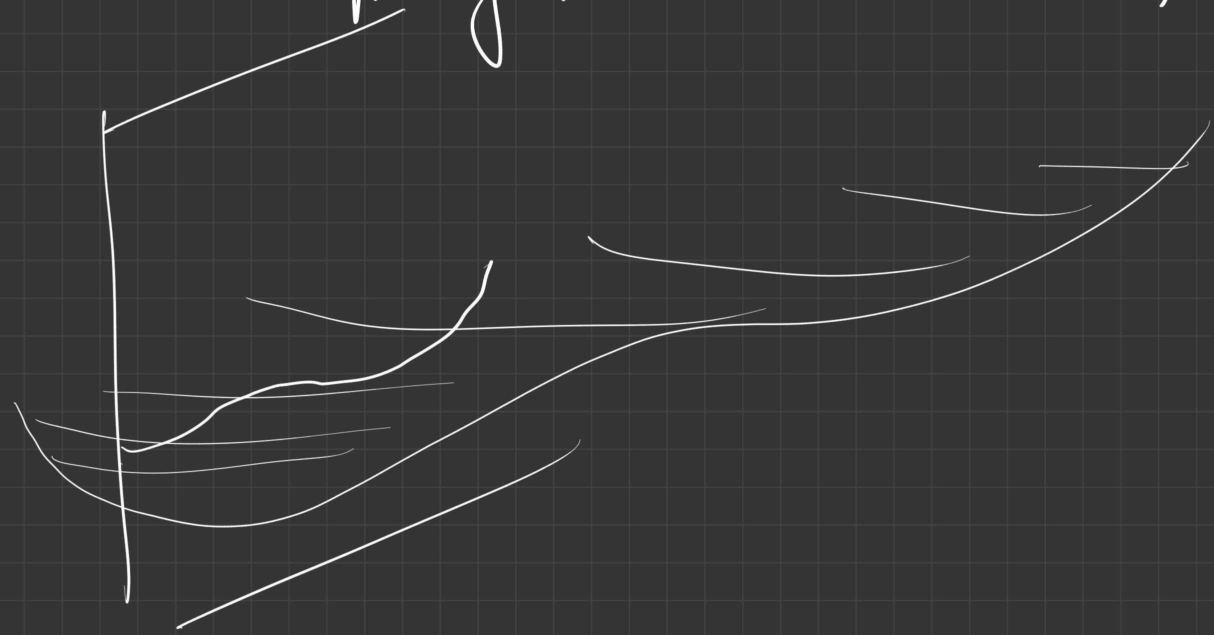
$$\alpha f(x_1) + (1-\alpha) f(x_2) \geq \underbrace{f(x)}_{= f(\alpha x_1 + (1-\alpha) x_2)} + \nabla f(x)^T \underbrace{\left[\alpha x_1 + (1-\alpha) x_2 - x \right]}_{0}$$

$\therefore$ f is convex

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

$$g(t) = f(x_1 + t(x_2 - x_1))$$

<u>Claim</u>: f is convex iff $g(\cdot)$ is convex $\forall x_1, x_2$



Second order characterization

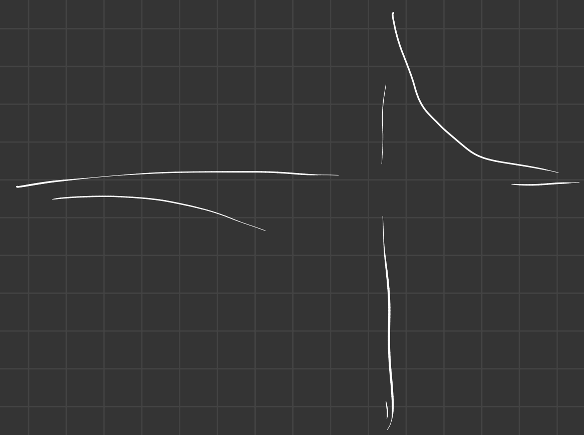
$f : \mathbb{R}^n \to \mathbb{R}$ & $\nabla^2 f$ exists.

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

Examples

① $f(x) = e^{\alpha x}$

$$f''(x) = \alpha^2 e^{\alpha x} \geq 0 \quad \forall x, \quad \forall \alpha$$



② x^α

$$f''(x) = \alpha(\alpha-1)\, x^{\alpha-2} \geq 0$$

(i) $\alpha \geq 1 \qquad x > 0$

$\alpha \geq 1 \rightarrow f(x)$ is convex over $(0, \infty)$

$x < 0 \rightarrow$ α even $\Rightarrow f$ is convex

$\alpha > 0$

α odd $\Rightarrow f$ is concave

$\alpha \leq 0$, $x \in (0, \infty) \Rightarrow f$ is convex

α integer $\quad x \in (-\infty, 0) \rightarrow f$ is concave for α odd

f is convex for α even

(3) $f(x) = \log x$ $\quad x > 0$

$$f''(x) = -\frac{1}{x^2} < 0 \quad x > 0$$

Concave

(4) $f : \mathbb{R}^n \to \mathbb{R}$

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

$$\begin{aligned} f(\alpha x_1 + (1-\alpha) x_2) &= \|\alpha x_1 + (1-\alpha) x_2\| \\ &\leq \|\alpha x_1\| + \|(1-\alpha) x_2\| \quad \text{Triangle ineq} \\ &= \alpha \|x_1\| + (1-\alpha) \|x_2\| \\ &= \alpha f(x_1) + (1-\alpha) f(x_2) \end{aligned}$$

(5) $|x|^\alpha \quad \alpha \geq 0$

(6) $x \log x \quad x \in (0, \infty)$

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

$$f(\underline{x}) = \max_{i=1,2,\dots,n} x_i$$

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

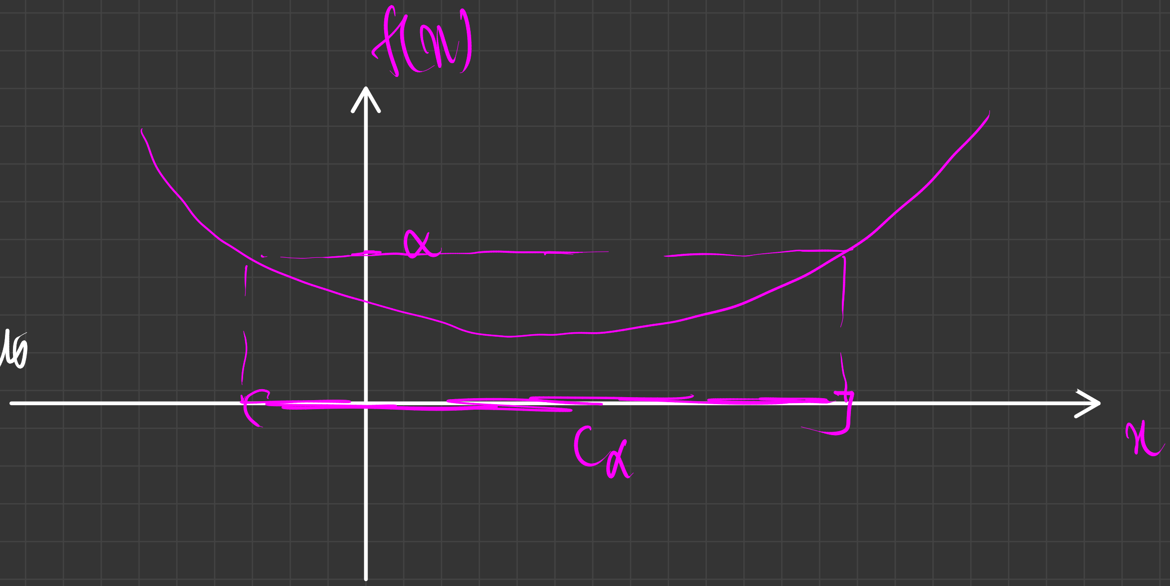
9. $$f(\underline{x}) = \left(\prod_{i=1}^{n} x_i \right)^{1/n}$$

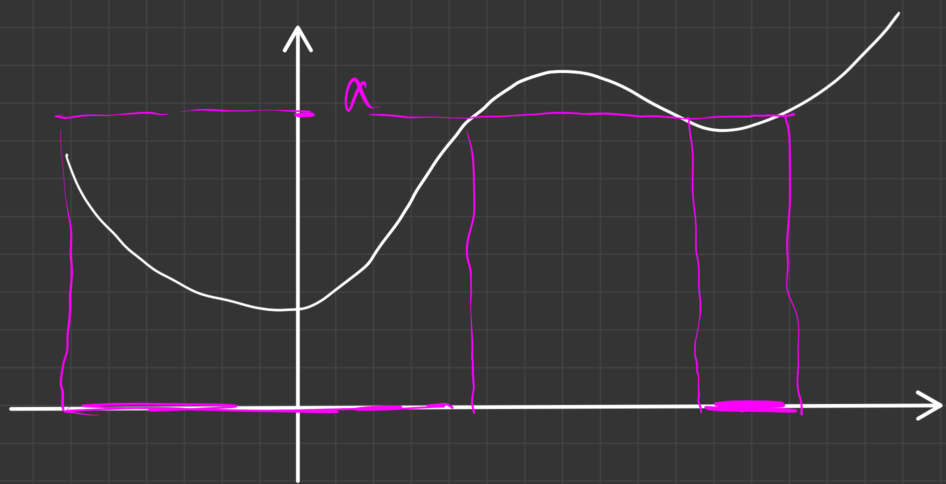
Sublevel set

For any $f: \mathbb{R}^n \to \mathbb{R}$, any $\alpha \in \mathbb{R}$,

$$C_\alpha = \{ \underline{x} \in \mathbb{R}^n : f(\underline{x}) \leq \alpha \}$$

<u>Claim</u>: If f is convex, then all sublevel sets C_α $\alpha \in \mathbb{R}$ is convex





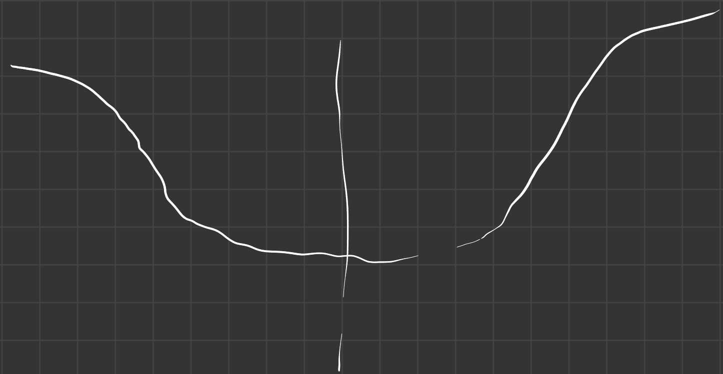
Q: If all sublevel sets of f

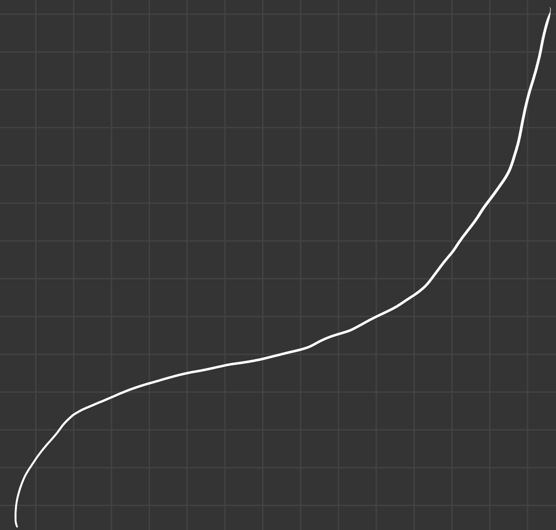
$$C_\alpha = \{ x \in \mathbb{R}^n : f(x) \le \alpha \}$$

is convex $\forall \alpha$

is f convex?

$f(x) = \log x$

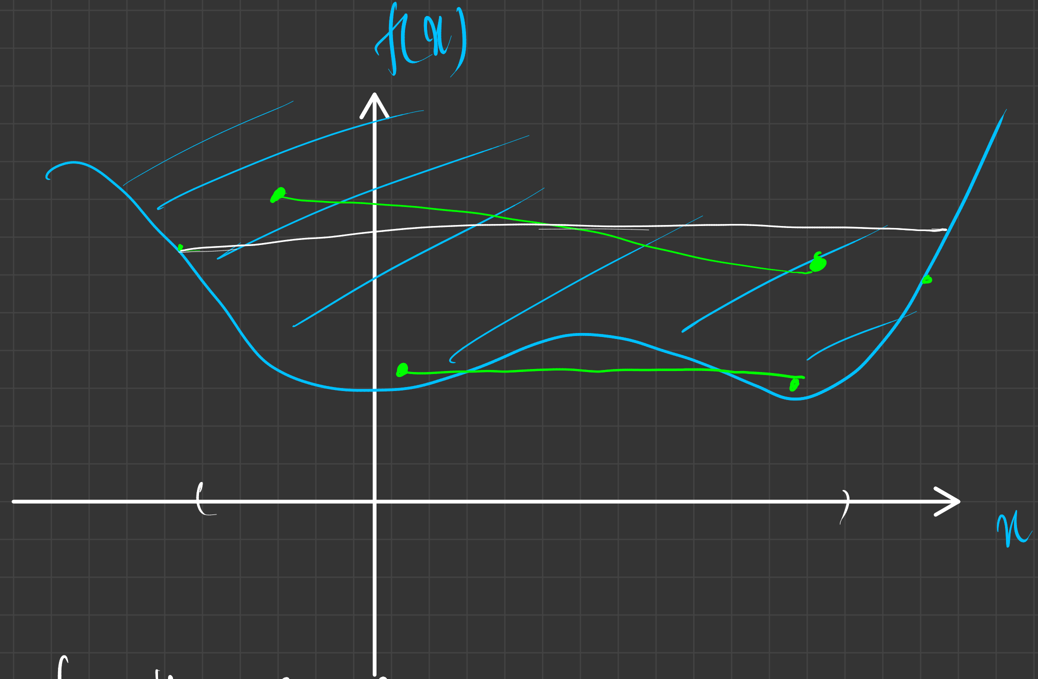




Epigraph: $f : \mathbb{R}^n \to \mathbb{R}$

$$\text{epi}(f) = \{ (\underline{x}, t) \in \mathbb{R}^{n+1} : f(\underline{x}) \leq t \}$$

$$\text{epi}(f) = \{ (x, y) : f(x) \leq y \}$$



Claim: f is convex $\Rightarrow$ epi(f) is convex

epi(f) is convex $\Rightarrow$ f is convex

Proof: Suppose f is convex

$(\underline{x}_1, t_1)$ $(x_2, t_2) \in \text{epi}(f)$

Does $(\alpha x_1 + (1-\alpha) x_2,\ \alpha t_1 + (1-\alpha) t_2) \in epi(f)$

$$f(\alpha x_1 + (1-\alpha) x_2) \leq \alpha f(x_1) + (1-\alpha) f(x_2)$$

as f is convex

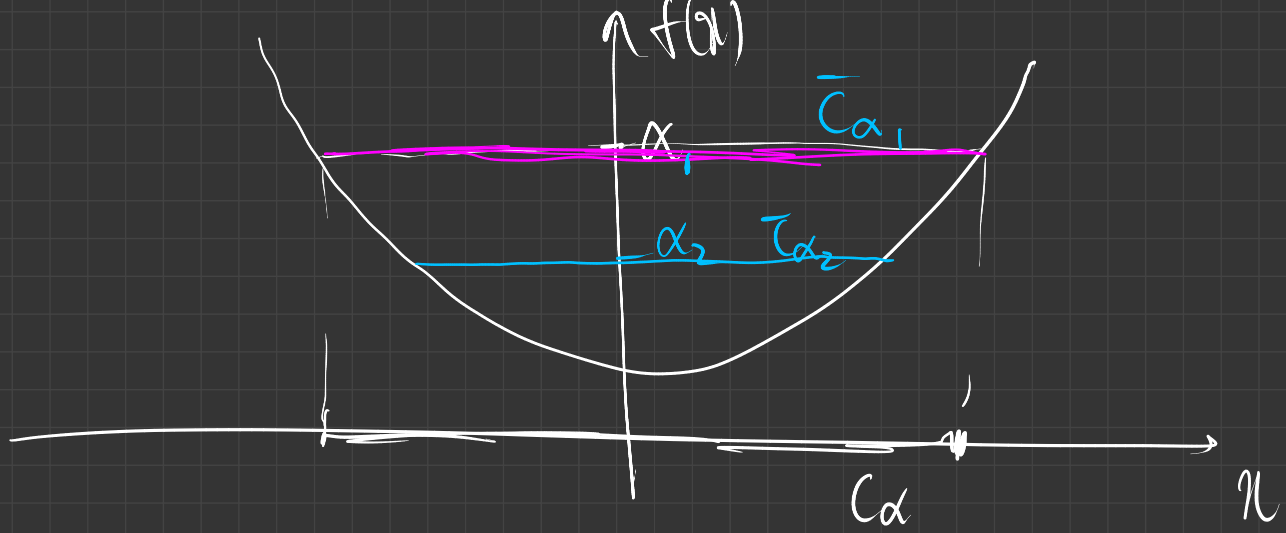
$$\leq \alpha t_1 + (1-\alpha) t_2$$

as $f(x_1) \leq t_1$ & $f(x_2) \leq t_2$

$(x_1, t_1) \in epi(f)$

$(x_2, t_2) \in epi(f)$

$(\{\alpha\}, C_\alpha)$

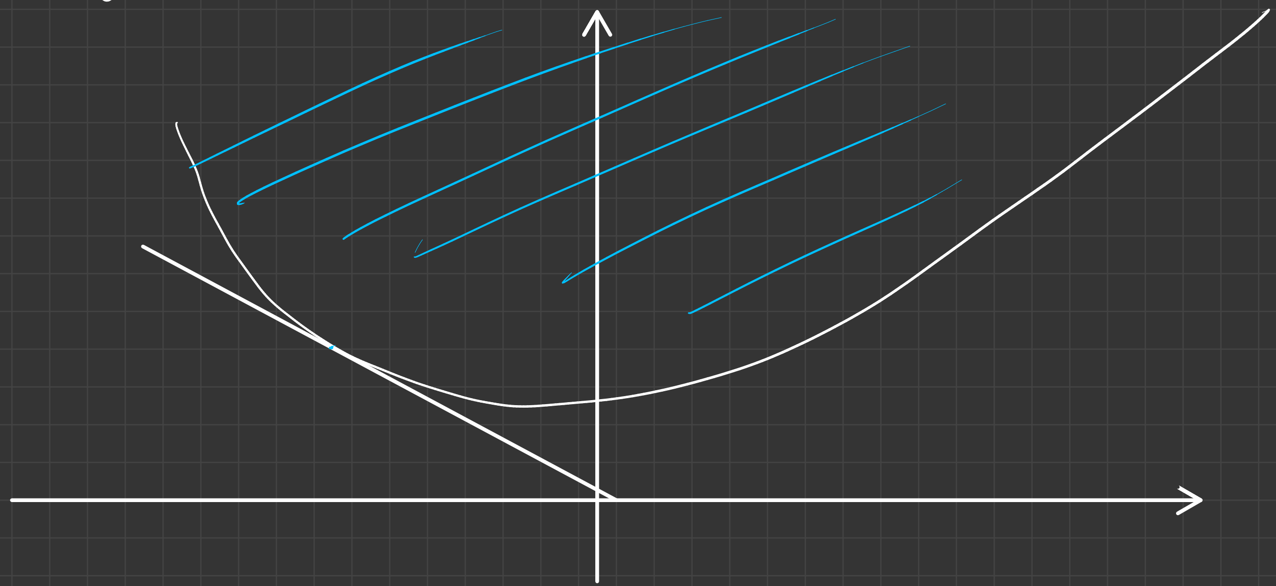


$$C_\alpha = \{ x : f(x) \leq \alpha$$

$$\bar{C}_\alpha = \{ (\alpha, x) : f(x) \leq \alpha \}$$

$$\mathrm{epi}(f) = \bigcup_{\alpha \in \mathbb{R}} \bar{C}_\alpha$$

First order test



If $f: R^n \to R$ is convex,

$$f(\alpha x_1 + (1-\alpha) x_2) \leq \alpha f(x_1) + (1-\alpha) f(x_2)$$

$$f\left(\sum_{i=1}^{m} \alpha_i x_i\right) \leq \sum_{i=1}^{m} \alpha_i f(x_i)$$

$\alpha_i \geq 0$ & $\sum_{i=1}^{m} \alpha_i = 1$

$X \quad p_X(x_i) = \alpha_i \longrightarrow \Pr[X = x_i]$

$$\mathbb{E}X = \sum_{i=1}^{m} \alpha_i x_i$$

$$\mathbb{E}f(X) = \sum_{i=1}^{m} \alpha_i f(x_i)$$

$f(\mathbb{E}X) \leq \mathbb{E}f(X)$ Jensen's inequality

AM-GM inequality

$$\sqrt{ab} \leq \frac{a+b}{2}$$

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

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

Apply Jensen inequality on $-\log x$

$$a^{\alpha} b^{1-\alpha} \leq \alpha a + (1-\alpha) b$$

Hölder inequality: $\underset{\sim}{x}, y \in \mathbb{R}^n$

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

$$\frac{1}{p} + \frac{1}{q} = 1$$

Conic combinations of convex functions are convex

$f_1, f_2 \ldots f_m : \mathbb{R}^n \to \mathbb{R}$

each $f_i : \mathbb{R}^n \to \mathbb{R}$ is convex,

then $f(\underline{x}) = \sum_{i=1}^{m} \theta_i f_i(\underline{x}) \qquad \theta_1 \ldots \theta_m \geq 0$

is convex.

Eg : $f(x) = x^2$ is convex

$f(\underline{x}) = \|x\|_2^2 = \sum_{i=1}^{m} x_i^2$

$f(\underline{x}) = \|\underline{x}\|_1 = \sum_{i=1}^{n} |x_i|$

Composition of affine and convex function

$g : \mathbb{R}^n \to \mathbb{R}^m$, $\quad g(\underline{x}) = A\underline{x} + \underline{b}$ is affine

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

Then, $h(g(\underline{x}))$ is convex.

Consider $\underline{x}_1, \underline{x}_2 \in \mathrm{Dom}(h \circ g)$

$0 \le \alpha \le 1$

$$h(g(\alpha \underline{x}_1 + (1-\alpha)\underline{x}_2)) = h(A(\alpha \underline{x}_1 + (1-\alpha)\underline{x}_2) + \underline{b})$$

$$= h(\alpha(A x_1) + (1-\alpha) A x_2 + b)$$

$$= h(\alpha \underbrace{(A x_1 + b)}_{y_1} + (1-\alpha) \underbrace{(A x_2 + b)}_{y_2})$$

$$\le \alpha h(A x_1 + b) + (1-\alpha) h(A x_2 + b)$$

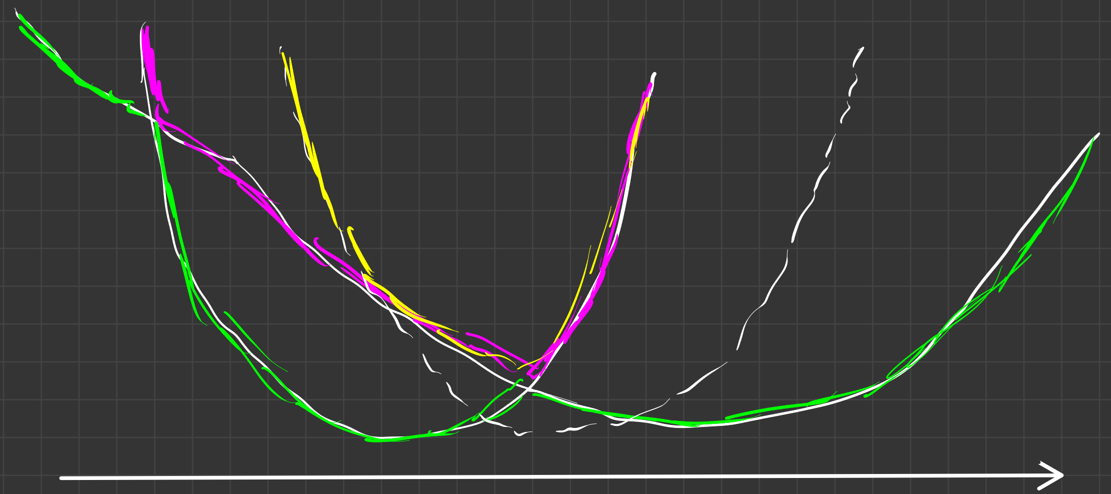
$$= \alpha h(g(x_1)) + (1-\alpha) h(g(x_2))$$

Maxima/Suprema of convex functions

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

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

is convex



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

Suppose $f(\underline{x}, y)$ is a convex fn of $\underline{x}$ for every y.

Then, $g(\underline{x}) = \sup_y f(\underline{x}, y)$ is convex

$$f(x,1) = f_1(x)$$

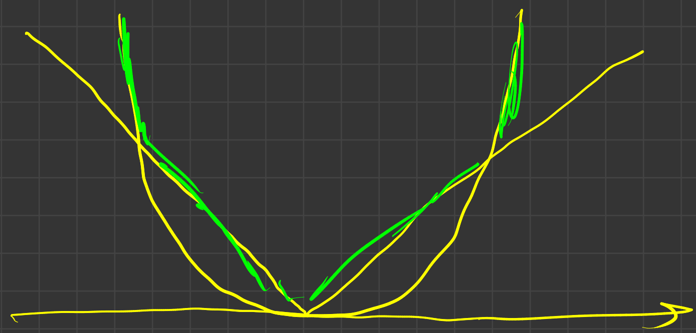
$$f(x,2) = f_2(x)$$

$$\max_{y \in \{1,2\}} f(x,y) = \max\{f_1(x), f_2(x)\}$$

$$f(x,y) = |x|^y$$

$$g(x) = \sup_{y \in \{1,2,\dots\}} f(x,y) = \begin{cases} |x| & \text{if } x \in [-1,1] \\ \infty & \text{if } x \neq 0 \end{cases}$$

$$f(x,y) = \max\{|x|, x^2\}$$



Recall,

$$f_0(x)$$

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

$$h_j(x) = 0 \qquad j = 1, 2, \ldots, k$$

$$L(x, \lambda, \nu) = f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{j=1}^{k} \nu_i h_i(x)$$

For each $x \in \mathbb{R}^n$, $L(x, \lambda, \nu)$ is an affine function of (λ, ν)

$$g(\lambda, \nu) = \inf_{x \in \mathbb{R}^n} L(x, \lambda, \nu) \quad \text{is concave.}$$

(even if the original problem is NOT convex)

For each y, $\quad f(\alpha x_1 + (1-\alpha) x_2, y) \leq \alpha f(x_1, y) + (1-\alpha) f(x_2, y)$

$\forall \alpha \in [0,1]$

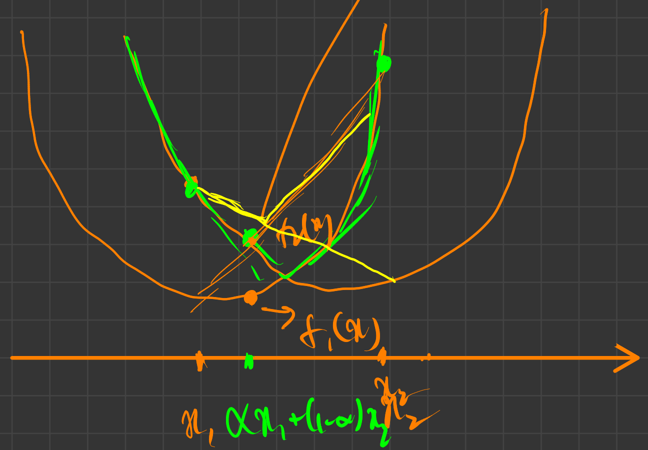
x_1, x_2

$\nabla_x^2 f(x, y)$ PSD $\quad \forall x, y$

Proof: Assume f_1, f_2 are convex

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

Take any $x_1, x_2,\ \alpha \in [0,1]$



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

$$\overset{f_1, f_2 \text{ convex}}{\leq} \max\{\alpha f_1(x_1) + (1-\alpha) f_1(x_2), \alpha f_2(x_1) + (1-\alpha) f_2(x_2)\}$$

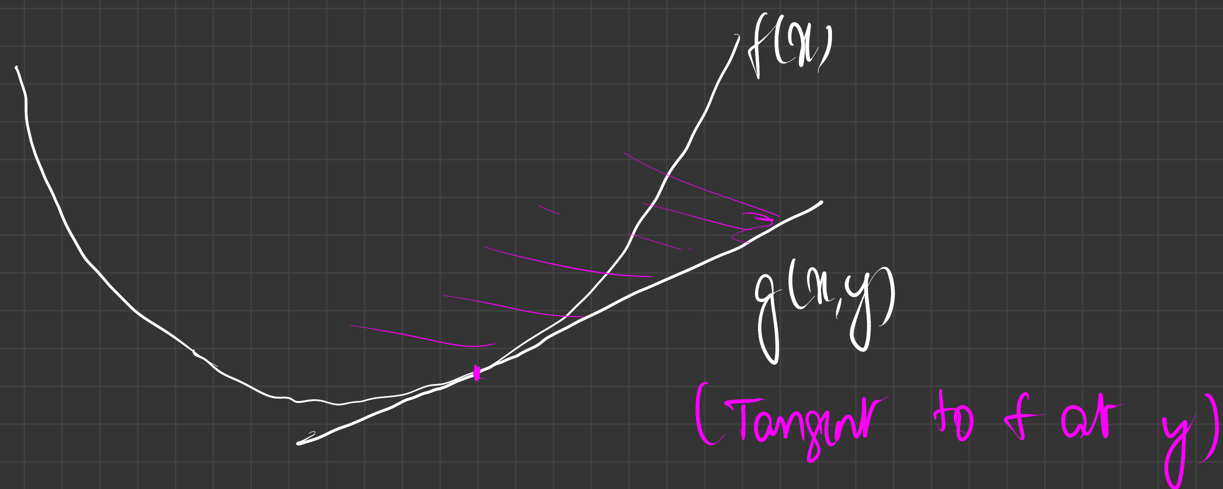
$$\leq \max\{\alpha f_1(x_1), \alpha f_2(x_1)\} + \max\{(1-\alpha) f_1(x_2), (1-\alpha) f_2(x_2)\}$$

$$\left(\max\{a+b, c+d\} \leq \max\{a,c\} + \max\{b,d\}\right)$$

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

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

* Maximum/ supremum of affine functions is convex



$$\bar{g}(x) = \sup_y g(x,y) \leq f(x) \quad \forall x$$
$$=$$

Composition of functions

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

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

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

* Suppose $n = m = 1$.

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

$$f''(x) = g''(h(x))(h'(x))^2 + g'(h(x))\, h''(x)$$

$$f''(x) = g''(h(x))(h'(x))^2 + g'(h(x))\, h''(x)$$

g	h	f
convex, $\uparrow$	convex	convex
convex $\downarrow$	concave	convex
concave $\uparrow$	concave	concave
concave $\downarrow$	convex	concave

Let g be convex ↑ h convex

$$f(\alpha x_1 + (1-\alpha) x_2) = g(h(\alpha x_1 + (1-\alpha) x_2))$$

$$\leq g(\alpha h(x_1) + (1-\alpha) h(x_2))$$

(g is ↑, h is convex)

$$\leq \alpha g(h(x_1)) + (1-\alpha) g(h(x_2))$$

- If g is convex, $e^{g(x)}$ is convex
- If g is concave & +ve, $\log g(x)$ is concave
- $\frac{1}{g(x)}$
- $(g(x))^p$
- $\left(\sum_{i=1}^{k} g_i(x)\right)^p$
- $f: \mathbb{R}^n \to \mathbb{R}$

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

If $k=1$, $f_i(\underline{x}) = x_i$

$f(\underline{x}) = \max\limits_{i=1,\dots n} f_i(\underline{x})$

$$f_I(x) = \sum_{j \in I} x_j \qquad I = \binom{[n]}{k}$$

$x \in \mathbb{R}^3$

$x_1 + x_2$

$x_1 + x_3$

$x_2 + x_3$

$$\max_I f_I(x) = f(x)$$