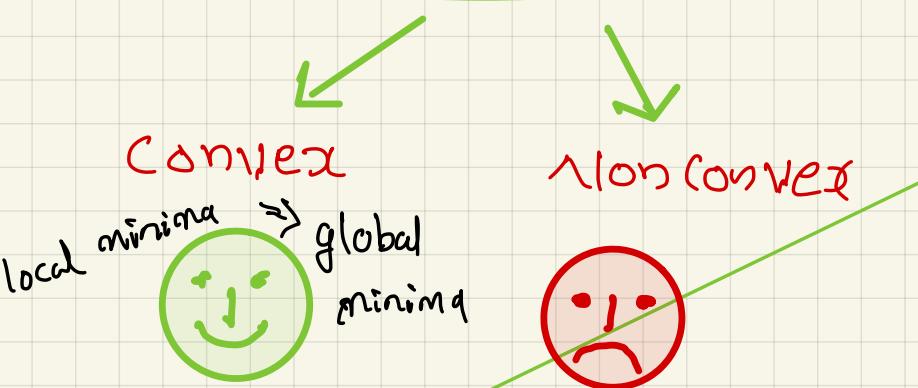
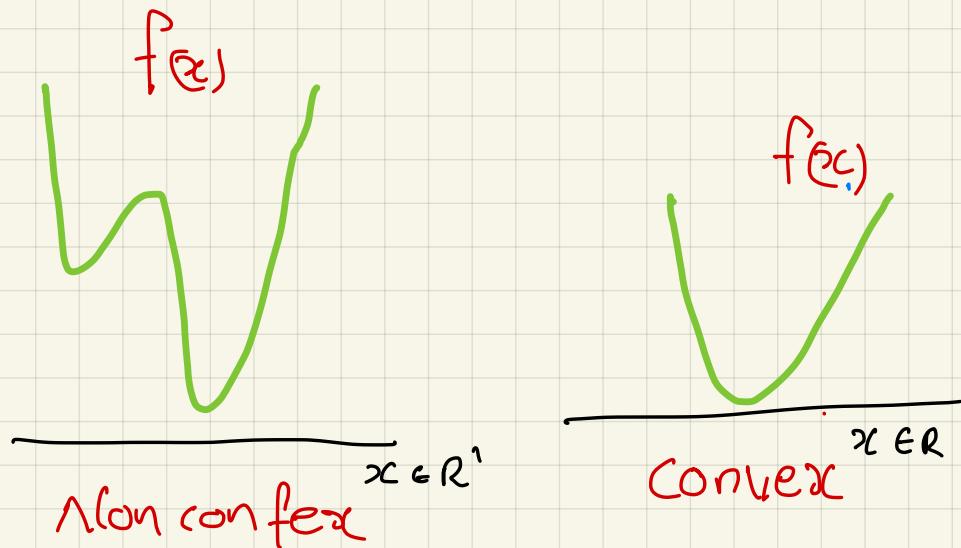


RECAP

Optimization problem



objective function

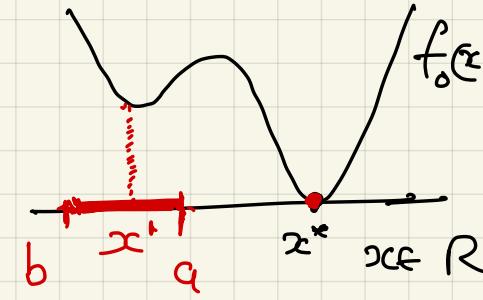


minimize

s.t.

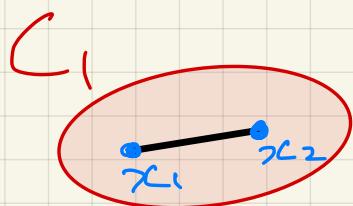
$$f_0(x)$$

$$f_i(x) \leq b_i, \quad i = 1, 2, \dots, m$$

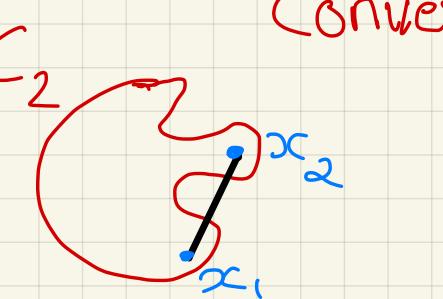
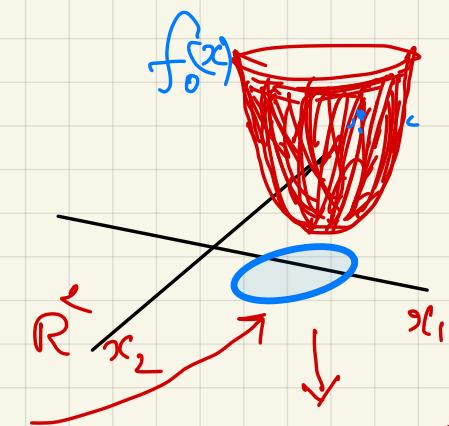


Constraints

forms a set



convex



non-convex

3. Convex functions

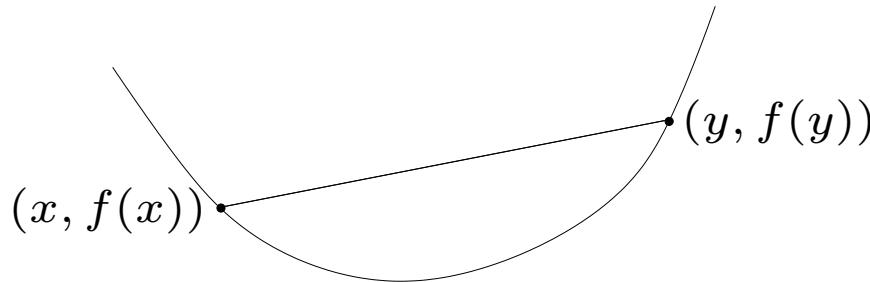
- basic properties and examples
- operations that preserve convexity
- the conjugate function
- quasiconvex functions
- log-concave and log-convex functions
- convexity with respect to generalized inequalities

Definition

$f : \mathbf{R}^n \rightarrow \mathbf{R}$ is convex if $\text{dom } f$ is a convex set and

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

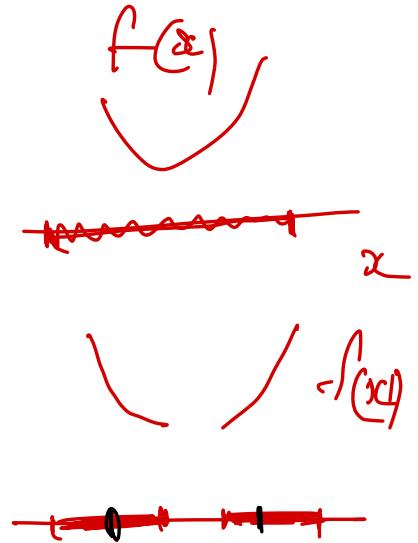
for all $x, y \in \text{dom } f$, $0 \leq \theta \leq 1$

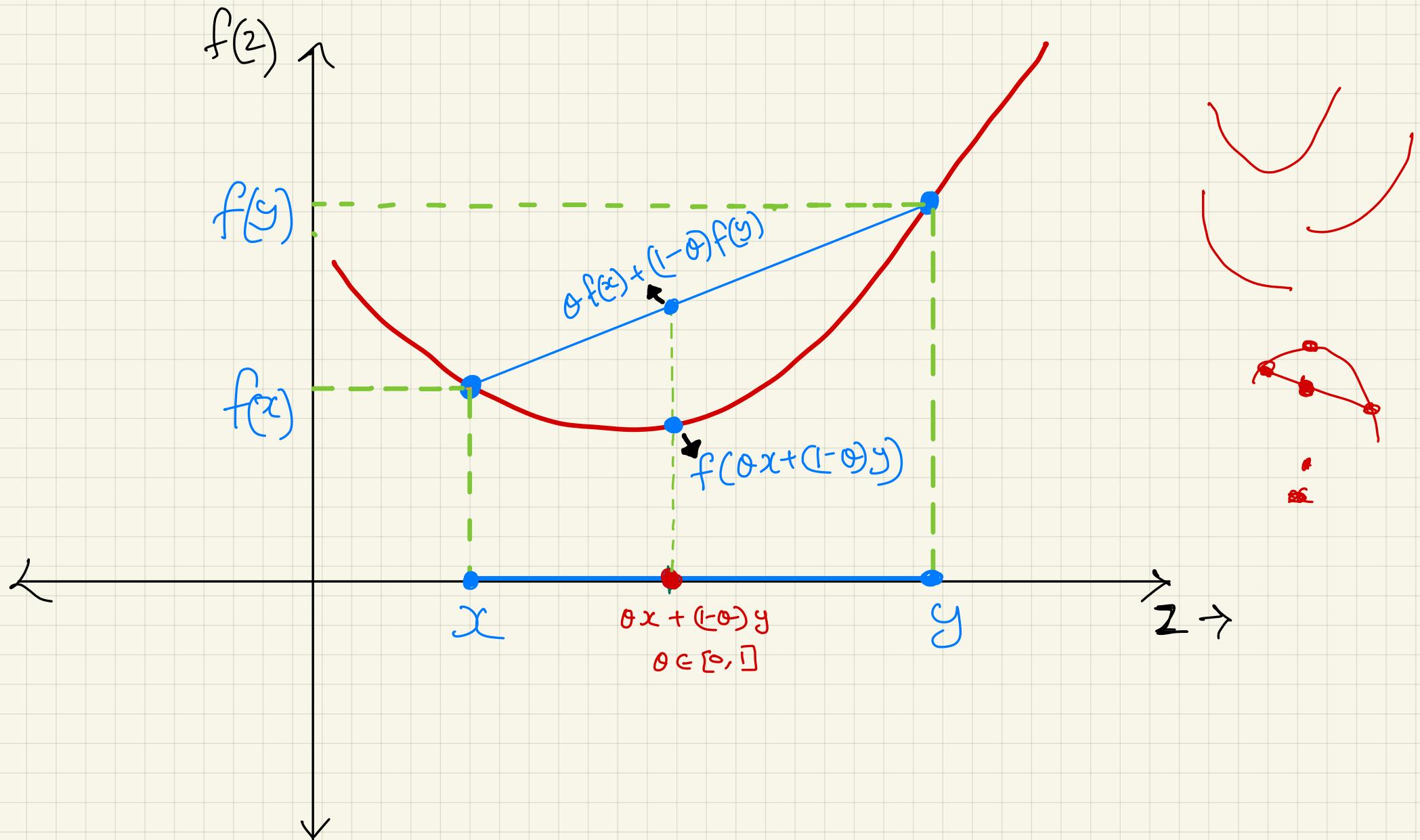


- f is concave if $-f$ is convex
- f is strictly convex if $\text{dom } f$ is convex and

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

for $x, y \in \text{dom } f$, $x \neq y$, $0 < \theta < 1$





affine $f(x) = ax + b$

Examples on \mathbf{R}

convex:

- affine: $ax + b$ on \mathbf{R} , for any $a, b \in \mathbf{R}$
- exponential: e^{ax} , for any $a \in \mathbf{R}$
- powers: x^α on \mathbf{R}_{++} , for $\alpha \geq 1$ or $\alpha \leq 0$
- powers of absolute value: $|x|^p$ on \mathbf{R} , for $p \geq 1$
- negative entropy: $x \log x$ on \mathbf{R}_{++}

concave:

- affine: $ax + b$ on \mathbf{R} , for any $a, b \in \mathbf{R}$
- powers: x^α on \mathbf{R}_{++} , for $0 \leq \alpha \leq 1$
- logarithm: $\log x$ on \mathbf{R}_{++}

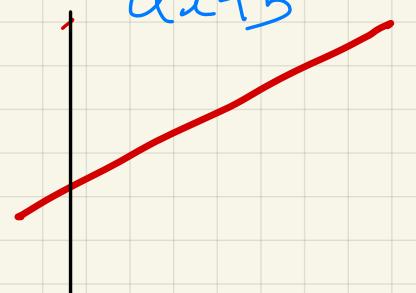
$$\begin{aligned}f(\theta x + (1-\theta)y) &= a(\theta x + (1-\theta)y) + b \\&= \theta(a x + (1-\theta)y) + \\&\quad (1-\theta)b\end{aligned}$$

$$\begin{aligned}&= \theta(ax+b) + (1-\theta)(ay+b) \\&= \theta f(x) + (1-\theta)f(y)\end{aligned}$$

Convex function: Examples

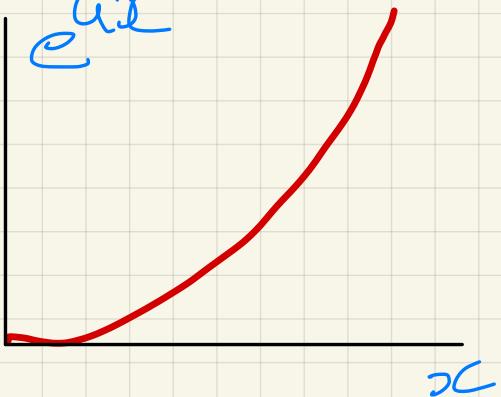
affine

$$ax+b$$



Exponential

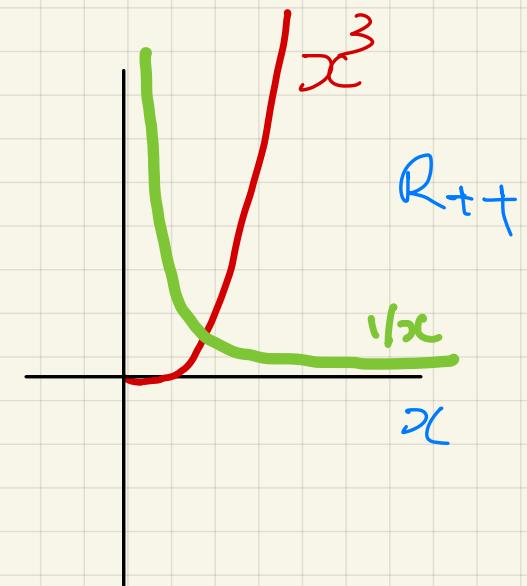
$$e^{ax}$$



$$\alpha \geq 1, \alpha \leq 0$$

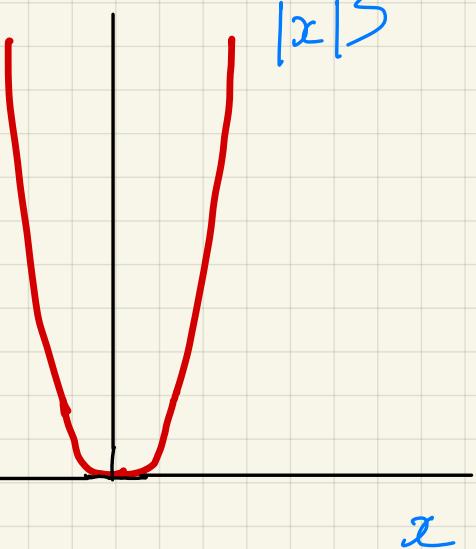
Powers

$$x^\alpha$$



Powers of absolute value

$$|x|^3$$



$$x \log x$$

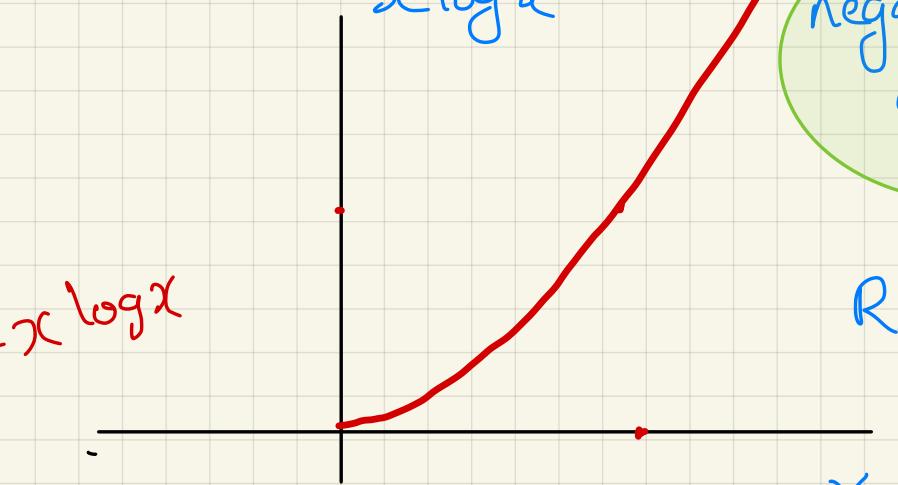
$$-x \log x$$

$$R_{++}$$

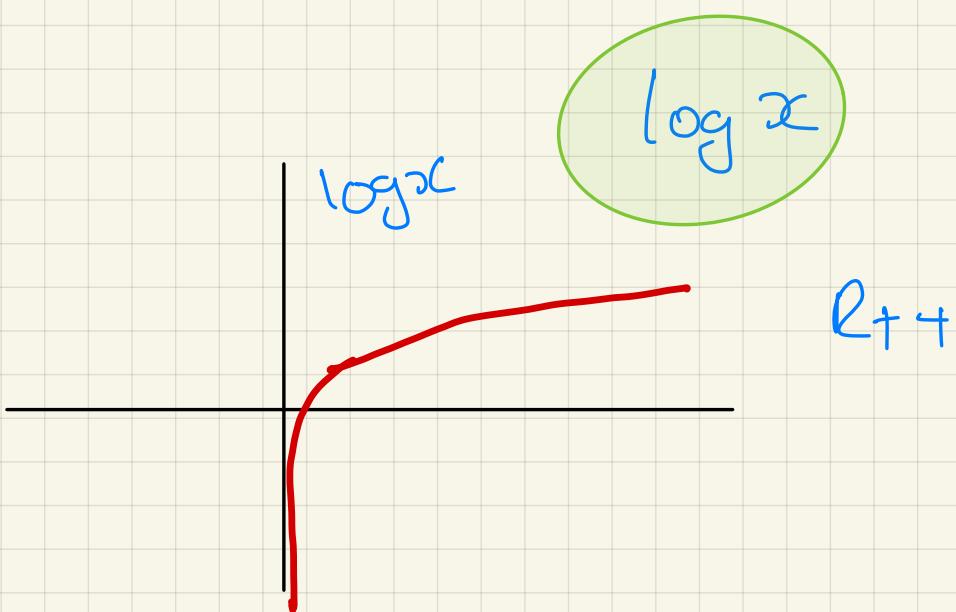
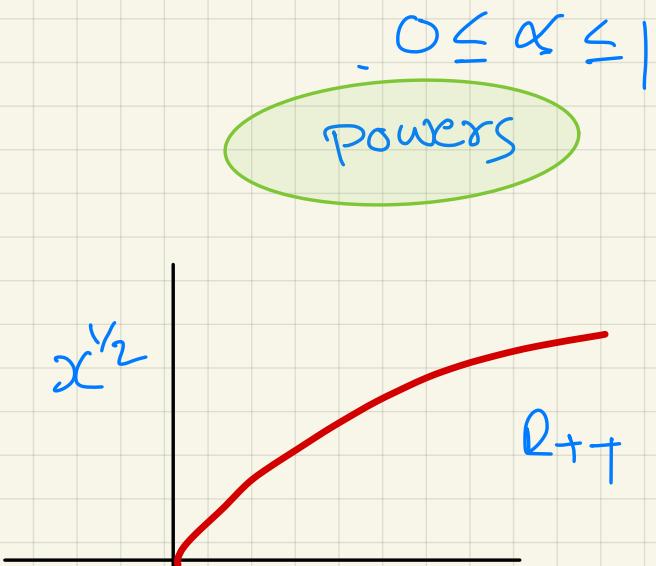
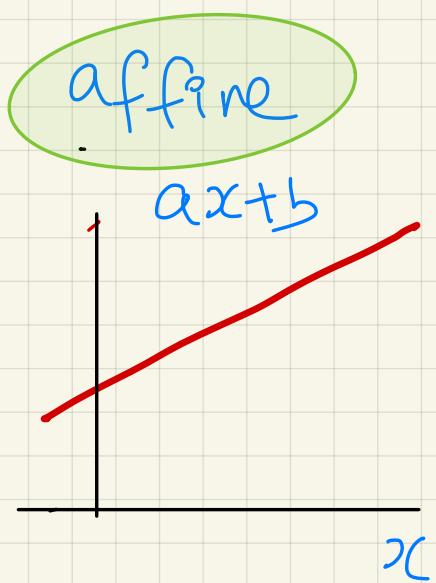
$$x$$

negative entropy

Entropy is Concave.



Concave function : Examples



Examples on \mathbb{R}^n and $\mathbb{R}^{m \times n}$

Affine function

$$\mathbb{R} \quad ax + b$$

$$\mathbb{R}^n \quad a^T x + b$$

$$\mathbb{R}^{m \times n} \quad \text{tr}(A^T x) + b$$

affine functions are convex and concave; all norms are convex

examples on \mathbb{R}^n

- affine function $f(x) = a^T x + b$

- norms: $\|x\|_p = (\sum_{i=1}^n |x_i|^p)^{1/p}$ for $p \geq 1$; $\|x\|_\infty = \max_k |x_k|$

examples on $\mathbb{R}^{m \times n}$ ($m \times n$ matrices)

- affine function

$$A = \begin{pmatrix} a_{11} & & \dots & \\ - & \ddots & & \\ \vdots & & \ddots & \\ x_{1n} & \dots & \dots & a_{nn} \end{pmatrix}$$

$$f(X) = \text{tr}(A^T X) + b = \sum_{i=1}^m \sum_{j=1}^n A_{ij} X_{ij} + b$$

- spectral (maximum singular value) norm

$$f(X) = \|X\|_2 = \sigma_{\max}(X) = (\lambda_{\max}(X^T X))^{1/2}$$

Review: Eigenvalue decomposition
Singular Value decomposition

Convex functions

$$\langle A, B \rangle = \text{tr}(A^T B)$$

$$\Rightarrow AC(:, :)^T B(:, :)$$

A

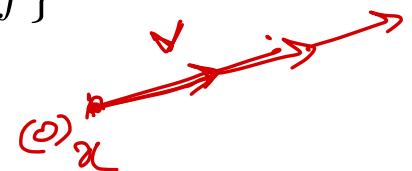
$$\sqrt{\text{eig}(A^T A)}$$

Restriction of a convex function to a line

$f : \mathbf{R}^n \rightarrow \mathbf{R}$ is convex if and only if the function $g : \mathbf{R} \rightarrow \mathbf{R}$,

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

is convex (in t) for any $x \in \text{dom } f$, $v \in \mathbf{R}^n$



can check convexity of f by checking convexity of functions of one variable

example. $f : \mathbf{S}^n \rightarrow \mathbf{R}$ with $f(X) = \log \det X$, $\text{dom } f = \mathbf{S}_{++}^n$

$$g(t) = \log \det(X + tV) = \log \det X + \log \det(I + tX^{-1/2}VX^{-1/2})$$

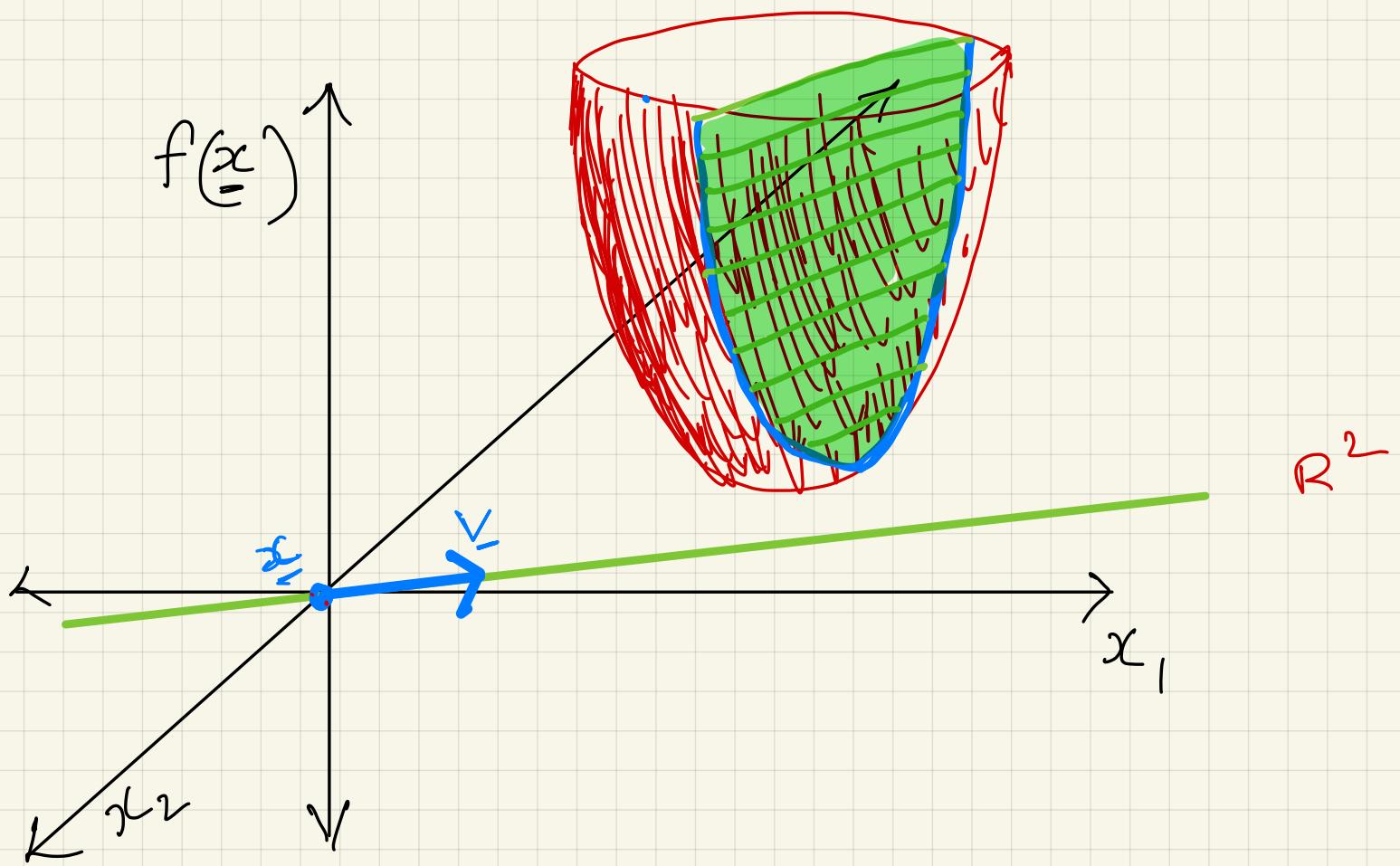
$$= \log \det X + \sum_{i=1}^n \log(1 + t\lambda_i)$$

$X + tV \in \mathbf{S}_{++}^n$

$\leftarrow \begin{matrix} X \in \mathbf{S}_{++}^n \\ t > 0 \end{matrix}$

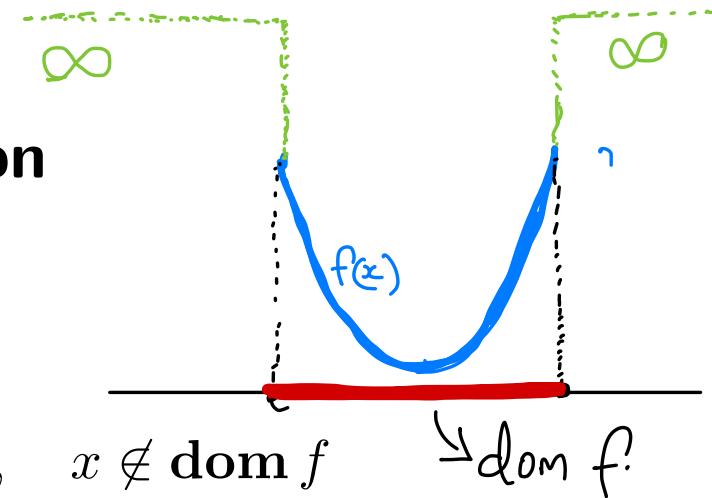
where λ_i are the eigenvalues of $X^{-1/2}VX^{-1/2}$

g is concave in t (for any choice of $X \succ 0$, V); hence f is concave



$$\begin{aligned}
 g(t) &= \log \det(X + tV) \\
 &= \log \det(X^{1/2} X^{1/2} + t X^{1/2} V X^{-1/2} V X^{-1/2}) \\
 &= \log \det[X^{1/2} (I + t X^{-1/2} V X^{-1/2}) X^{1/2}] \stackrel{\det(ABC)}{\rightarrow} \det(A)\det(B)\det(C) \\
 &= \log [(\det X^{1/2}) [\det(I + t X^{-1/2} V X^{-1/2})] (\det X^{1/2})] \\
 &\stackrel{\det(A^n) = (\det A)^n}{=} \log [(\det X)^{1/2} [\det(I + t X^{-1/2} V X^{-1/2})] (\det X)^{1/2}] \\
 &= \log(\det X) + \log[\det(I + t \underbrace{X^{-1/2} V X^{1/2}}_{\text{Eigen Values} = \lambda_1, \lambda_2, \dots, \lambda_n})] \\
 &= \log(\det X) + \log \prod_{i=1}^n (1 + t \lambda_i)
 \end{aligned}$$

$$g(t) = \log(\det X) + \sum_{i=1}^n \underbrace{\log(1 + t \lambda_i)}_{\log \text{ is Concave function}}$$



Extended-value extension

extended-value extension \tilde{f} of f is

$$\tilde{f}(x) = f(x), \quad x \in \text{dom } f, \quad \tilde{f}(x) = \infty, \quad x \notin \text{dom } f$$

$\Downarrow \text{dom } f$

often simplifies notation; for example, the condition

$$0 \leq \theta \leq 1 \implies \tilde{f}(\theta x + (1 - \theta)y) \leq \theta \tilde{f}(x) + (1 - \theta) \tilde{f}(y)$$

(as an inequality in $\mathbf{R} \cup \{\infty\}$), means the same as the two conditions

- $\text{dom } f$ is convex
- for $x, y \in \text{dom } f$,

$$0 \leq \theta \leq 1 \implies f(\theta x + (1 - \theta)y) \leq \theta f(x) + (1 - \theta)f(y)$$

First-order condition

Revise

gradient gives vector
Hessian gives matrix

f is **differentiable** if $\text{dom } f$ is open and the gradient

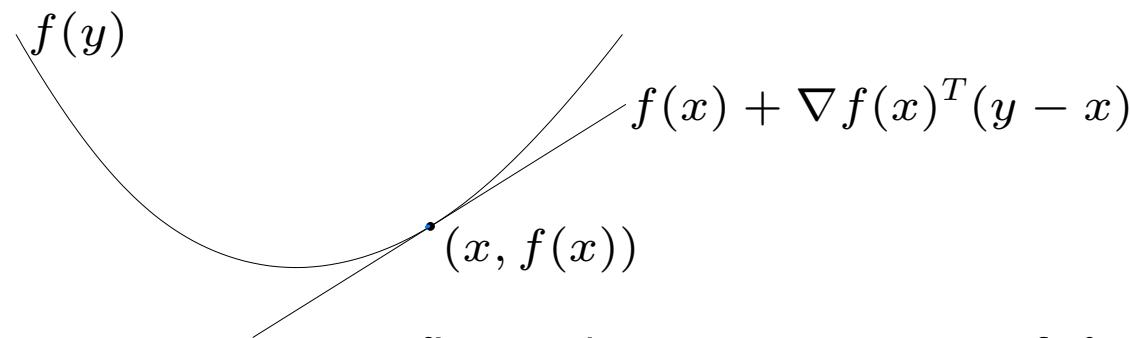
$$\nabla f(x) = \left(\frac{\partial f(x)}{\partial x_1}, \frac{\partial f(x)}{\partial x_2}, \dots, \frac{\partial f(x)}{\partial x_n} \right)$$



exists at each $x \in \text{dom } f$

1st-order condition: differentiable f with convex domain is convex iff

$$f(y) \geq f(x) + \nabla f(x)^T(y - x) \quad \text{for all } x, y \in \text{dom } f$$



first-order approximation of f is global underestimator

Convex functions

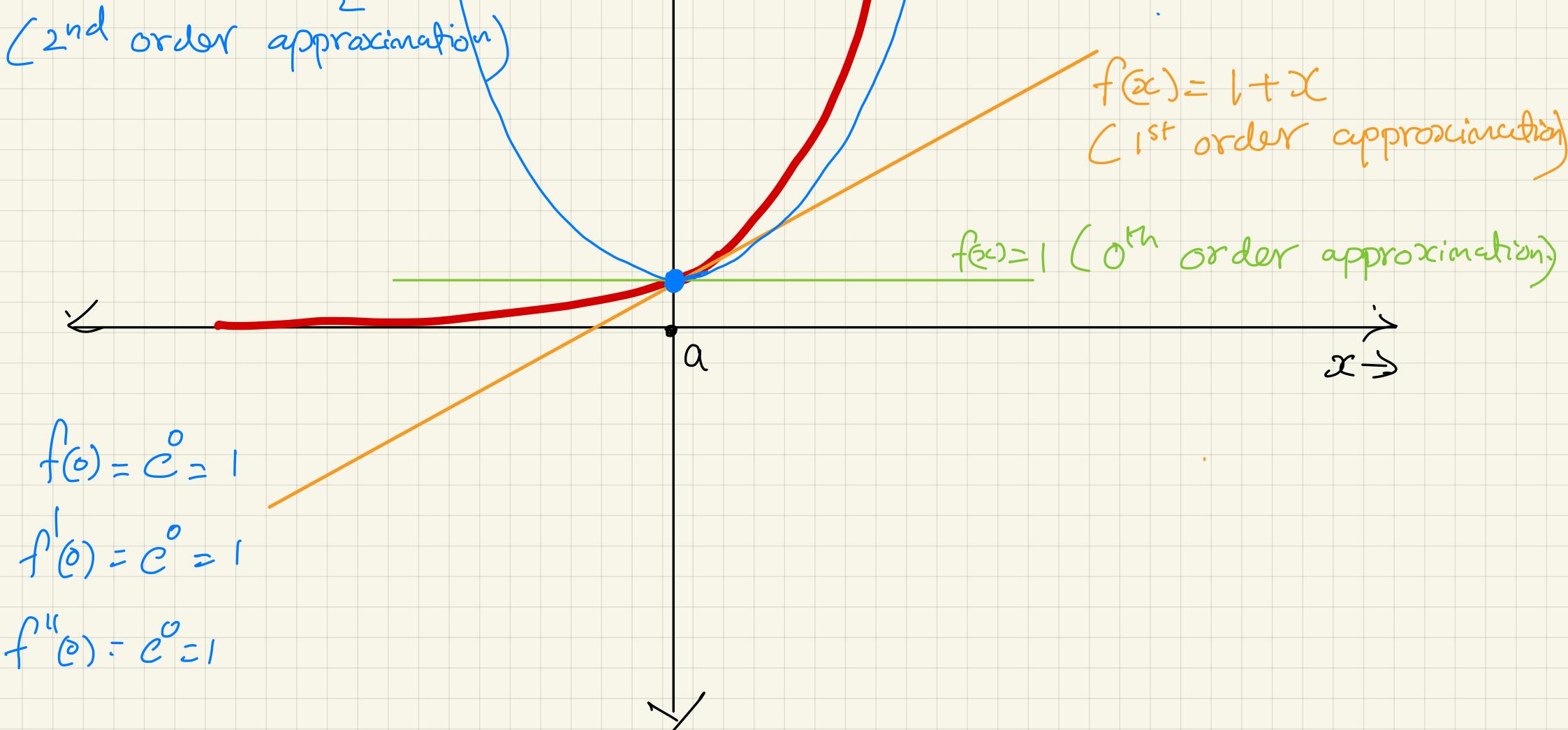
$$f(y) = f(x) + \frac{f'(x)(y-x)}{1!} + \frac{f''(x)(y-x)^2}{2!} + \dots$$

Taylor series

$$f(x) = f(a) + \frac{f'(a)(x-a)}{1!} + \frac{f''(a)(x-a)^2}{2!} + \frac{f'''(a)(x-a)^3}{3!} + \dots$$

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

(2nd order approximation)

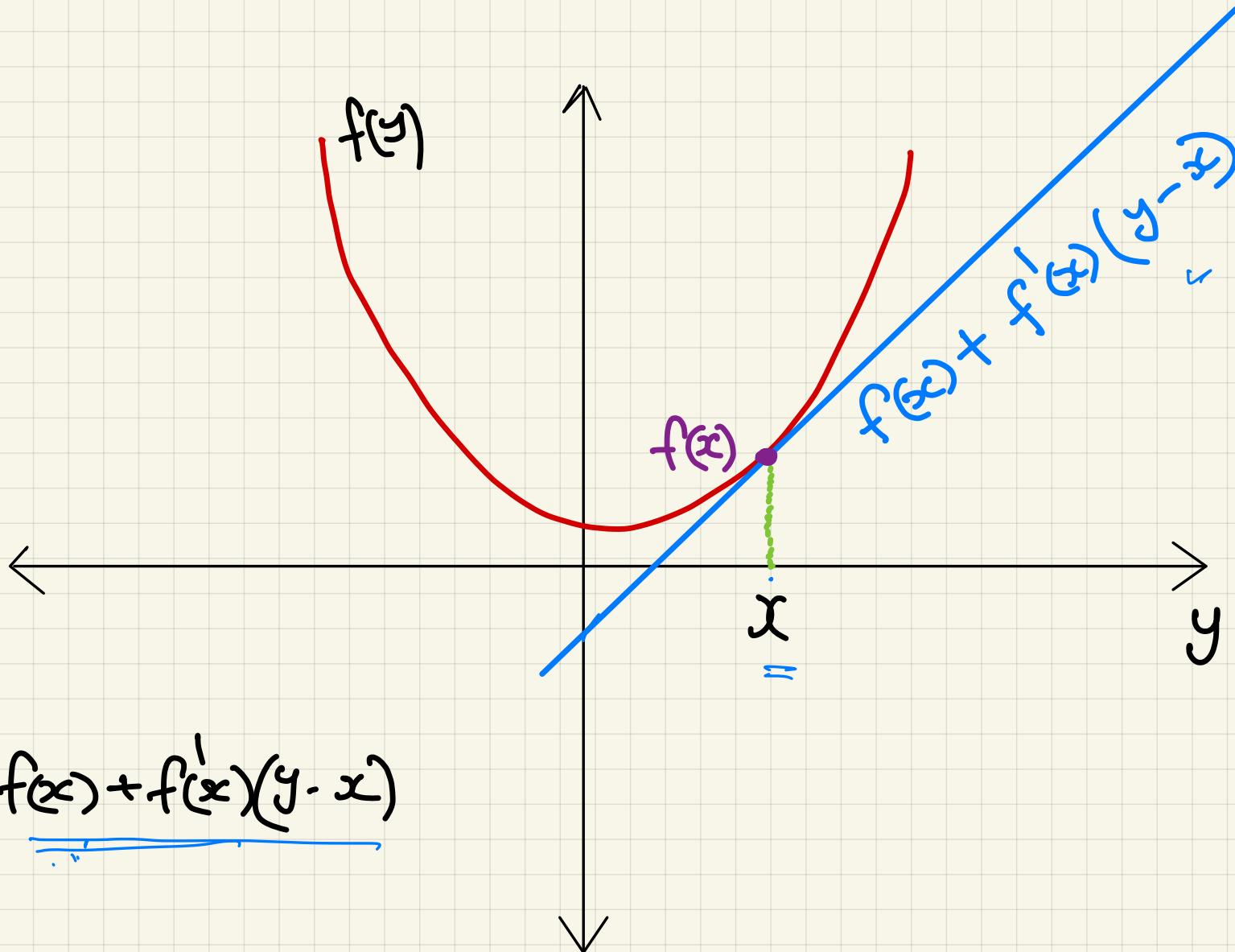


$$f(x) = 1 \quad (0^{\text{th}} \text{ order approximation})$$

$$f(0) = c^0 = 1$$

$$f'(0) = c^1 = 1$$

$$f''(0) = c^2 = 1$$



$$\underline{f(y) \geq f(x) + f'(x)(y - x)}$$

using

local information $\underline{f(x), f'(x)}$.

we can get a global conclusion.

Second-order conditions

f is **twice differentiable** if $\text{dom } f$ is open and the Hessian $\nabla^2 f(x) \in \mathbf{S}^n$,

$$\nabla^2 f(x)_{ij} = \frac{\partial^2 f(x)}{\partial x_i \partial x_j}, \quad i, j = 1, \dots, n,$$

exists at each $x \in \text{dom } f$



$$\begin{bmatrix} \cdot & \cdot & \cdot & \vdots \\ \cdot & \ddots & & \\ \cdot & & \ddots & \\ \cdot & \cdots & \cdots & \frac{\partial^2 f(x)}{\partial x_i \partial x_j} \end{bmatrix}$$

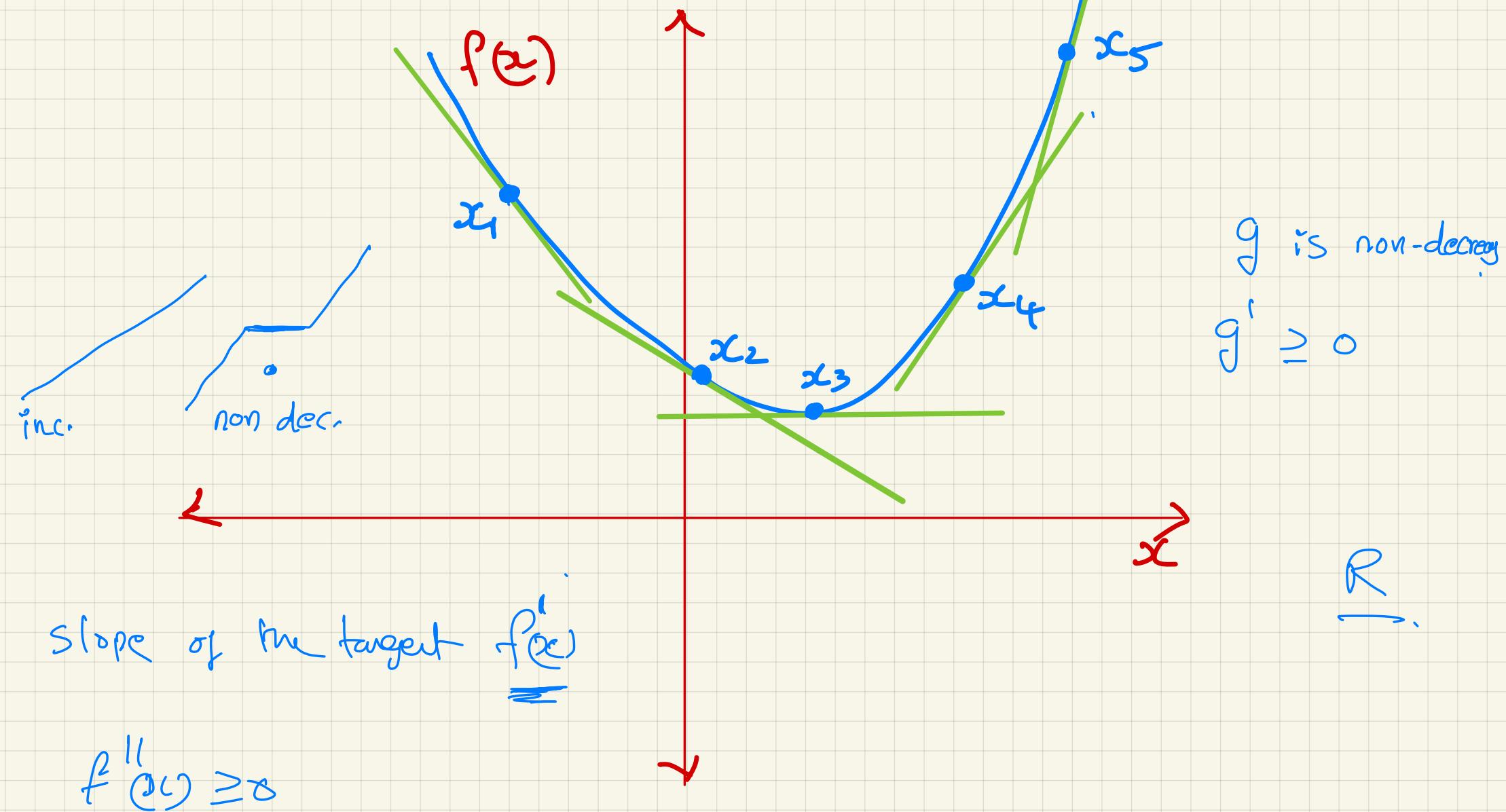
2nd-order conditions: for twice differentiable f with convex domain

- f is convex if and only if

$$\nabla^2 f(x) \succeq 0 \quad \text{for all } x \in \text{dom } f$$

- if $\nabla^2 f(x) \succ 0$ for all $x \in \text{dom } f$, then f is strictly convex

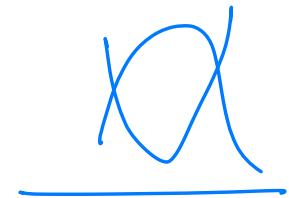
Interpretation of 2nd order condition in \mathbb{R}^1



Slope of the tangent is "increasing" as x increases

Examples

1 **quadratic function:** $f(x) = (1/2)x^T Px + q^T x + r$ (with $P \in \mathbf{S}^n$)



$$\nabla f(x) = Px + q, \quad \nabla^2 f(x) = P$$

~~convex if $P \succeq 0$~~

2 **least-squares objective:** $f(x) = \|\underline{Ax} - \underline{b}\|_2^2 = (\underline{Ax} - \underline{b})^T (\underline{Ax} - \underline{b})$

$$\nabla f(x) = 2A^T(Ax - b), \quad \nabla^2 f(x) = 2A^T A$$

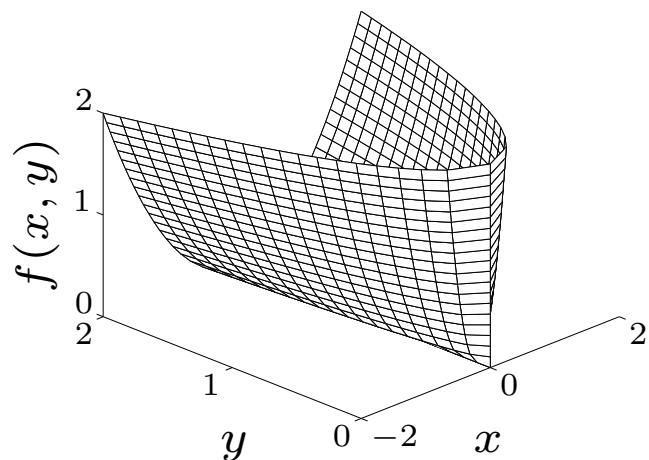
convex (for any A)

$\|\underline{x}\|_2 = \sqrt{\sum_{i=1}^n x_i^2}$

3 **quadratic-over-linear:** $f(x, y) = \underline{x^2/y}$

$$\nabla^2 f(x, y) = \frac{2}{y^3} \begin{bmatrix} y \\ -x \end{bmatrix} \begin{bmatrix} y \\ -x \end{bmatrix}^T \succeq 0$$

convex for $y > 0$



Quadratic over linear

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

$$\nabla f(x) = \begin{bmatrix} \frac{\partial x}{y} & -\frac{x^2}{y^2} \end{bmatrix}$$

$$; \quad \nabla^2 f(x) = \begin{bmatrix} \frac{2}{y} & -\frac{2x}{y^2} \\ -\frac{2x}{y^2} & \frac{+2x^2}{y^3} \end{bmatrix}$$

$$= \frac{2}{y^3} \begin{bmatrix} y^2 & -xy \\ -xy & x^2 \end{bmatrix}$$

$$= \frac{2}{y^3} \begin{bmatrix} y \\ -x \end{bmatrix} \begin{bmatrix} y & -x \end{bmatrix} \geq 0$$

soft max function

log-sum-exp: $f(x) = \log \sum_{k=1}^n \exp x_k$ is convex

$$\frac{1}{e^1} + \frac{5}{e^5} + \frac{10}{e^{10}} \approx 1.0$$

$$\nabla^2 f(x) = \frac{1}{\mathbf{1}^T z} \mathbf{diag}(z) - \frac{1}{(\mathbf{1}^T z)^2} z z^T \quad (z_k = \exp x_k)$$

to show $\nabla^2 f(x) \succeq 0$, we must verify that $v^T \nabla^2 f(x) v \geq 0$ for all v :

$$v^T \nabla^2 f(x) v = \frac{(\sum_k z_k v_k^2)(\sum_k z_k) - (\sum_k v_k z_k)^2}{(\sum_k z_k)^2} \geq 0$$

since $(\sum_k v_k z_k)^2 \leq (\sum_k z_k v_k^2)(\sum_k z_k)$ (from Cauchy-Schwarz inequality)

geometric mean: $f(x) = (\prod_{k=1}^n x_k)^{1/n}$ on \mathbf{R}_{++}^n is concave

(similar proof as for log-sum-exp)

log-sum-exp

$$f(x) = \log \sum_{k=1}^n e^{x_k}$$

$$\nabla f(x) = \frac{1}{\sum_{k=1}^n e^{x_k}} \left[e^{x_1} e^{x_2} \dots e^{x_k} \right] = \underline{\underline{z}}$$

$$\nabla^2 f(x) = \frac{1}{(\underline{\underline{z}})^2} \begin{bmatrix} \frac{\partial^2}{\partial x_1^2} \\ \frac{\partial^2}{\partial x_2^2} \\ \vdots \\ \frac{\partial^2}{\partial x_k^2} \end{bmatrix} - \frac{\underline{\underline{z}}}{(\underline{\underline{z}})^2} \cdot \begin{bmatrix} \frac{\partial (\underline{\underline{z}})}{\partial x_1} \\ \vdots \\ \frac{\partial (\underline{\underline{z}})}{\partial x_k} \end{bmatrix}$$

$$= \frac{\text{diag}(\underline{\underline{z}})}{(\underline{\underline{z}})^2} - \frac{\underline{\underline{z}} \underline{\underline{z}}^T}{(\underline{\underline{z}})^2}$$

$$= \left[(\underline{\underline{z}})^2 \text{diag}(\underline{\underline{z}}) - \underline{\underline{z}} \underline{\underline{z}}^T \right] / (\underline{\underline{z}})^2.$$

to check positive definiteness

$$\underline{z}^T \underline{V} = [z_1 v_1 z_2 v_2 \dots z_k v_k]$$

$$\underline{V}^T \nabla^2 f(\underline{z}) \underline{V} = \left(\underline{V}^T (\underline{1}^T \underline{z}) \text{diag}(\underline{z}) \underline{V} - \underbrace{\underline{V}^T \underline{z} \underline{z}^T \underline{V}}_{(\underline{z}^T \underline{V})^T (\underline{z}^T \underline{V})} \right) / (\underline{1}^T \underline{z})^2$$

$$= \left[(\underline{1}^T \underline{z}) \begin{bmatrix} V_1 & V_2 & \dots & V_k \end{bmatrix} \begin{bmatrix} z_1 & 0 & 0 & \dots & 0 \\ 0 & z_2 & 0 & 0 & 0 \\ \vdots & 0 & 0 & \dots & z_k \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \\ \vdots \\ V_k \end{bmatrix} \right] - \sum_{i=1}^k (z_i V_i)^2 / \left(\sum_{i=1}^k z_i \right)^2$$

$$\leq \left[\left(\sum z_i \right) \sum_{i=1}^k z_i V_i^2 - \sum_{i=1}^k (z_i V_i)^2 \right] / \left(\sum_{i=1}^k z_i \right)^2$$

$\sum \circ$

Cauchy-Schwarz inequality.
 $|\langle \underline{u}, \underline{v} \rangle| \leq \|\underline{u}\|_2 \|\underline{v}\|_2$

$$\underline{u} = [z_1^{v_1} z_2^{v_2} \dots z_k^{v_k}]$$

$$\underline{v} = [z_1^{v_2} v_1 \dots z_k^{v_2} v_k]$$

Epigraph and sublevel set

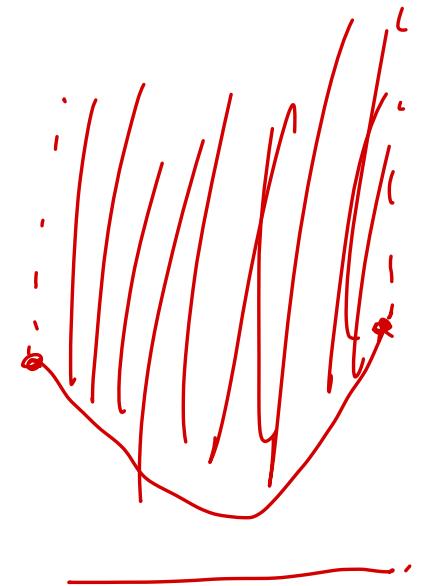
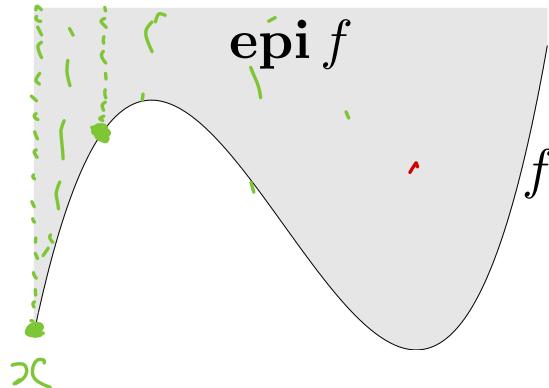
α -sublevel set of $f : \mathbf{R}^n \rightarrow \mathbf{R}$:

$$C_\alpha = \{x \in \text{dom } f \mid f(x) \leq \alpha\}$$

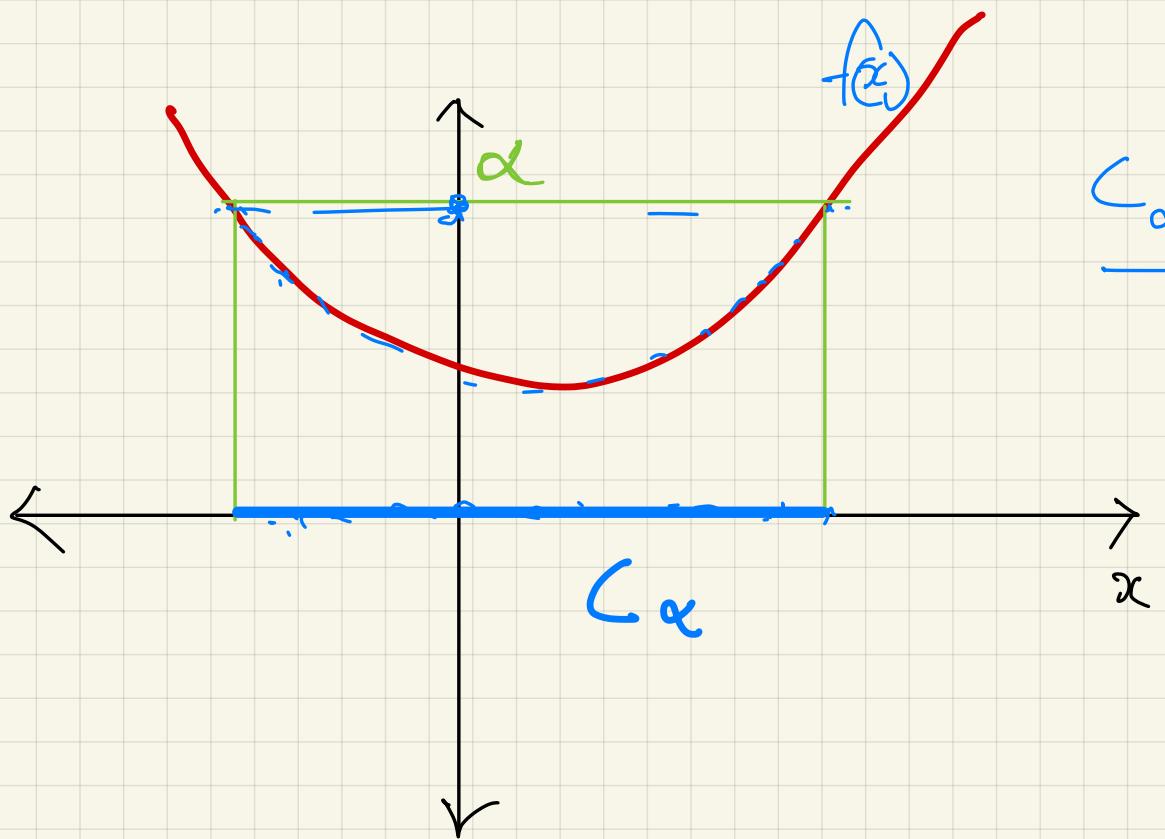
sublevel sets of convex functions are convex (converse is false)

epigraph of $f : \mathbf{R}^n \rightarrow \mathbf{R}$:

$$\text{epi } f = \{(x, t) \in \mathbf{R}^{n+1} \mid x \in \text{dom } f, f(x) \leq t\}$$

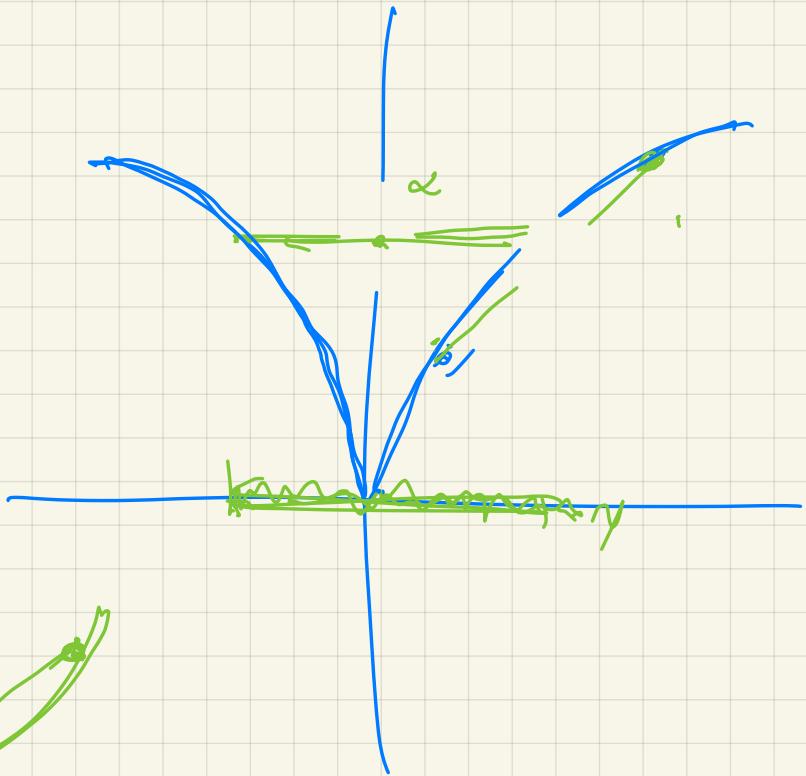


f is convex if and only if $\text{epi } f$ is a convex set



$$L_\alpha = \{x \in \text{dom } f, f(x) \leq \alpha\}$$

Convex
not differentiable.



1st and 2nd order Conditions (RECAP)

function $f : \mathbb{R}^n \rightarrow \mathbb{R}$.

f is differentiable

f is not differentiable

① 1st-order condition.

f is differentiable with convex domain

then f is convex iff

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

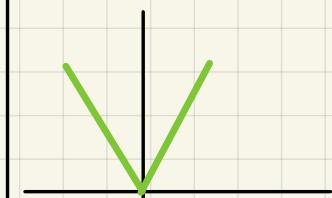
② 2nd-order condition.

f is twice differentiable with convex domain

then f is convex iff

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

- * f can be convex or non convex.
- * Convexity is checked using other methods.



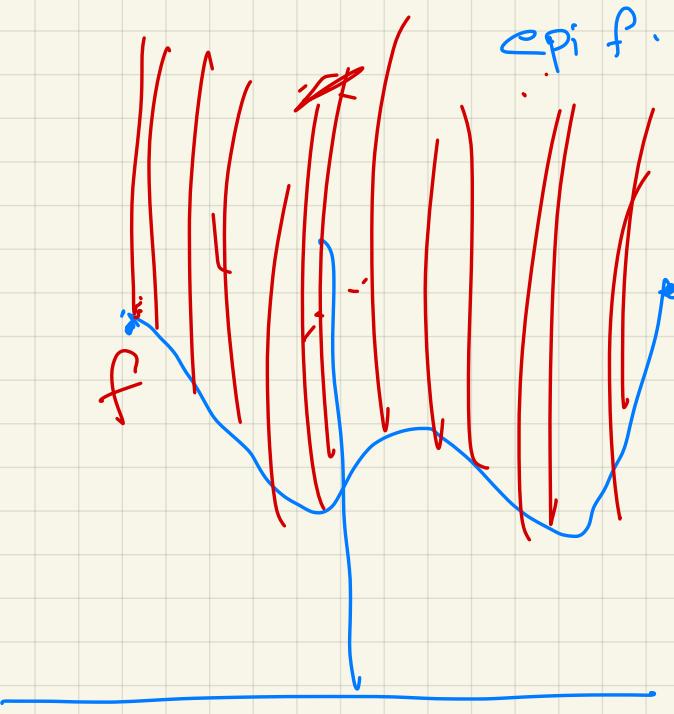
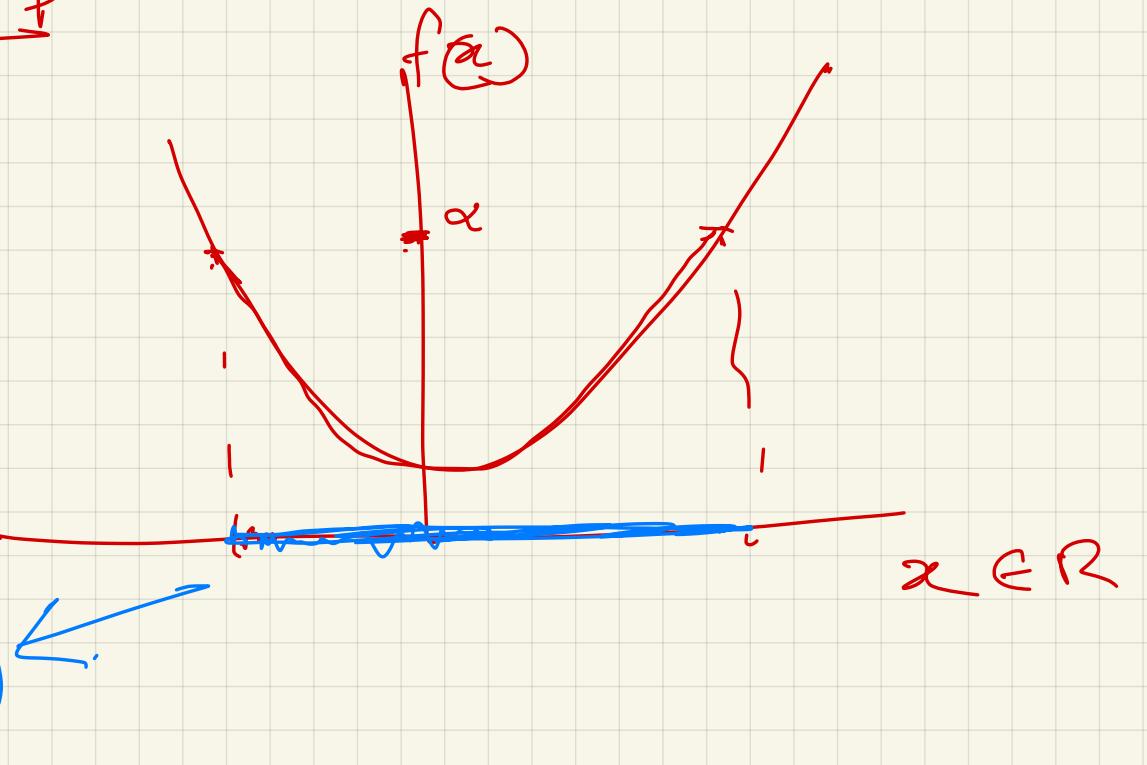
Convex



Non convex

Recap

α -sublevel set



Jensen's inequality

basic inequality: if f is convex, then for $0 \leq \theta \leq 1$,

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

$f(Ez) \leq E(f(z))$

extension: if f is convex, then

$$f(\mathbf{E} z) \leq \mathbf{E} f(z)$$

for any random variable z

basic inequality is special case with discrete distribution

$$\text{prob}(z = x) = \theta, \quad \text{prob}(z = y) = 1 - \theta$$

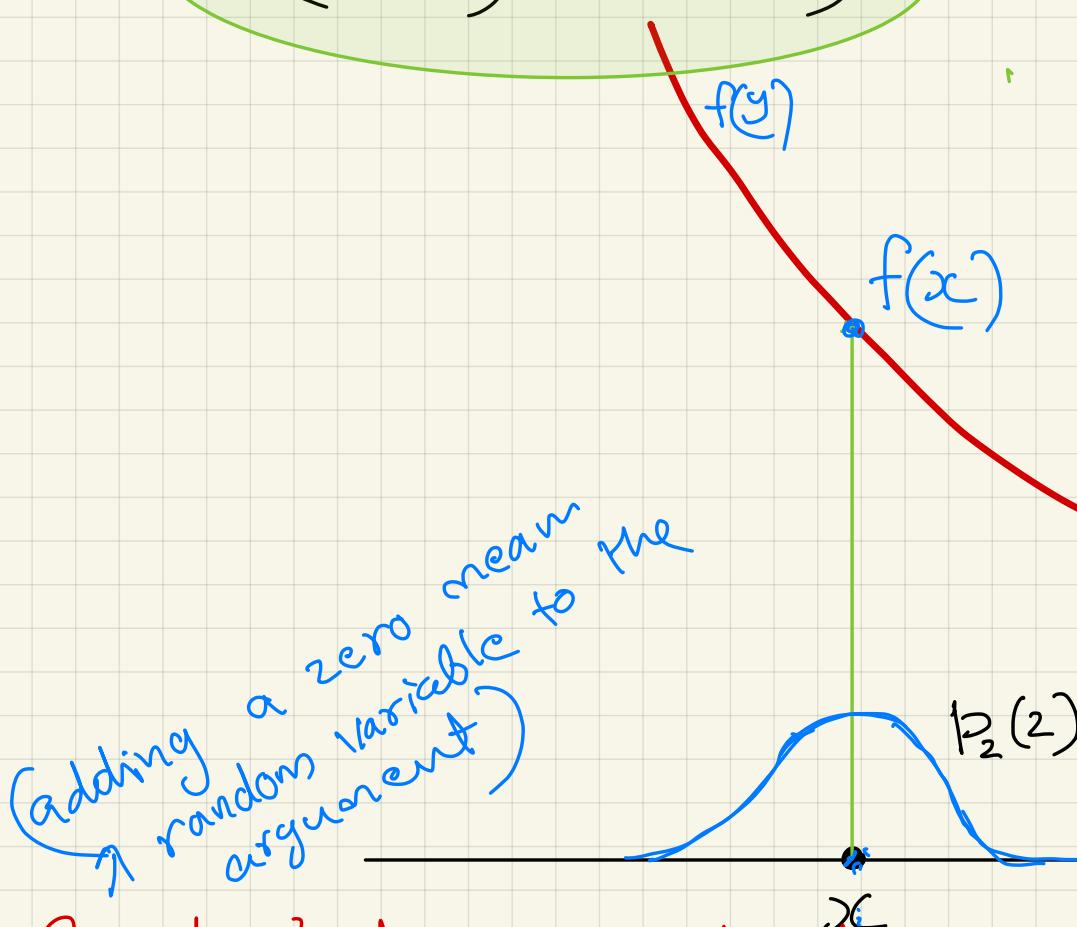
$$f(\theta_1 \underline{x}_1 + \theta_2 \underline{x}_2 + \dots + \theta_n \underline{x}_n) \leq \theta_1 f(\underline{x}_1) + \theta_2 f(\underline{x}_2) + \dots + \theta_n f(\underline{x}_n)$$

$$f\left(\int_{\text{dom } f} \underline{x} p(\underline{x}) d\underline{x}\right) \leq \int_{\text{dom } f} f(\underline{x}) p(\underline{x}) d\underline{x}$$

$$\sum \theta_i = 1 \\ \theta_i \in [0, 1]$$

$$\Rightarrow f(E[\underline{x}]) \leq E(f(\underline{x}))$$

convex function.



$$f(x+z), z \sim N(0, 1)$$

$$f(E(x+z)) \leq E f(x+z)$$

$$\Rightarrow f(\underbrace{E}_{\text{un}} \underbrace{x + Ez}_{0}) \leq E f(x+z)$$

$$\Rightarrow f(x) \leq E f(x+z)$$

Randomization cannot decrease the value of a crx function on average

Operations that preserve convexity

practical methods for establishing convexity of a function

1. verify definition (often simplified by restricting to a line)
2. for twice differentiable functions, show $\nabla^2 f(x) \succeq 0$
3. show that f is obtained from simple convex functions by operations that preserve convexity
 - nonnegative weighted sum
 - composition with affine function
 - pointwise maximum and supremum
 - composition
 - minimization
 - perspective

Positive weighted sum & composition with affine function

nonnegative multiple: αf is convex if f is convex, $\alpha \geq 0$

sum: $f_1 + f_2$ convex if f_1, f_2 convex (extends to infinite sums, integrals)

composition with affine function: $f(Ax + b)$ is convex if f is convex

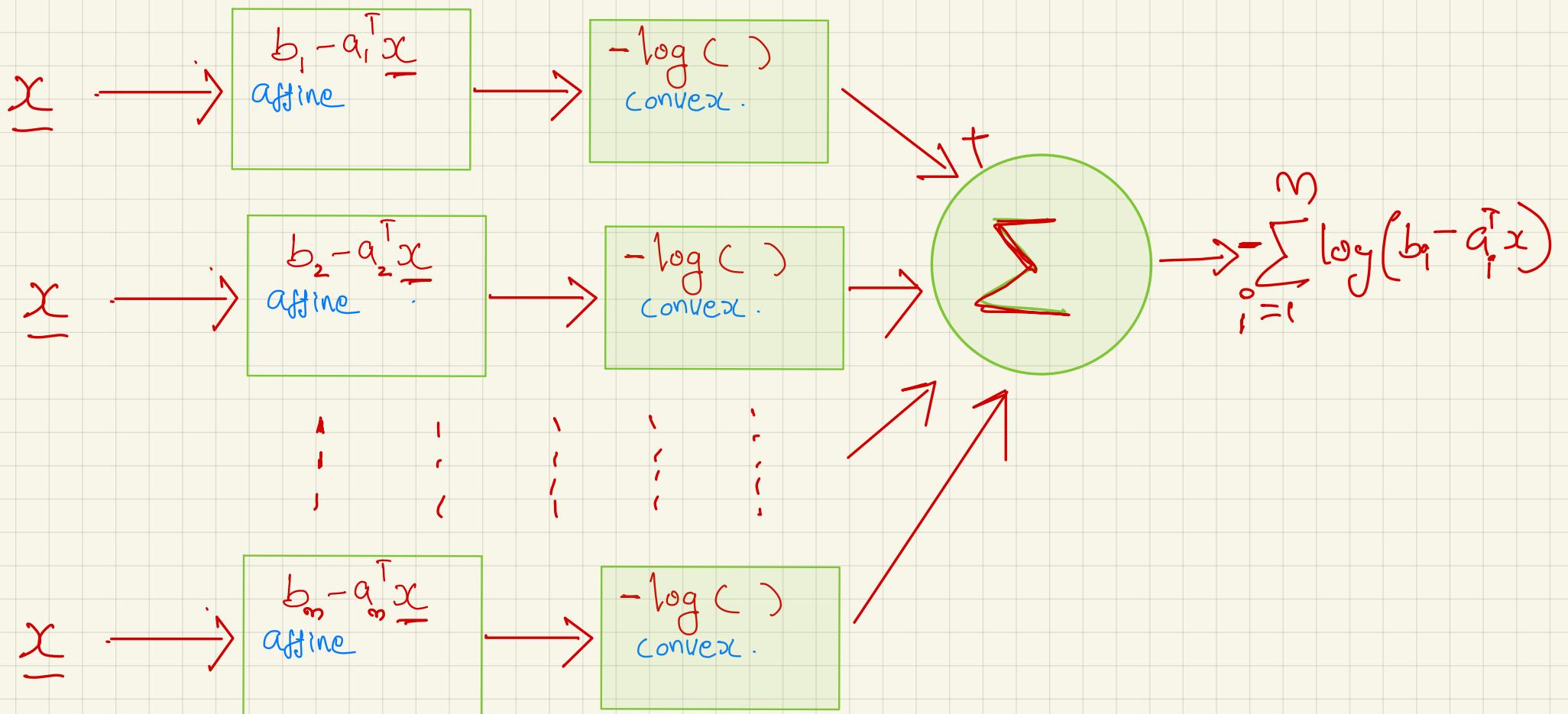
examples

- log barrier for linear inequalities

$$f(x) = - \sum_{i=1}^m \log(b_i - a_i^T x), \quad \text{dom } f = \{x \mid a_i^T x < b_i, i = 1, \dots, m\}$$

- (any) norm of affine function: $f(x) = \|Ax + b\|$

$$\|Ax + b\|^2$$



Pointwise maximum

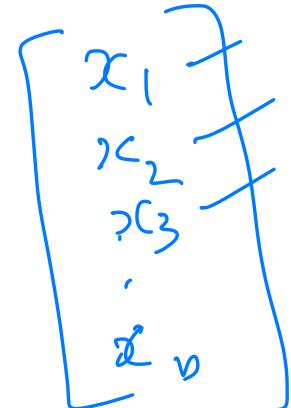
if f_1, \dots, f_m are convex, then $f(x) = \max\{f_1(x), \dots, f_m(x)\}$ is convex

with $\text{dom } f = \bigcap_{i=1}^m \text{dom } f_i$

examples

- piecewise-linear function: $f(x) = \max_{i=1,\dots,m} (a_i^T x + b_i)$ is convex
- sum of r largest components of $x \in \mathbf{R}^n$:

$$f(x) = x_{[1]} + x_{[2]} + \cdots + x_{[r]}$$

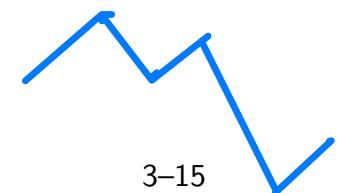


is convex ($x_{[i]}$ is i th largest component of x)

proof:

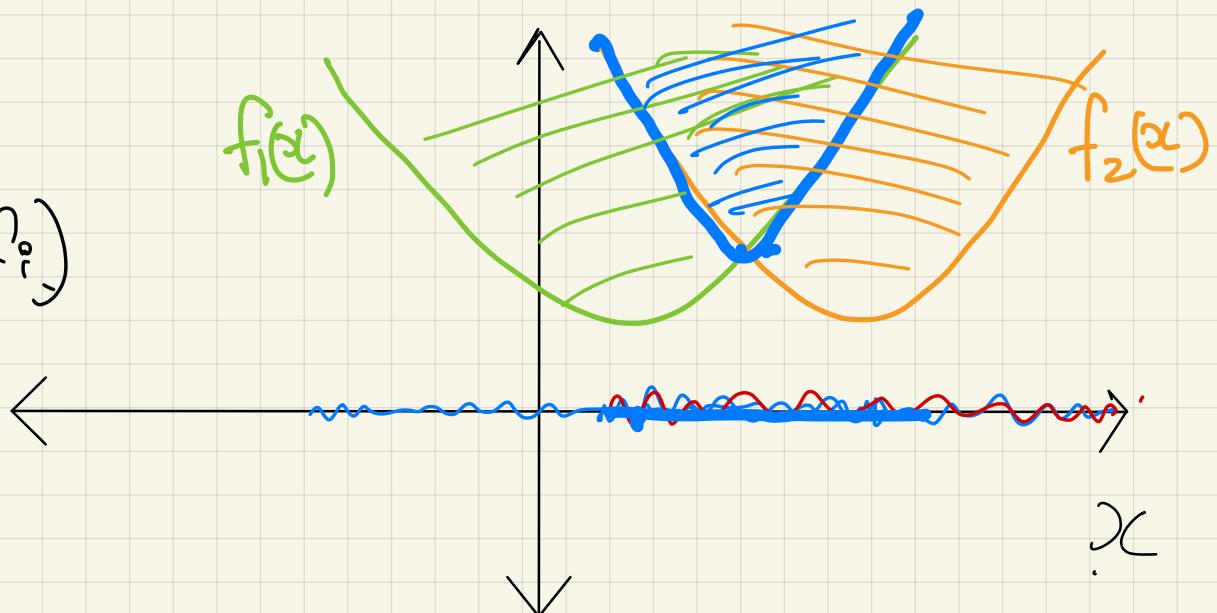
$$f(x) = \max\{x_{i_1} + x_{i_2} + \cdots + x_{i_r} \mid 1 \leq i_1 < i_2 < \cdots < i_r \leq n\}$$

NOTE: All piecewise-linear functions are not convex



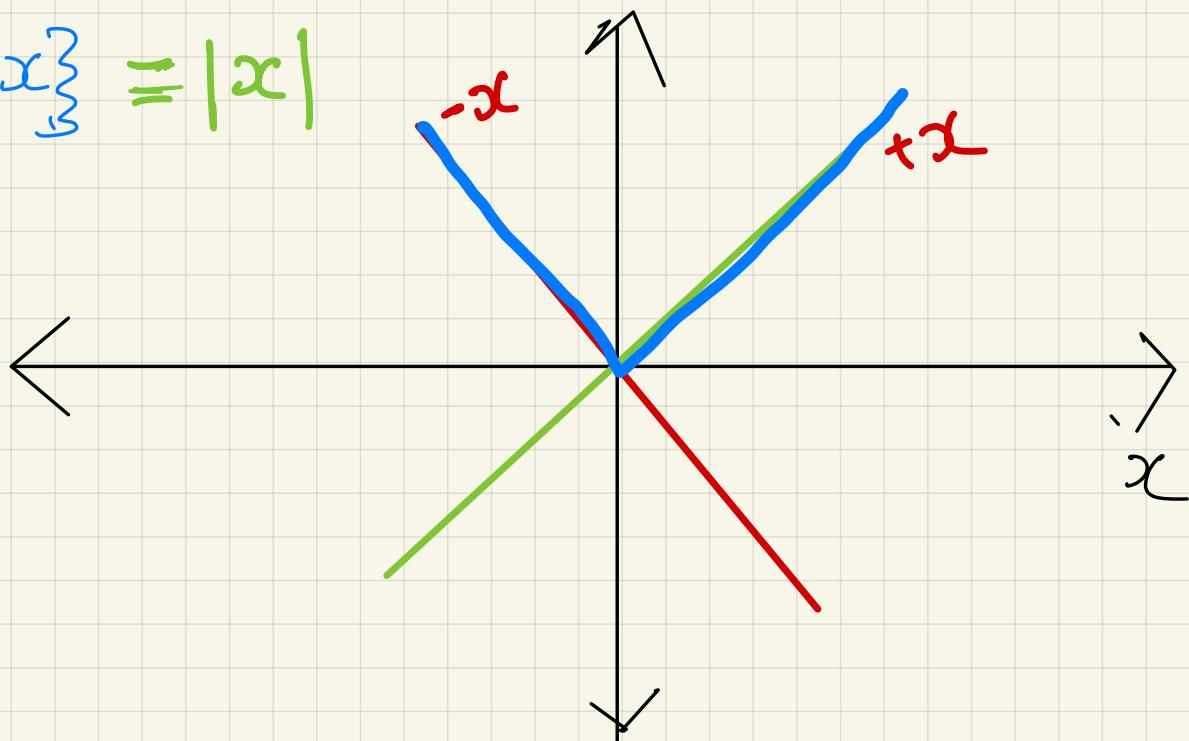
$$f(x) = \max(f_1(x), f_2(x))$$

$$\text{cpi}(\max_i f_i) = ? \text{ cpi}(f_i)$$



piece-wise linear function.

$$\text{e.g. } f(x) = \max\{x, -x\} \equiv |x|$$



Point wise maximum preserves convexity (proof)

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

$$\theta \in [0, 1]$$

$$f(\theta \underline{x} + (1-\theta) \underline{y}) = \max \{ f_1(\theta \underline{x} + (1-\theta) \underline{y}), f_2(\theta \underline{x} + (1-\theta) \underline{y}) \}$$

$$\leq \max \{ \theta f_1(\underline{x}) + (1-\theta) f_1(\underline{y}), \theta f_2(\underline{x}) + (1-\theta) f_2(\underline{y}) \}$$

(convexity of f_1 and f_2)

$$\max \{ a+b, c+d \} = 7$$

$$\leq \max \{ a, c \} + \max \{ b, d \}$$

$$2+6=8$$

$$\leq \max \{ \theta f_1(\underline{x}), \theta f_2(\underline{x}) \} + \max \{ (1-\theta) f_1(\underline{y}), (1-\theta) f_2(\underline{y}) \}$$

$$= \theta \max \{ f_1(\underline{x}), f_2(\underline{x}) \} + (1-\theta) \max \{ f_1(\underline{y}), f_2(\underline{y}) \}$$

$$= \theta f(\underline{x}) + (1-\theta) f(\underline{y})$$

Sum of r largest component $\equiv \max_{\underline{x}} \underline{a}_1^T \underline{x}$

$$\begin{bmatrix} 1 & 1 & 0 & 1 & \dots & 0 \end{bmatrix} \underbrace{\underline{a}_1^T}_{r=3} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ \vdots \\ x_n \end{bmatrix} = x_1 + x_2 + x_4.$$

choose $\binom{n}{r}$ different
 a_i 's by varying the
positions of '1's

$\binom{5}{3}$

Pointwise supremum

if $f(x, y)$ is convex in x for each $y \in \mathcal{A}$, then

A need not convex

$$g(x) = \sup_{y \in \mathcal{A}} f(x, y)$$

is convex



examples

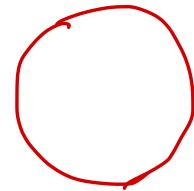
- support function of a set C : $S_C(x) = \sup_{y \in C} y^T x$ is convex $S_C(x)$
- distance to farthest point in a set C :

$$f(x) = \sup_{y \in C} \|x - y\|$$

- maximum eigenvalue of symmetric matrix: for $X \in \mathbf{S}^n$,

$$f(X) = \lambda_{\max}(X) = \sup_{\|y\|_2=1} y^T X y$$

≤ 1



Supremum

Supremum of a set: smallest upper bound of the set.

$$S \subseteq \mathbb{R}$$

Upper bound of S : b is an upper bound of S if
for each $x \in S$, $x \leq b$

e.g①: $S = [1, 5)$

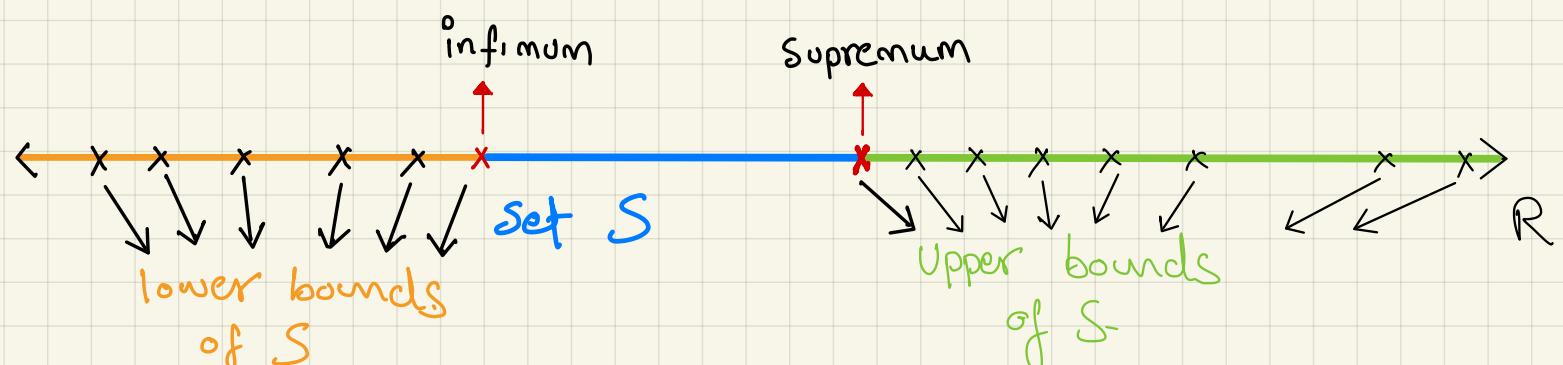
$\max S = 2$

$\sup S = 5$

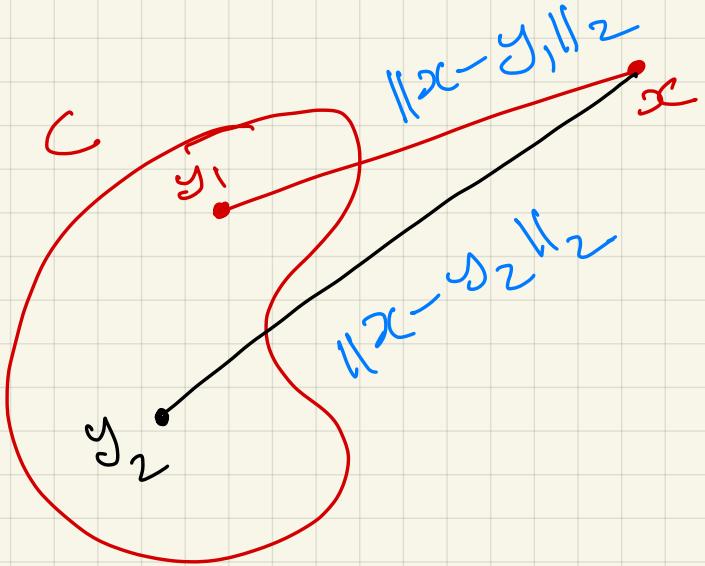
e.g.②: $S = [1, 5]$

$\max S = 5$

$\sup S = 5$



Infimum: largest lower bound.



$$\sup_{y \in C} \|x - y\|_2$$

$\underline{y}_1^T \times y_1$
 $\underline{y}_2^T \times y_2$
 \vdots
 \vdots

} \rightarrow linear function in X
 ↓
 Convex

$$\lambda_{\max}(X) = \sup_{\|\underline{y}\|_2=1} \underline{y}^T X \underline{y}$$

Composition with scalar functions

composition of $g : \mathbf{R}^n \rightarrow \mathbf{R}$ and $h : \mathbf{R} \rightarrow \mathbf{R}$:

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

f is convex if g convex, h convex, \tilde{h} nondecreasing
 g concave, h convex, \tilde{h} nonincreasing

$$f(Ax+b)$$

\downarrow \downarrow

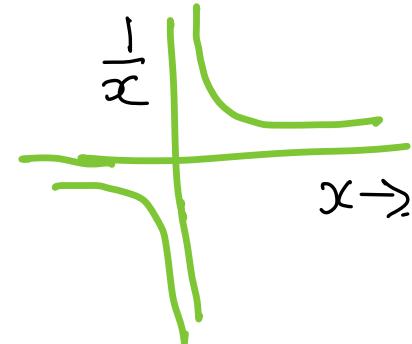
h g

h	g
\exp	$g(\cdot)$
↑	↑
cvx	cvx

- proof (for $n = 1$, differentiable g, h)

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

- note: monotonicity must hold for extended-value extension \tilde{h}



examples

- $\exp g(x)$ is convex if g is convex
- $1/g(x)$ is convex if g is concave and positive

① g is concave.
 ② g is +ve $\Rightarrow h$ is \downarrow
 h is cvx

$$h(g(y)) = \frac{1}{g(y)}$$

$$h(y) = 1/y$$

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

$$\begin{aligned} g &: \mathbb{R}^n \rightarrow \mathbb{R} \\ h &: \mathbb{R} \rightarrow \mathbb{R} \end{aligned}$$

Proof ($n=1$)

$Ax+b$

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

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

for $f(x)$ to be convex.

≥ 0

$h''() \geq 0$
 $\Rightarrow h$ is convex

$$g''() \geq 0, h'() \geq 0$$

g is convex
 h is non-decreasing

$$g''() \leq 0, h'() \leq 0$$

g is concave
 h is non-increasing

$$f(Ax+b)$$

$$h \rightarrow f$$

$$g(x) = Ax + b$$

$$\begin{matrix} h \text{ cvx} \\ g \text{ cvx} \end{matrix} \Rightarrow \text{cvx}$$

Vector composition

composition of $g : \mathbf{R}^n \rightarrow \mathbf{R}^k$ and $h : \mathbf{R}^k \rightarrow \mathbf{R}$:

$$\cancel{f''(Ax+b) A A^T}$$

f is convex if

g_i convex, h convex, \tilde{h} nondecreasing in each argument
 g_i concave, h convex, \tilde{h} nonincreasing in each argument

proof (for $n = 1$, differentiable g, h)

$$f''(x) = g'(x)^T \underbrace{\nabla^2 h(g(x))}_{\leq 0} g'(x) + \nabla h(g(x))^T \cancel{g''(x)} \leq 0$$

① \rightarrow

≤ 0

≥ 0

≤ 0

≥ 0

examples ② \rightarrow

≥ 0

≥ 0

≥ 0

- ① • $\sum_{i=1}^m \log g_i(x)$ is **concave** if g_i are concave and positive

- ② • $\log \sum_{i=1}^m \exp g_i(x)$ is convex if g_i are convex

$$h(x)$$

$g \rightarrow$ concave

$h \rightarrow$ concave.

$h \rightarrow ?$

Minimization

if $f(x, y)$ is convex in (x, y) and C is a convex set, then

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

is convex

examples

- $f(x, y) = x^T Ax + 2x^T By + y^T Cy$ with

$$\begin{bmatrix} x \\ y \end{bmatrix}^T \begin{bmatrix} A & B \\ B^T & C \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

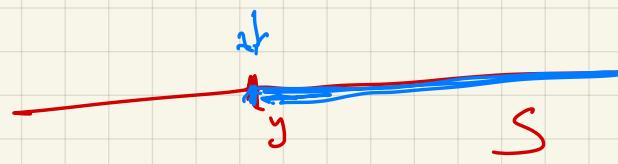
$$\begin{bmatrix} A & B \\ B^T & C \end{bmatrix} \succeq 0, \quad C \succ 0$$

minimizing over y gives $g(x) = \inf_y f(x, y) = x^T(A - BC^{-1}B^T)x$

g is convex, hence Schur complement $A - BC^{-1}B^T \succeq 0$

- distance to a set: $\text{dist}(x, S) = \inf_{y \in S} \|x - y\|$ is convex if S is convex

Proof



Let $\epsilon > 0$

Then there are $y_1, y_2 \in C$ such that $f(x_i, y_i) \leq g(x_i) + \epsilon$ for $i = 1, 2$. Now let $\theta \in [0, 1]$. We have

$$\begin{aligned} g(\theta x_1 + (1 - \theta)x_2) &= \inf_{y \in C} f(\theta x_1 + (1 - \theta)x_2, y) \\ &\leq f(\theta x_1 + (1 - \theta)x_2, \theta y_1 + (1 - \theta)y_2) \\ &\leq \theta f(x_1, y_1) + (1 - \theta)f(x_2, y_2) \\ &\leq \theta g(x_1) + (1 - \theta)g(x_2) + \epsilon. \end{aligned}$$

Since this holds for any $\epsilon > 0$, we have

$$g(\theta x_1 + (1 - \theta)x_2) \leq \theta g(x_1) + (1 - \theta)g(x_2).$$

$$g \leq g + \epsilon$$

$$a \leq b + \epsilon$$

$$a \leq b$$

$$f(\underline{x}, \underline{y}) = \underline{x}^T A \underline{x} + 2 \underline{x}^T B \underline{y} + \underline{y}^T C \underline{y}$$

$$\nabla_{\underline{y}} f = 0 + 2B^T \underline{x} + 2C \underline{y}^* = 0$$

$$\underline{y}^* = -C^{-1} B^T \underline{x}$$

$$f(\underline{x}, \underline{y}^*) = \underline{x}^T A \underline{x} - 2 \underline{x}^T B C^{-1} B^T \underline{x} + (C^{-1} B^T \underline{x})^T C (C^{-1} B^T \underline{x})$$

$$= \underline{x}^T A \underline{x} - 2 \underline{x}^T B C^{-1} B^T \underline{x} + \underline{x}^T B (C^{-1})^T C (C^{-1} B^T \underline{x})$$

$$(C^{-1})^T = C = \underline{x}^T A \underline{x} - 2 \underline{x}^T B C^{-1} B^T \underline{x} + \underline{x}^T B C^{-1} B^T \underline{x}$$

Since C is symmetric.

$$= \underline{x}^T A \underline{x} - \underline{x}^T B C^{-1} B^T \underline{x}$$

$$= \underline{x}^T (A - B C^{-1} B^T) \underline{x}$$

Schur complement

Consider a matrix $X \in \mathbf{S}^n$ partitioned as

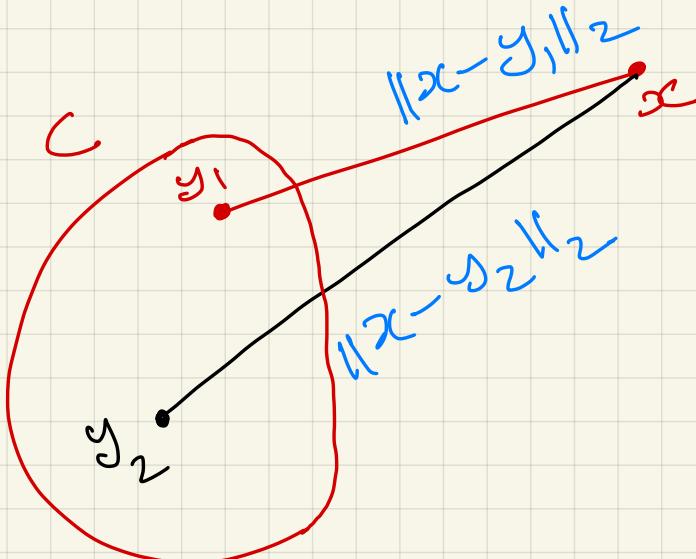
$$X = \begin{bmatrix} A & B \\ B^T & C \end{bmatrix},$$

where $A \in \mathbf{S}^k$. If $\det A \neq 0$, the matrix

$$S = C - B^T A^{-1} B$$

is called the *Schur complement* of A in X . Schur complements arise in several contexts, and appear in many important formulas and theorems. For example, we have

$$\det X = \det A \det S.$$



$$\inf_{y \in C} \|x-y\|_2$$

Perspective

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \end{bmatrix} \in \mathbb{R}^n$$

the **perspective** of a function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is the function $g : \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}$,

$$g(x, t) = tf(x/t), \quad \text{dom } g = \{(x, t) \mid x/t \in \text{dom } f, t > 0\}$$

g is convex if f is convex

examples

- $f(x) = x^T x$ is convex; hence $g(x, t) = x^T x / t$ is convex for $t > 0$
- negative logarithm $f(x) = -\log x$ is convex; hence relative entropy $g(x, t) = t \log t - t \log x$ is convex on \mathbb{R}_{++}^2
- if f is convex, then

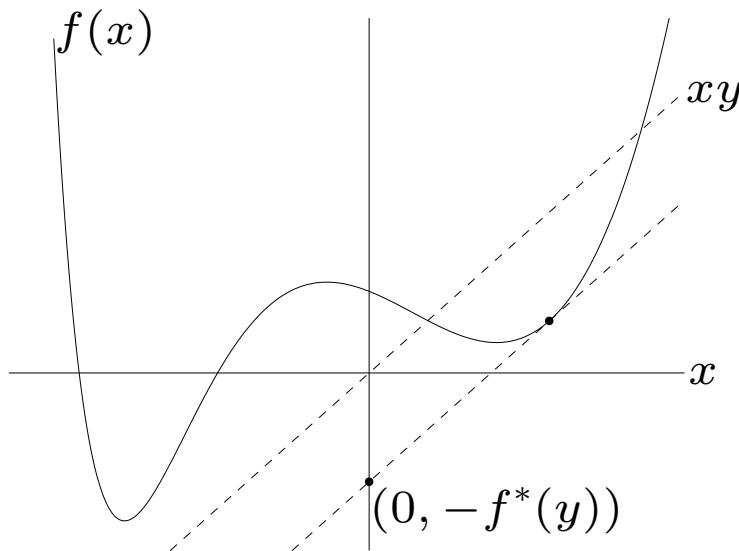
$$g(x) = (c^T x + d) f \left((Ax + b) / (c^T x + d) \right)$$

is convex on $\{x \mid c^T x + d > 0, (Ax + b) / (c^T x + d) \in \text{dom } f\}$

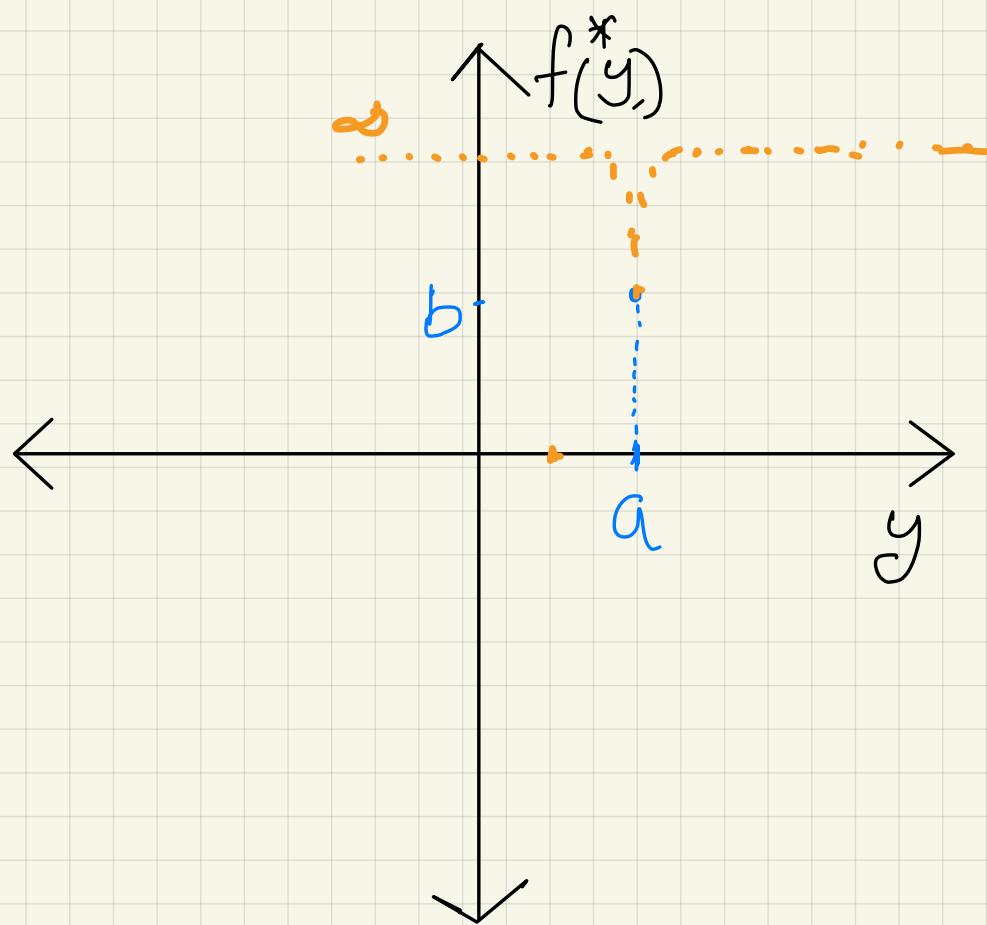
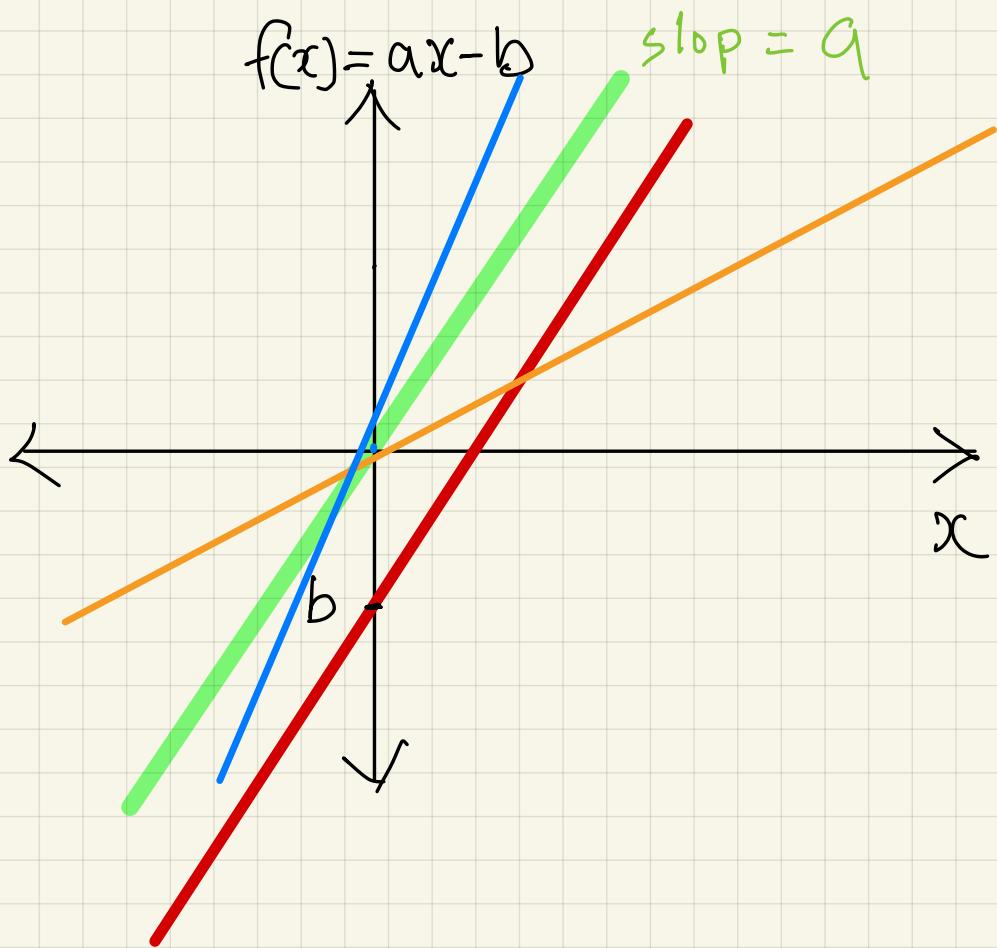
The conjugate function

the **conjugate** of a function f is

$$f^*(y) = \sup_{x \in \text{dom } f} (y^T x - f(x))$$



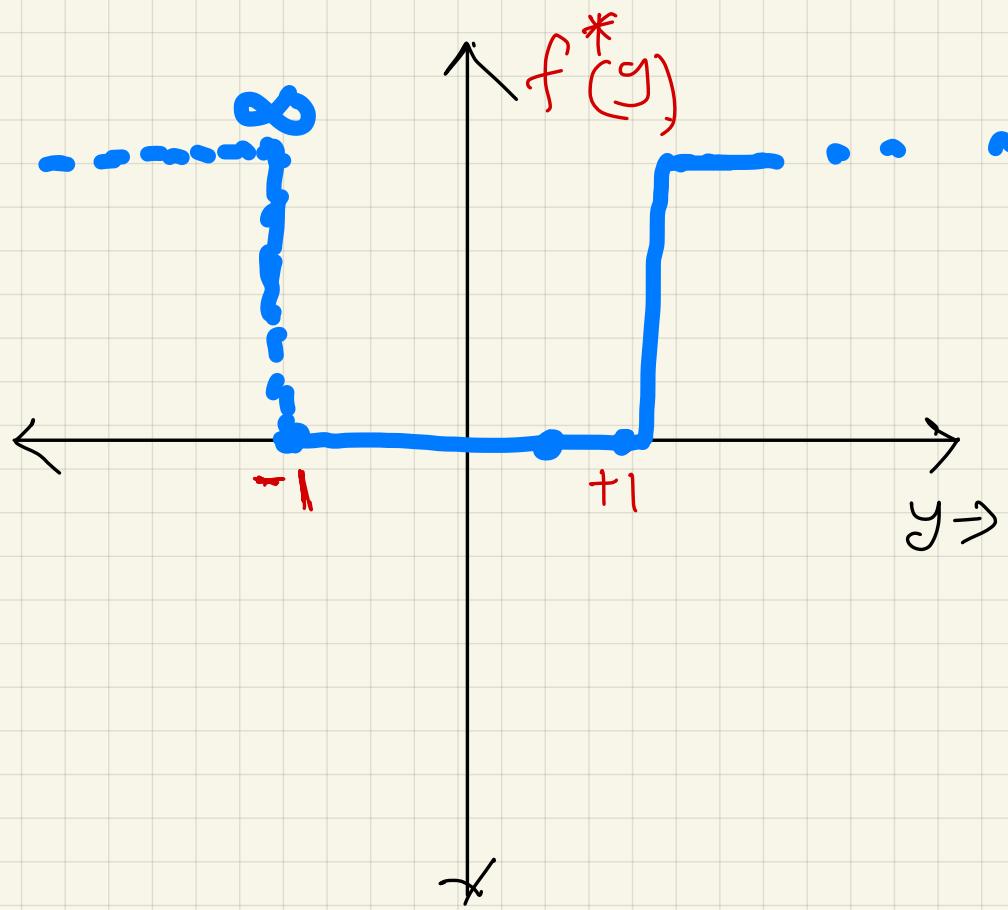
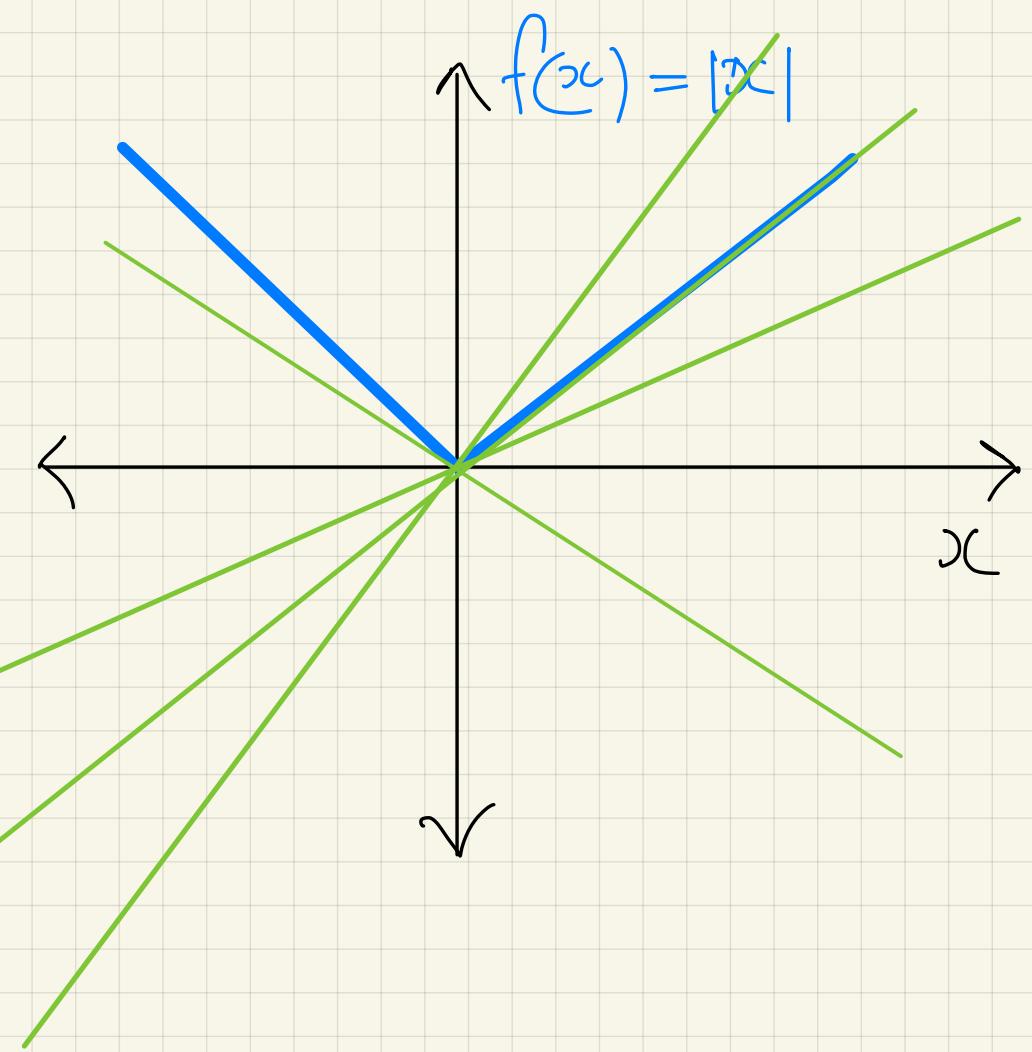
- f^* is convex (even if f is not)
- will be useful in chapter 5



$$f^*(y) = \sup_{x \in \mathbb{R}} (yx - (ax - b))$$

$$= \sup_{x \in \mathbb{R}} ((y-a)x + b) =$$

$$\begin{cases} b & \text{if } y = a \\ \infty & \text{if } y \neq a \end{cases}$$



examples

- negative logarithm $f(x) = -\log x$

$$f^*(y) = \sup_{x>0} (xy + \underbrace{\log x}_{g(x)})$$

$$= \begin{cases} -1 - \log(-y) & y < 0 \\ \infty & \text{otherwise} \end{cases}$$

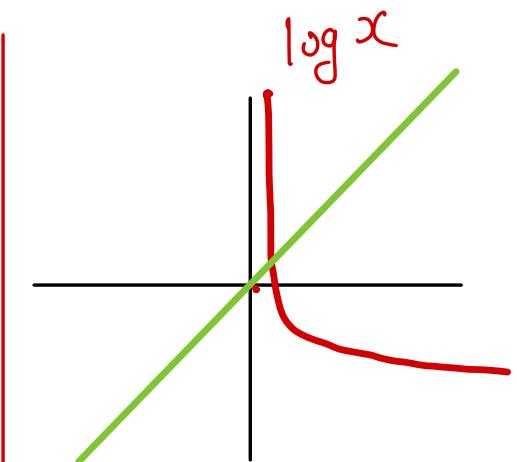
$$\begin{aligned} g(x^*) &= y + \frac{1}{x^*} = 0 \\ x^* &= (-1/y) \\ g(x^*) &= -1 + \log\left(\frac{-1}{y}\right) \\ &= -1 - \log(-y) \end{aligned}$$

- strictly convex quadratic $f(x) = (1/2)x^T Qx$ with $Q \in \mathbf{S}_{++}^n$

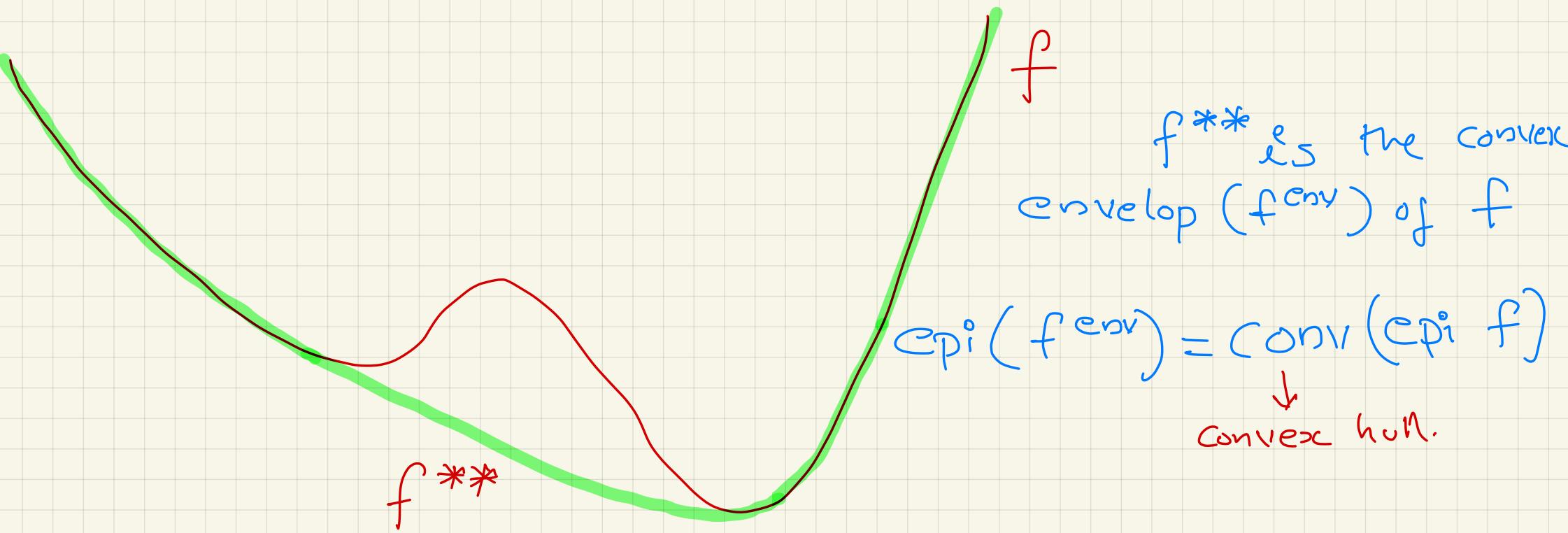
$$\begin{aligned} f^*(y) &= \sup_x (y^T x - (1/2)x^T Qx) \\ &= \frac{1}{2} y^T Q^{-1} y \end{aligned}$$

$$\begin{aligned} g(x^*) &= y - Qx^* = 0 \\ \Rightarrow x^* &= Q^{-1}y \end{aligned}$$

$$\begin{aligned} g(x^*) &= y^T Q^{-1} y - \frac{1}{2} (Q^{-1}y)^T Q (Q^{-1}y) \\ &= y^T Q^{-1} y - \frac{1}{2} y^T Q^{-1} Q Q^{-1} y \\ &= \frac{1}{2} y^T Q^{-1} y \end{aligned}$$



Conjugate of a Conjugate: (f^{**})



if f is convex

$$f^{***} = f$$

(Note: f should be closed.)

Quasiconvex functions

$f : \mathbf{R}^n \rightarrow \mathbf{R}$ is quasiconvex if $\text{dom } f$ is convex and the sublevel sets

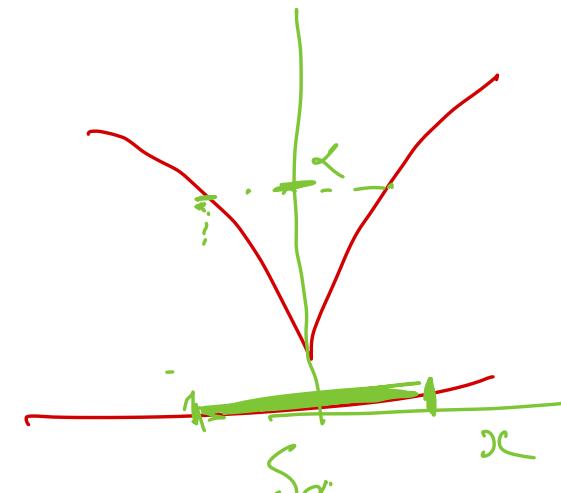
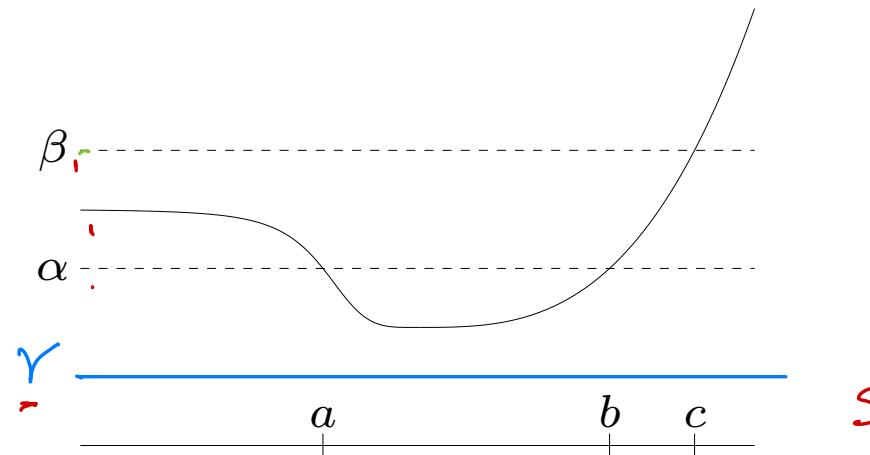
$$S_\alpha = \{x \in \text{dom } f \mid f(x) \leq \alpha\}$$

are convex for all α

$$S_\beta = [\omega, c] \quad \checkmark$$

$$S_\alpha = [a, b] \quad \checkmark$$

$$S_r = \emptyset \quad -$$



- f is quasiconcave if $-f$ is quasiconvex or α -super level set is convex
- f is quasilinear if it is quasiconvex and quasiconcave

$$S_\alpha^{\text{sup}} = \left\{ x \in \text{dom } f \mid f(x) \geq \alpha \right\}$$

Examples

- $\sqrt{|x|}$ is quasiconvex on \mathbf{R}
- $\text{ceil}(x) = \inf\{z \in \mathbf{Z} \mid z \geq x\}$ is quasilinear
- $\log x$ is quasilinear on \mathbf{R}_{++}
- $f(x_1, x_2) = x_1 x_2$ is quasiconcave on \mathbf{R}_{++}^2
- linear-fractional function

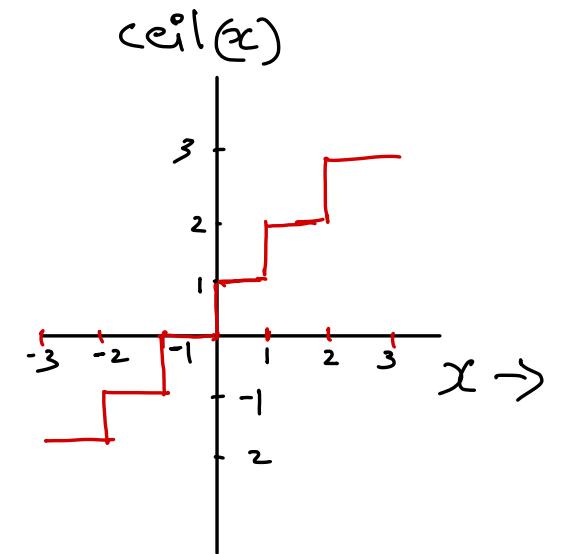
$$f(x) = \frac{a^T x + b}{c^T x + d}, \quad \text{dom } f = \{x \mid c^T x + d > 0\}$$

is quasilinear

- distance ratio

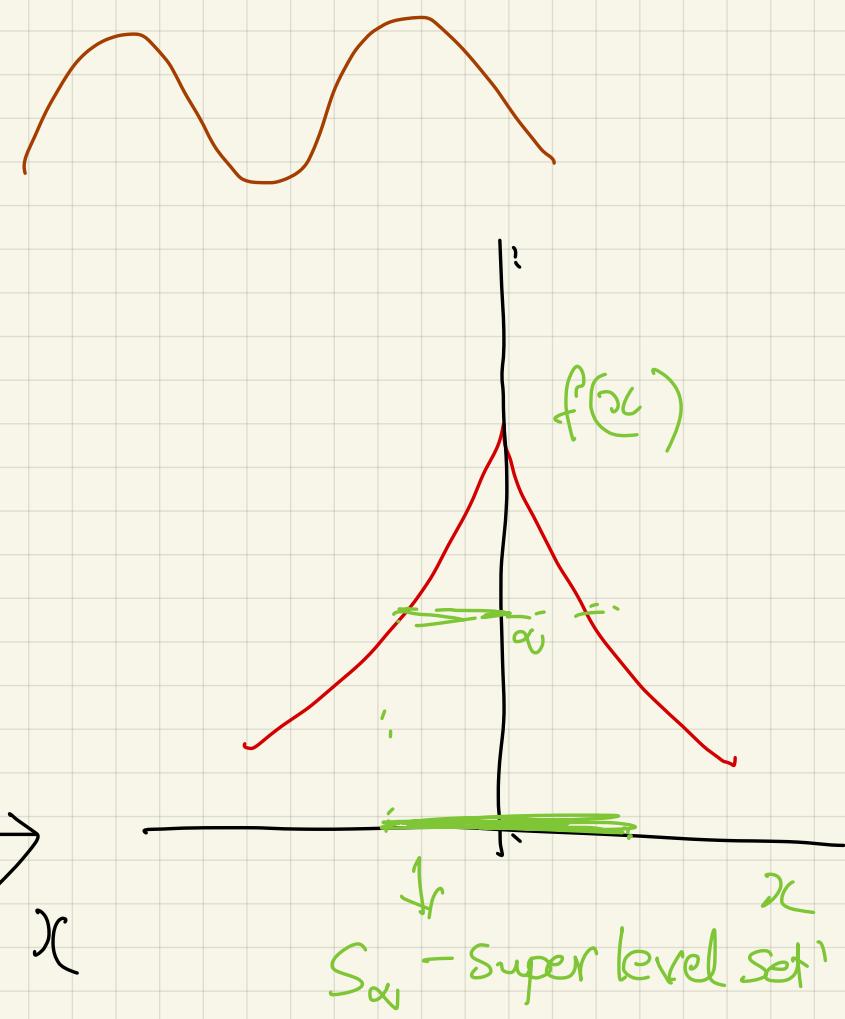
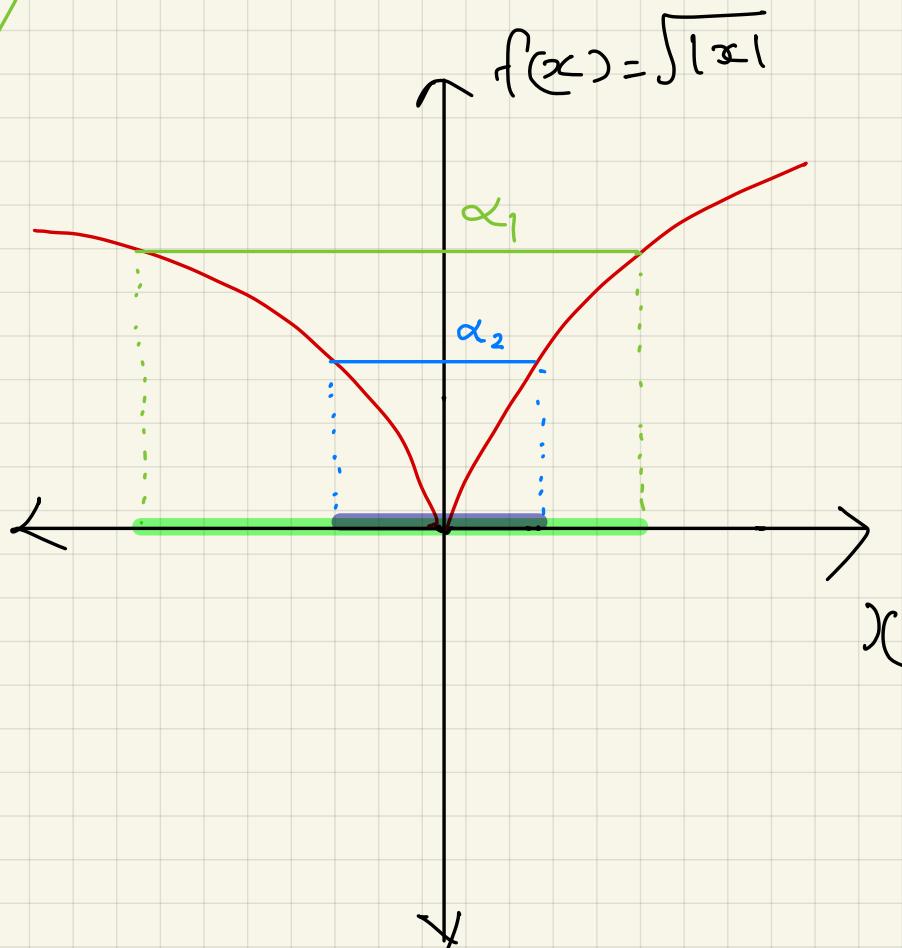
$$f(x) = \frac{\|x - a\|_2}{\|x - b\|_2}, \quad \text{dom } f = \{x \mid \|x - a\|_2 \leq \|x - b\|_2\}$$

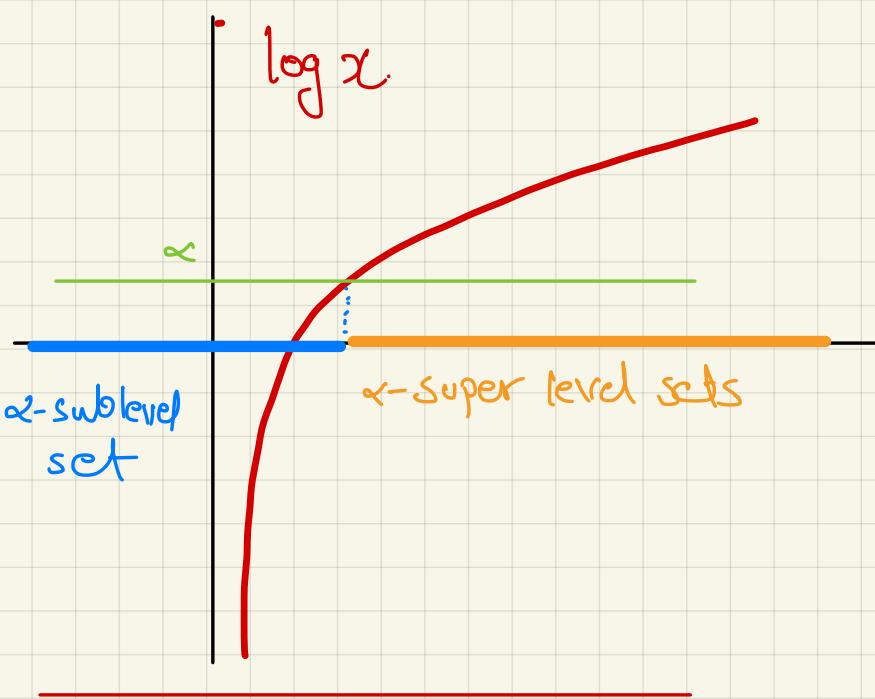
is quasiconvex



Quasi Convex

$C(x)$





$$f(x_1, x_2) = x_1 x_2, \quad x \in \mathbb{R}_{++}^2$$

α -super level set $\rightarrow S_\alpha^{\text{sup}}$

$$\begin{aligned} \text{Let } \underline{x} &\in S_\alpha^{\text{sup}} \Rightarrow x_1 x_2 \geq \alpha \\ \underline{y} &\in S_\alpha^{\text{sup}} \Rightarrow y_1 y_2 \geq \alpha \end{aligned}$$

$$\theta \underline{x} + (1-\theta) \underline{y} \in S_\alpha^{\text{sup}} ?$$

$$\theta \in [0, 1]$$

$$\begin{aligned}
& (\theta x_1 + (1-\theta)y_1)(\theta x_2 + (1-\theta)y_2) \\
&= \theta^2 x_1 x_2 + (1-\theta)^2 y_1 y_2 + \theta(1-\theta)(x_1 y_2 + x_2 y_1) \\
&= \theta^2 x_1 x_2 + (1-\theta)^2 y_1 y_2 + \theta(1-\theta) \left(\frac{x_1}{x_2} x_2 y_2 + \frac{x_2}{x_1} x_1 y_1 \right) \\
&\geq \theta^2 + \theta(1-\theta)^2 + \theta(1-\theta) \alpha \left(\frac{x_1}{x_2} + \frac{x_2}{x_1} \right) \\
&= \theta^2 + \theta(1-\theta)^2 + \theta(1-\theta) \alpha \left(\frac{x_1^2 + x_2^2}{x_1 x_2} \right) \\
&\geq \theta^2 + \theta(1-\theta)^2 + \theta(1-\theta) \alpha (2) \\
&= \alpha (\theta + 1 - \theta)^2 = \alpha.
\end{aligned}$$

$$\nabla f(x_1, x_2) = \begin{bmatrix} x_2 & x_1 \end{bmatrix}$$

$$\nabla^2 f(x_1, x_2) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

$$\text{eigen values} = \{-1, +1\}$$

$$f(x) = \frac{a^T x + b}{c^T x + d},$$

with $\text{dom } f = \{x \mid c^T x + d > 0\}$, is quasiconvex, and quasiconcave, i.e., quasilinear.
Its α -sublevel set is

$$\begin{aligned} S_\alpha &= \{x \mid c^T x + d > 0, (a^T x + b)/(c^T x + d) \leq \alpha\} \\ &= \{x \mid c^T x + d > 0, a^T x + b \leq \alpha(c^T x + d)\}, \end{aligned}$$

which is convex, since it is the intersection of an open halfspace and a closed halfspace.
(The same method can be used to show its superlevel sets are convex.)

$$f(x) = \frac{\|x - a\|_2}{\|x - b\|_2},$$

i.e., the ratio of the Euclidean distance to a to the distance to b . Then f is quasiconvex on the halfspace $\{x \mid \|x - a\|_2 \leq \|x - b\|_2\}$. To see this, we consider the α -sublevel set of f , with $\alpha \leq 1$ since $f(x) \leq 1$ on the halfspace $\{x \mid \|x - a\|_2 \leq \|x - b\|_2\}$. This sublevel set is the set of points satisfying

$$\|x - a\|_2 \leq \alpha \|x - b\|_2.$$

Squaring both sides, and rearranging terms, we see that this is equivalent to

$$(1 - \alpha^2)x^T x - 2(a - \alpha^2 b)^T x + a^T a - \alpha^2 b^T b \leq 0.$$

$B(\underline{x}_c, r)$

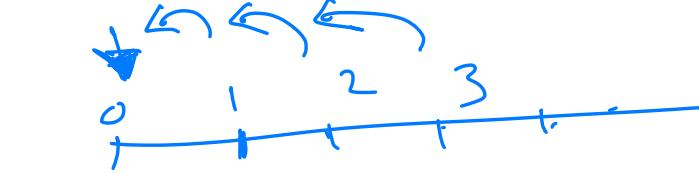
This describes a convex set (in fact a Euclidean ball) if $\alpha \leq 1$.

internal rate of return

$$r = 8\% \text{ per year} \quad x \rightarrow x(1+r)$$

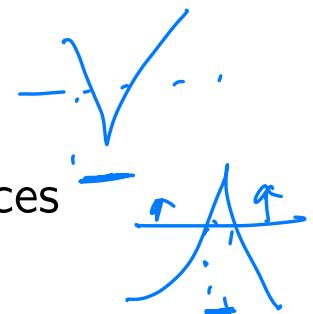
- cash flow $x = (x_0, \dots, x_n)$; x_i is payment in period i (to us if $x_i > 0$)
- we assume $x_0 < 0$ and $x_0 + x_1 + \dots + x_n > 0$
- present value of cash flow x , for interest rate r :

$$PV(x, r) = \sum_{i=0}^n (1+r)^{-i} x_i = -10000 + \frac{2000}{1+r} + \frac{3000}{(1+r)^2} + \dots$$



- internal rate of return is smallest interest rate for which $PV(x, r) = 0$:

$$\text{IRR}(x) = \inf\{r \geq 0 \mid PV(x, r) = 0\}$$



IRR is quasiconcave: superlevel set is intersection of open halfspaces

$$\text{IRR}(x) \geq R \iff \sum_{i=0}^n (1+r)^{-i} x_i > 0 \text{ for } 0 \leq r < R$$

possible cash flows for which smallest interest rate for zero present value (PV) is $\geq R$.

Convex functions

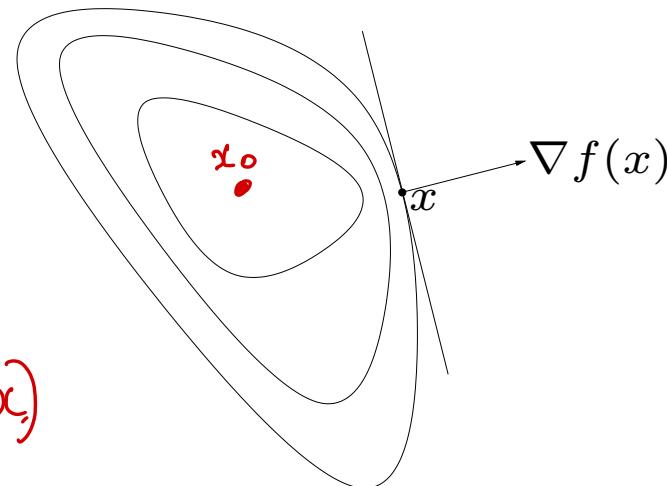
Properties

modified Jensen inequality: for quasiconvex f

$$0 \leq \theta \leq 1 \implies f(\theta x + (1 - \theta)y) \leq \max\{f(x), f(y)\}$$

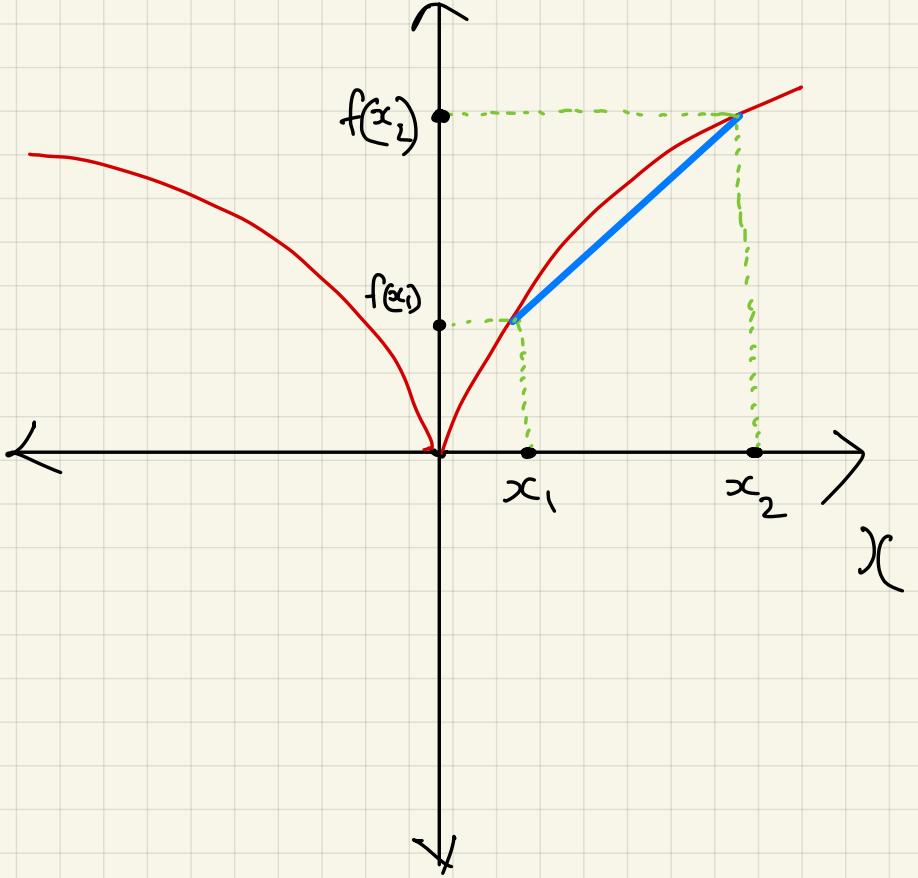
first-order condition: differentiable f with cvx domain is quasiconvex iff

$$f(y) \leq f(x) \implies \nabla f(x)^T(y - x) \leq 0$$



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

sums of quasiconvex functions are not necessarily quasiconvex



$\max \{f(x_1), f(x_2)\}$
= $f(x_2)$

Log-concave and log-convex functions

a positive function f is log-concave if $\log f$ is concave:

$$f(\theta x + (1 - \theta)y) \geq f(x)^\theta f(y)^{1-\theta} \quad \text{for } 0 \leq \theta \leq 1$$

f is log-convex if $\log f$ is convex

$$f(x) \geq 0 \quad \forall x \in \text{dom } f.$$

$$f(x) = x^a$$

$$\log f(x) =$$

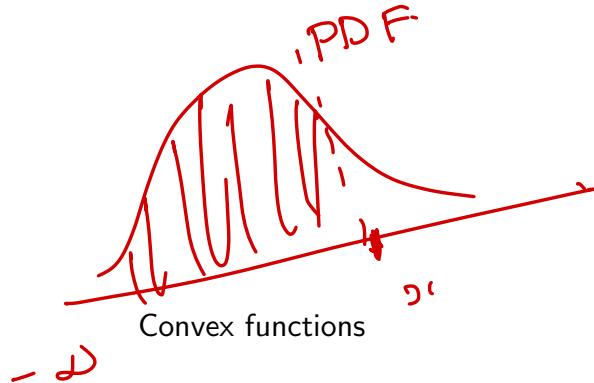
$$a \log x$$

- powers: x^a on \mathbb{R}_{++} is log-convex for $a \leq 0$, log-concave for $a \geq 0$
- many common probability densities are log-concave, e.g., normal:

KC

$$f(x) = \frac{1}{\sqrt{(2\pi)^n \det \Sigma}} e^{-\frac{1}{2}(x-\bar{x})^T \Sigma^{-1}(x-\bar{x})} \rightarrow -\frac{1}{2}(x-\bar{x})^T \Sigma^{-1}(x-\bar{x})$$

- cumulative Gaussian distribution function Φ is log-concave



$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-u^2/2} du$$

if PDF is
log concave
then CDF is
log concave.

log-concave function.

$$\log f(\theta \underline{x} + (1-\theta) \underline{y}) \geq \theta \log f(\underline{x}) + (1-\theta) \log f(\underline{y})$$

$$\Rightarrow \log f(\theta \underline{x} + (1-\theta) \underline{y}) \geq \log [f(\underline{x})]^\theta + \log [f(\underline{y})]^{(1-\theta)}$$

$$\Rightarrow \log f(\theta \underline{x} + (1-\theta) \underline{y}) \geq \log [f(\underline{x})^\theta \cdot f(\underline{y})^{(1-\theta)}]$$

$$f(\theta \underline{x} + (1-\theta) \underline{y}) \geq f(\underline{x})^\theta f(\underline{y})^{(1-\theta)}$$

Properties of log-concave functions

- twice differentiable f with convex domain is log-concave if and only if

$$f(x)\nabla^2 f(x) \preceq \nabla f(x)\nabla f(x)^T$$

for all $x \in \text{dom } f$

- product of log-concave functions is log-concave
- sum of log-concave functions is not always log-concave
- integration: if $f : \mathbf{R}^n \times \mathbf{R}^m \rightarrow \mathbf{R}$ is log-concave, then

$$g(x) = \int f(x, y) dy$$

is log-concave (not easy to show)

Further reading: properties of log-concave functions

$f_1, f_2 \rightarrow \text{log-concave}$
 $\log f_1 \quad \log f_2 \rightarrow \text{convex}$
 $g = f_1 \cdot f_2$
 $\log g = \underbrace{\log f_1 + \log f_2}_{\text{Sum of 2 concave functions is concave}}$
 $\Rightarrow g \text{ is log concave}$

$$\underline{g(x)} = \log f(x)$$

$$\nabla \underline{g(x)} = \frac{1}{f(x)} \nabla f(x)$$

$$\nabla^2 \underline{g(x)} = \frac{1}{f(x)} \nabla^2 f(x) - \frac{1}{(f(x))^2} (\nabla f(x)) (\nabla f(x))^T$$

$\underline{g(x)}$ is log concave $\Leftrightarrow \nabla^2 \underline{g(x)} \leq 0$

$$f(x) \nabla^2 f(x) \leq (\nabla f(x)) (\nabla f(x))^T$$

consequences of integration property

- convolution $f * g$ of log-concave functions f, g is log-concave

$$(f * g)(x) = \int f(x - y)g(y)dy$$

- if $C \subseteq \mathbf{R}^n$ convex and y is a random variable with log-concave pdf then

$$f(x) = \text{prob}(x + y \in C)$$

is log-concave

proof: write $f(x)$ as integral of product of log-concave functions

$$f(x) = \int g(x + y)p(y) dy, \quad g(u) = \begin{cases} 1 & u \in C \\ 0 & u \notin C, \end{cases}$$

p is pdf of y

example: yield function

$$Y(x) = \mathbf{prob}(x + w \in S)$$

- $x \in \mathbf{R}^n$: nominal parameter values for product
- $w \in \mathbf{R}^n$: random variations of parameters in manufactured product
- S : set of acceptable values

if S is convex and w has a log-concave pdf, then

- Y is log-concave
- yield regions $\{x \mid Y(x) \geq \alpha\}$ are convex

Convexity with respect to generalized inequalities

$f : \mathbf{R}^n \rightarrow \mathbf{R}^m$ is K -convex if $\text{dom } f$ is convex and

$$f : \mathbf{R}^n \rightarrow \mathbf{R}$$

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

↗ proper cone.

for $x, y \in \text{dom } f$, $0 \leq \theta \leq 1$

example $f : \mathbf{S}^m \rightarrow \mathbf{S}^m$, $f(X) = X^2$ is \mathbf{S}_+^m -convex

proof: for fixed $z \in \mathbf{R}^m$, $z^T X^2 z = \|Xz\|_2^2$ is convex in X , i.e.,

$$z^T (\theta X + (1 - \theta)Y)^2 z \leq \theta z^T X^2 z + (1 - \theta) z^T Y^2 z$$

for $X, Y \in \mathbf{S}^m$, $0 \leq \theta \leq 1$

$$\text{therefore } (\theta X + (1 - \theta)Y)^2 \preceq \theta X^2 + (1 - \theta)Y^2$$

$$\left. \begin{array}{l} f: S^n \rightarrow S^n \\ f(x) = x^2 \end{array} \right\} \rightarrow S_+^n - \text{convex}.$$

PROOF

$$h(x) = z^T x^2 z = \|xz\|_2^2 \text{ is convex in } x \in S^n \text{ for fixed } z \in \mathbb{R}^n$$

$\stackrel{z^T x^2 z = (x^2)^T z}{=}$

$$x, y \in S^n$$

$$\Rightarrow h(\theta x + (1-\theta)y) \leq \theta h(x) + (1-\theta) h(y) \xrightarrow{\theta \in [0,1]} \text{definition of convexity}$$

$$z^T (\theta x + (1-\theta)y)^2 z \leq \theta z^T x^2 z + (1-\theta) z^T y^2 z$$

$$\Rightarrow \theta z^T x^2 z + (1-\theta) z^T y^2 z - z^T (\theta x + (1-\theta)y)^2 z \geq 0$$

$$\Rightarrow z^T [\theta x^2 + (1-\theta)y^2 - (\theta x + (1-\theta)y)^2] \geq 0$$

PSD

$$\Rightarrow \theta x^2 + (1-\theta)y^2 - (\theta x + (1-\theta)y)^2 \stackrel{S_+^n}{\succeq} 0$$

$$\Rightarrow (\theta x + (1-\theta)y)^2 \stackrel{S_+^n}{\leq} \theta x^2 + (1-\theta)y^2$$

$$\Rightarrow f(x) = x^2 \text{ is } S_+^n - \text{convex}$$