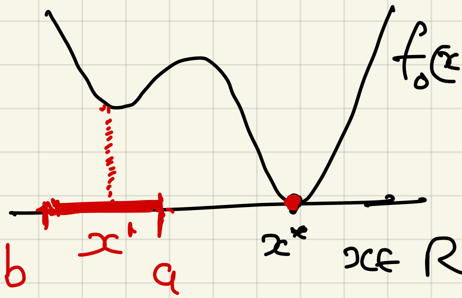
RECAP

Optimization problem

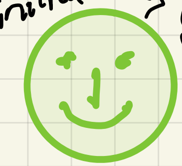
minimize $f_0(x)$

s.t. $f_i(x) \leq b_i$, $i = 1, 2, \ldots m$



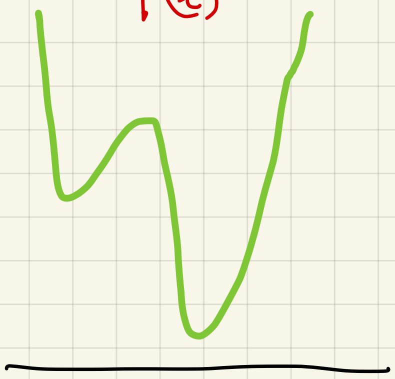
Convex

local minima => global minima

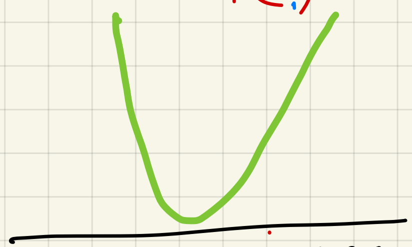


Non convex

objective function



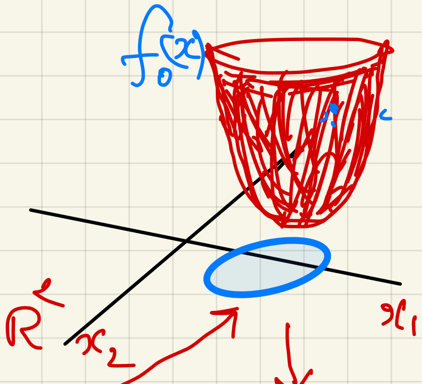
Non confex



Convex

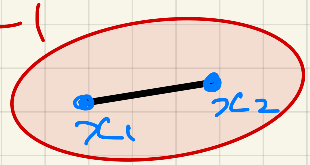
Constraints

forms a set



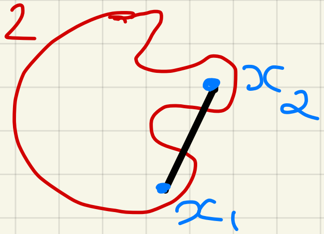
convex?

C_1



Convex

C_2

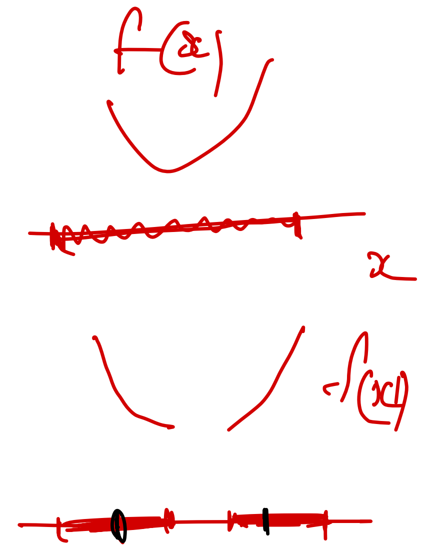


Non-convex.

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

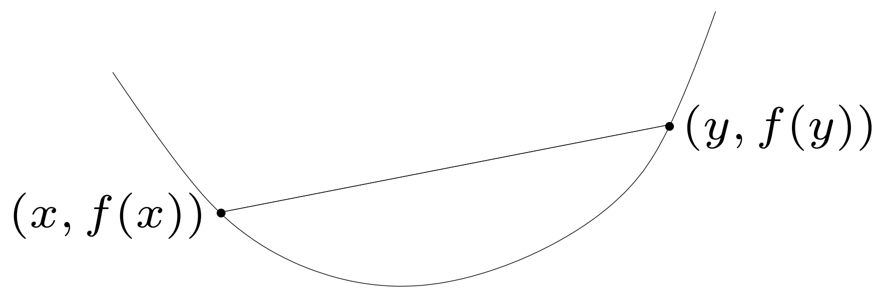
Definition



$f : \mathbf{R}^n \to \mathbf{R}$ is convex if $\textbf{dom}\, f$ is a convex set and

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

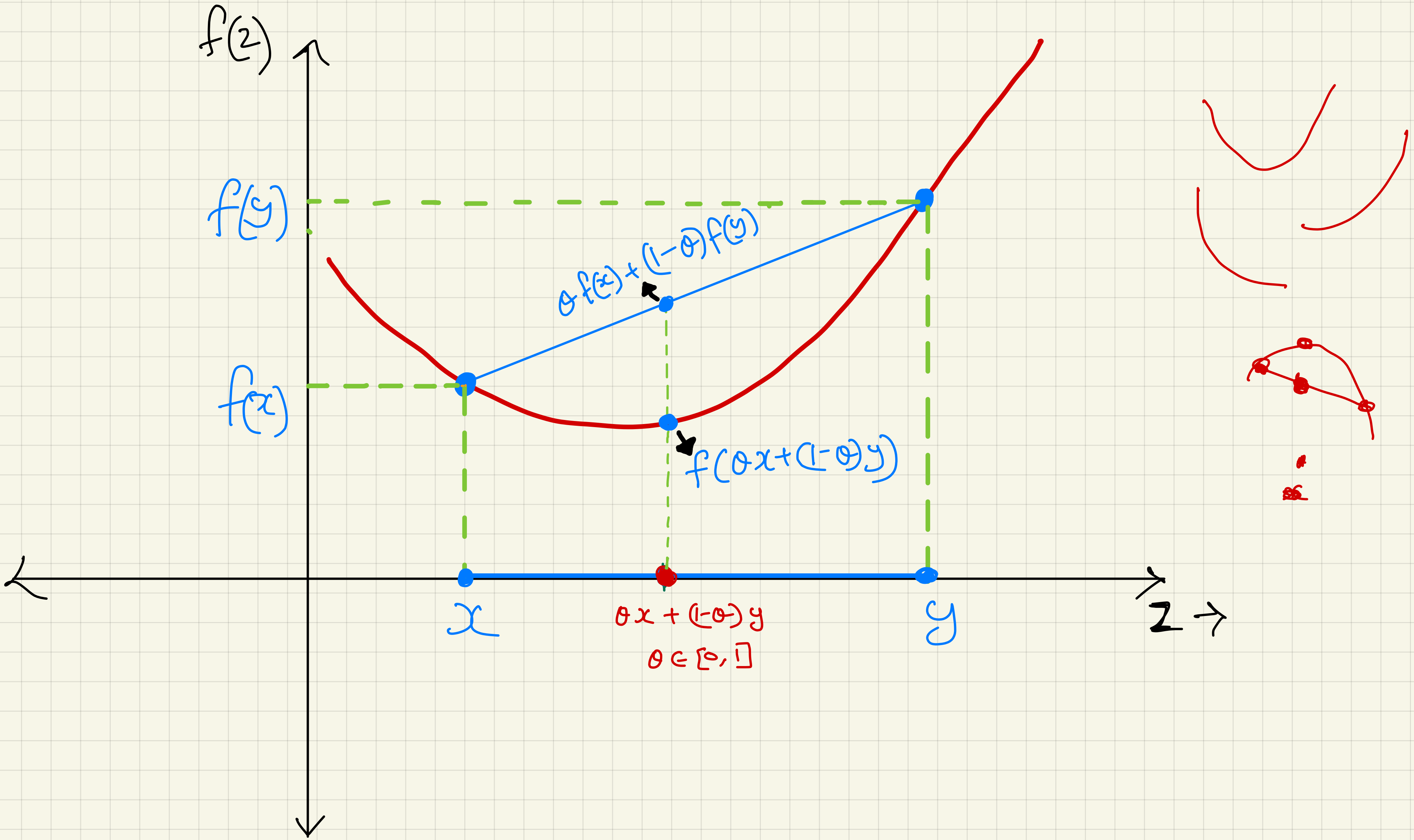
for all $x, y \in \textbf{dom}\, f$, $0 \le \theta \le 1$



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

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

for $x, y \in \textbf{dom}\, f$, $x \ne y$, $0 < \theta < 1$



f(z)
f(y)
f(x)
θf(x)+(1−θ)f(y)
f(θx+(1−θ)y)
x
θx + (1−θ) y
θ ∈ [0, 1]
y
z

Examples on R

convex:

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

concave:

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

$\boxed{\text{affine} \quad f(x) = ax + b}$

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

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

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

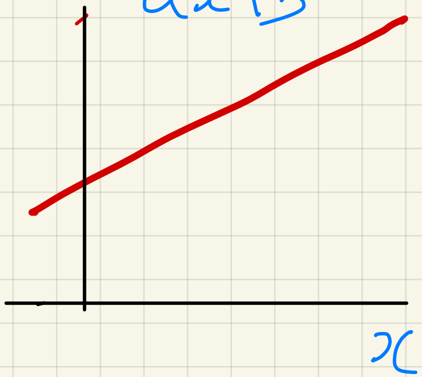
$= \theta(ax+b) + (1-\theta)(ay+b)$

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

Convex function: Examples

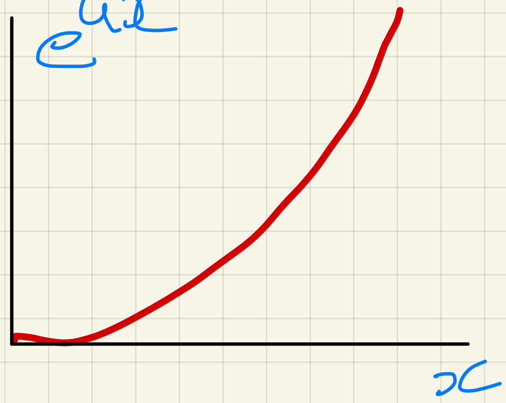
affine

$ax+b$



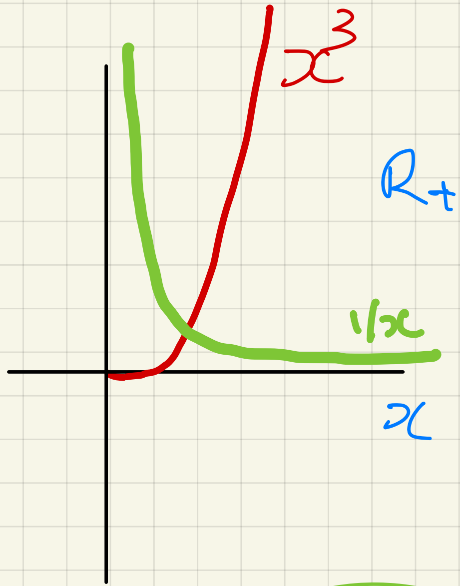
exponential

e^{ax}



powers

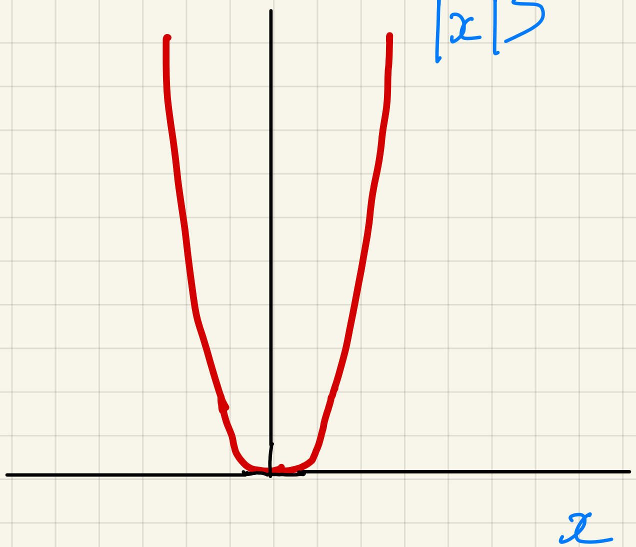
$\alpha \geq 1,\ \alpha \leq 0$



R_{++}

powers of absolute value

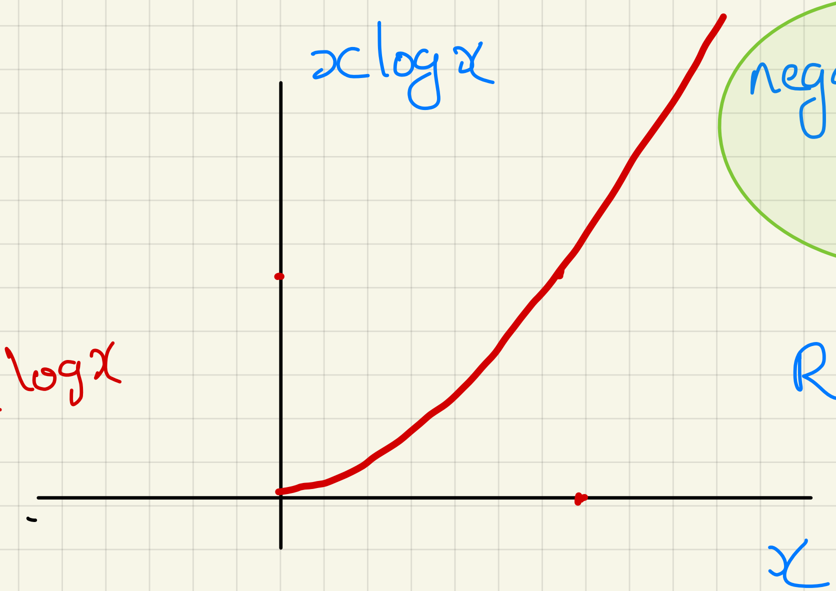
$|x|^3$



negative entropy

$x \log x$

$-x \log x$

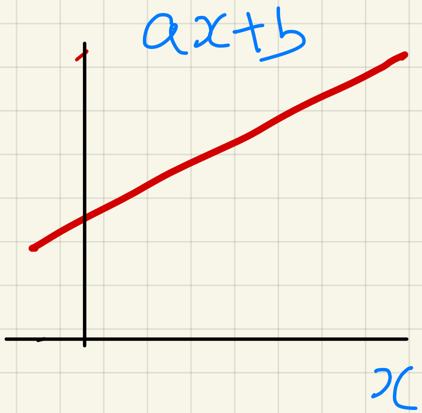


R_{++}

entropy is concave.

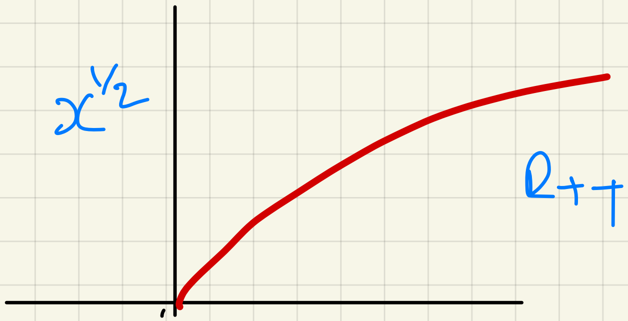
Concave function : Examples

affine



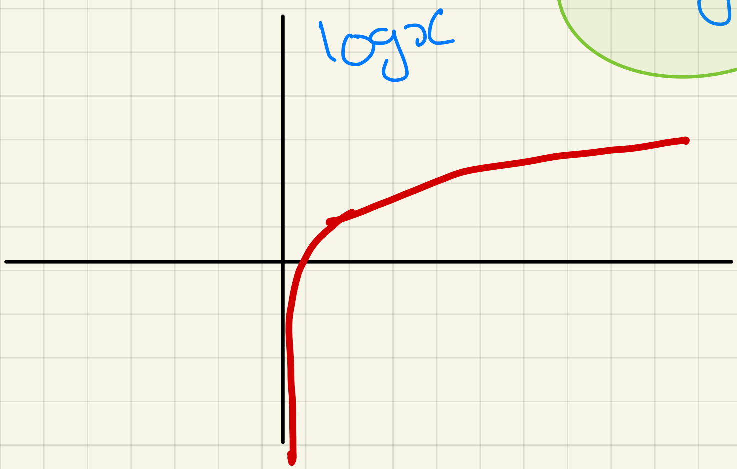
$0 \le \alpha \le 1$

powers



$\mathbb{R}_{++}$

$\log x$



$\mathbb{R}_{++}$

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

Affine function

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

$\mathbb{R}^n$ a^Tx+b

$\mathbb{R}^{m\times n}$ $tr(A^TX)+b$

not convex

affine functions are convex and concave; all norms are convex

examples on $\mathbf{R}^n$

$\|x\|_0$

$\|X\|_{1/2}$

- 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 $\mathbf{R}^{m\times n}$ ($m \times n$ matrices)

$\langle A, B\rangle = tr(A^TX)$

$\Rightarrow A(:)^T B(:)$

- affine function

$A = \begin{pmatrix} a_{11} & \cdots \\ \vdots & \end{pmatrix}$

$X = \begin{pmatrix} x_{11} & \cdots \\ \vdots & \end{pmatrix}$

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

- spectral (maximum singular value) norm

A

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

$\sqrt{eig(A^TA)}$

Revise: eigen value decomposition.
singular value decomposition.

Restriction of a convex function to a line

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

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

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

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

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

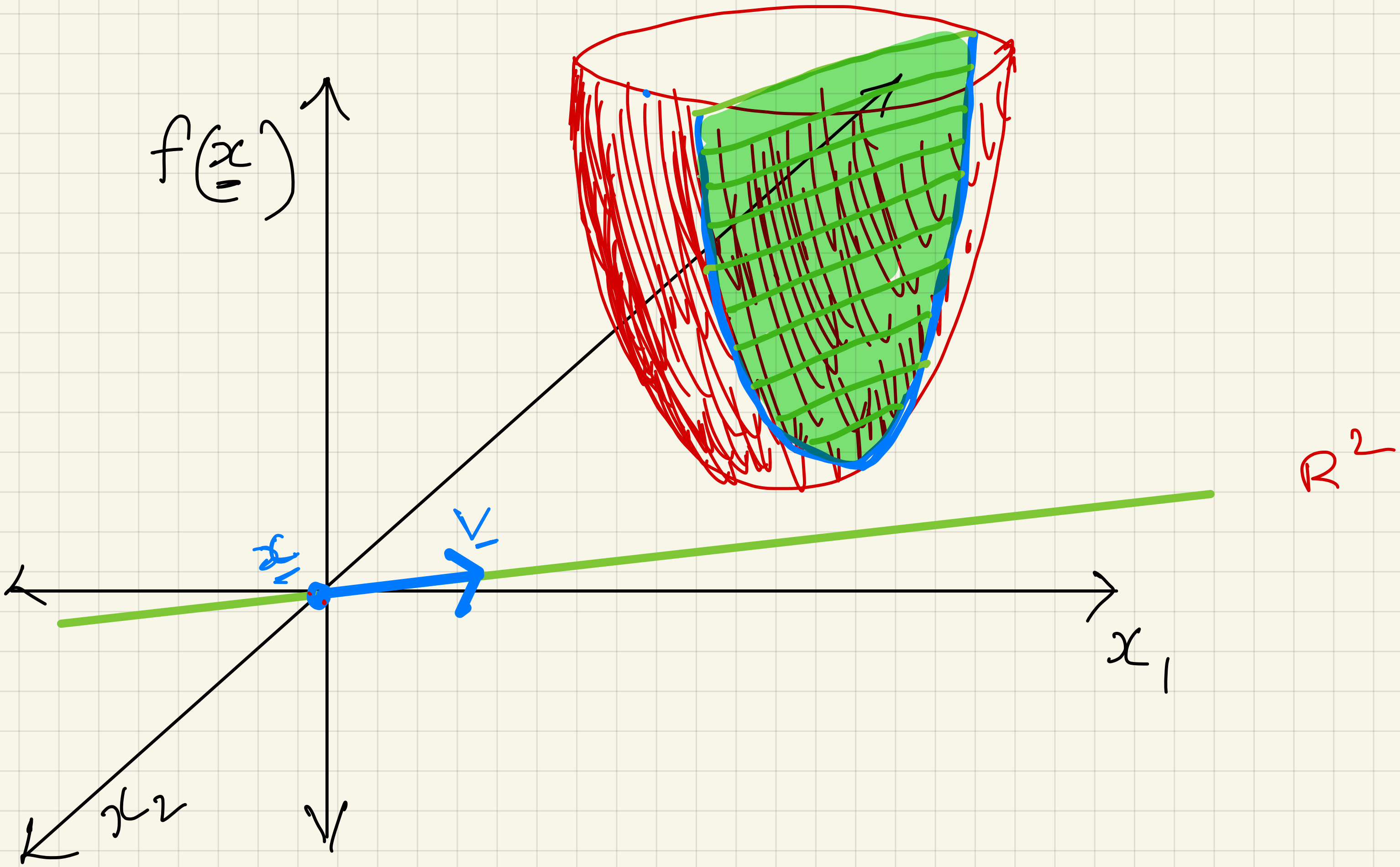
$$\begin{aligned} g(t) = \log\det(X + tV) &= \log\det X + \log\det(I + tX^{-1/2}VX^{-1/2}) \\ &= \log\det X + \sum_{i=1}^{n} \log(1 + t\lambda_i) \end{aligned}$$

$X+tV \in S^n_{++}$

$t=0$ $X \in S^n_{++}$

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



$f(\underline{x})$
$\underline{x}$
$\underline{v}$
x_1
x_2
R^2

$$g(t) = \log\det(X + tV)$$

$$= \log\det\left(X^{1/2}X^{1/2} + t\, X^{1/2}X^{-1/2}\, V\, X^{-1/2}X^{1/2}\right)$$

$$= \log\det\left[X^{1/2}\left(I + t\, X^{-1/2} V X^{-1/2}\right) X^{1/2}\right]$$

det(ABC) = det(A) det(B) det(C)

$$= \log\left[\left(\det X^{1/2}\right)\left[\det\left(I + t\, X^{-1/2} V X^{-1/2}\right)\right]\left(\det X^{1/2}\right)\right]$$

$\det(A^n) = (\det A)^n$

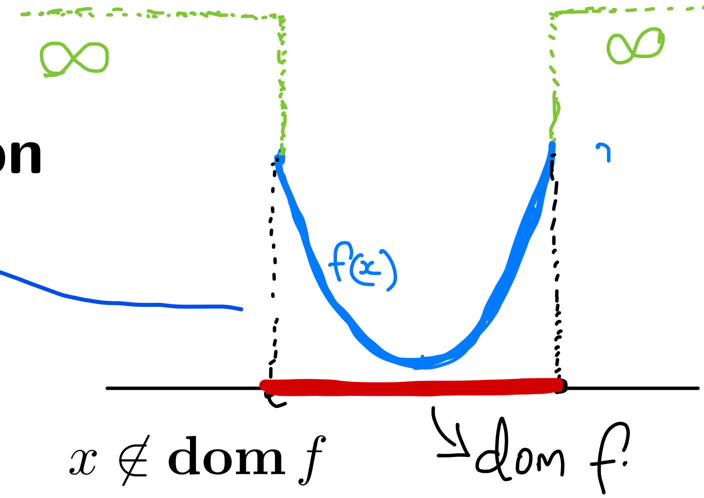
$$= \log\left[(\det X)^{1/2}\left[\det\left(I + t\, X^{-1/2} V X^{-1/2}\right)\right](\det X)^{1/2}\right]$$

$$= \log(\det X) + \log\left[\det\left(I + t\, \underbrace{X^{-1/2} V X^{-1/2}}_{\text{eigven values} = \lambda_1, \lambda_2 \cdots \lambda_n}\right)\right]$$

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

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

Extended-value extension



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

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

often simplifies notation; for example, the condition

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

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

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

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

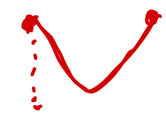
Revise

gradien gives vector.
Hessian gives matrix

First-order condition

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

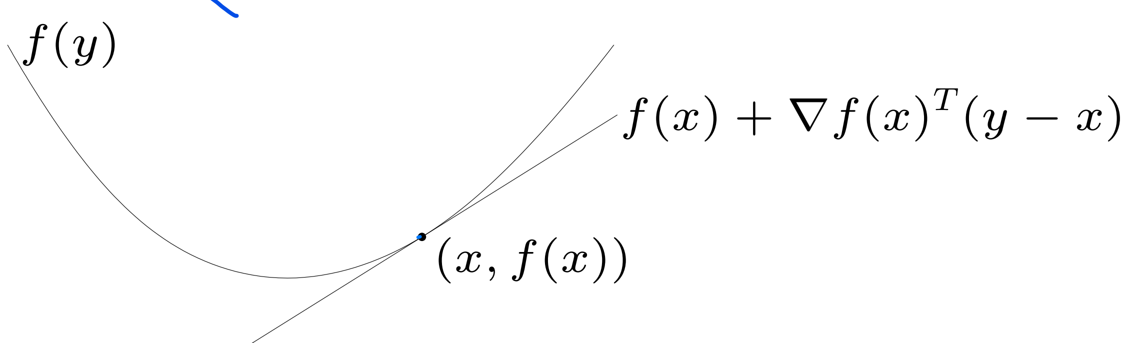
$$\nabla f(x) = \left(\frac{\partial f(x)}{\partial x_1}, \frac{\partial f(x)}{\partial x_2}, \ldots, \frac{\partial f(x)}{\partial x_n} \right)$$



exists at each $x \in \mathbf{dom}\, f$

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

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

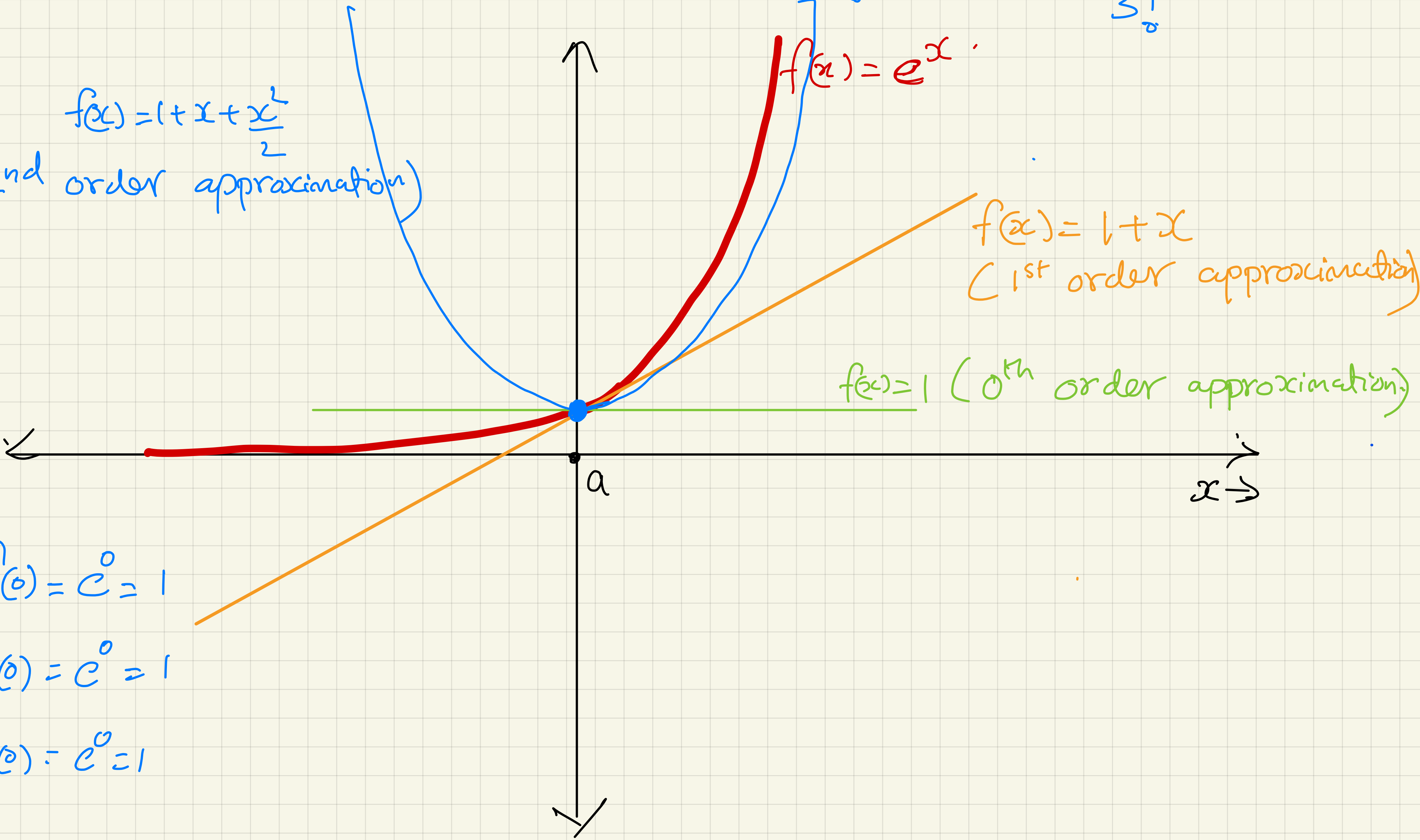


first-order approximation of f is global underestimator

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

Tailor series

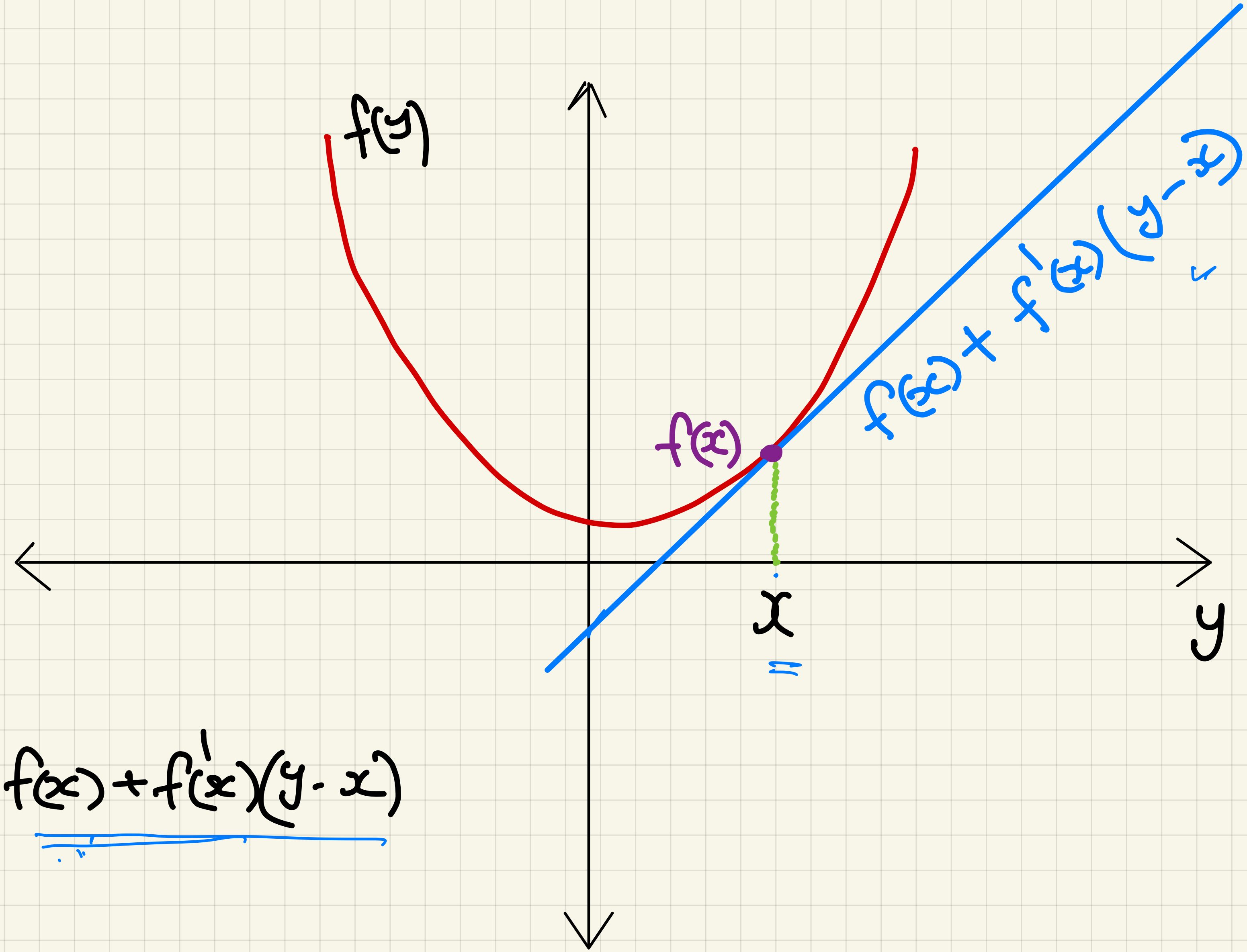
$$f(x) = f(a) + \frac{f'(a)(x-a)}{1!} + \frac{f''(a)(x-a)^2}{2!} + \frac{f'''(a)(x-a)^3}{3!} + \cdots\cdots$$



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

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

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



$$f(y) \geq f(x) + f'(x)(y - x)$$

using
local information $f(x)$, $f'(x)$ we can get a global conclusion.

Second-order conditions

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

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

exists at each $x \in \mathbf{dom}\, f$

$$\begin{array}{cc} & j \\ i \rightarrow & \begin{bmatrix} \cdots & \vdots & \\ \cdots & \frac{\partial^2 f(x)}{\partial x_i \partial x_j} & \end{bmatrix} \end{array}$$

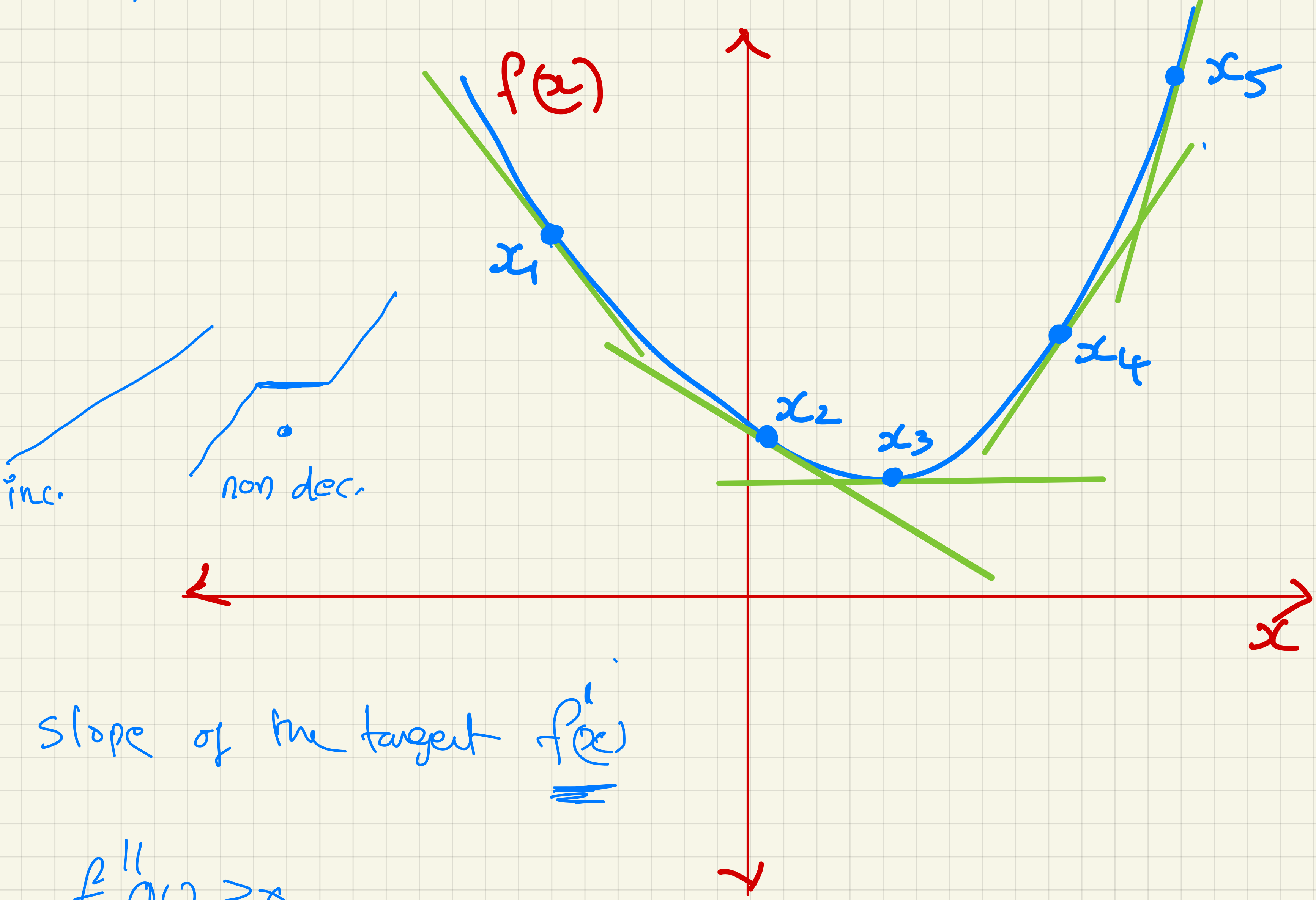
2nd-order conditions: for twice differentiable f with convex domain

- f is convex if and only if

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

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

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



g is non-decreasing

$g' \geq 0$

$\mathbb{R}$

slope of the tangent $f'(x)$

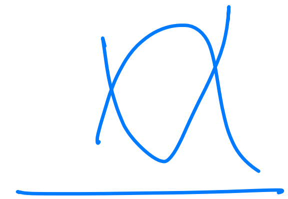
$f''(x) \geq 0$

Slope of the tangent is "increasing" as x increases

The Matrix cookbook
(KB Petersen)

Examples

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



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

$x^T A^T A x$

convex if $P \succeq 0$

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

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

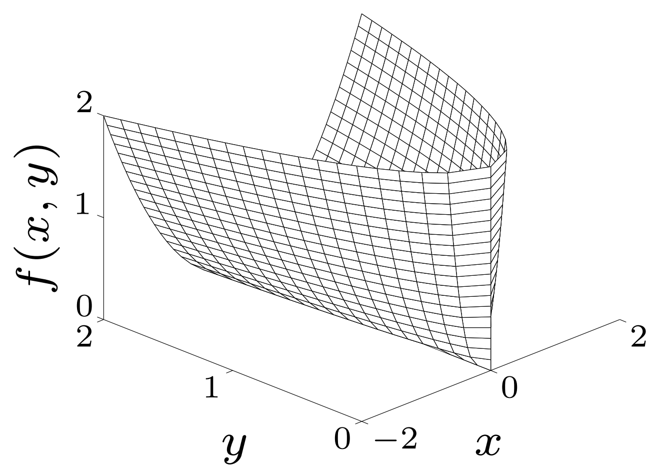
$\|x\|_2^2 = \sqrt{x_1^2 + x_2^2 + \dots}$

convex (for any A)

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

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

convex for $y > 0$



Quadratic-over-linear:

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

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

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

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

soft max function.

$e^1 \quad e^5 \quad e^{10}$

≈ 10

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

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

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

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

$$\nabla f(x) = \frac{1}{\sum_{k=1}^{n} e^{x_k}} \underbrace{\begin{bmatrix} e^{x_1} & e^{x_2} & \cdots & e^{x_k} \end{bmatrix}}_{z} = \frac{z}{1^T z}$$

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

$$= \frac{\operatorname{diag}(z)}{1^T z} - \frac{z}{(1^T z)^2} z^T$$

$$= \left[(1^T z)^2 \operatorname{diag}(z) - z z^T \right] / (1^T z)^2$$

to check positive definiteness

$$\boxed{(z^T v)^T = [z_1 v_1 \; z_2 v_2 \cdots z_k v_k]}$$

$$v^T \nabla^2 f(x)\, v = \left(v^T (1^T z)\, \mathrm{diag}(z)\, v - \underbrace{v^T z\, z^T v}_{(z^T v)^T (z^T v)} \right) \Big/ (1^T z)^2$$

$$= \left[(1^T z) \begin{bmatrix} v_1 & v_2 & \cdots & v_k \end{bmatrix} \begin{bmatrix} z_1 & 0 & 0 \cdots 0 \\ 0 & z_2 & 0 \quad 0 \quad 0 \\ \vdots & & \\ 0 & 0 \cdots & z_k \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_k \end{bmatrix} - \sum_{i=1}^{k} (z_i v_i)^2 \right] \Big/ \left(\sum_{i=1}^{k} z_i \right)^2$$

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

$$\geq 0$$

Cauchy-Swartz inequality.

$$|\langle \underline{u}, \underline{v} \rangle| \leq \|u\|_2 \|v\|_2$$

$$\underline{u} = [z_1^{1/2} \; z_2^{1/2} \cdots z_k^{1/2}]$$

$$\underline{v} = [z_1^{1/2} v_1 \cdots z_k^{1/2} v_k]$$

Epigraph and sublevel set

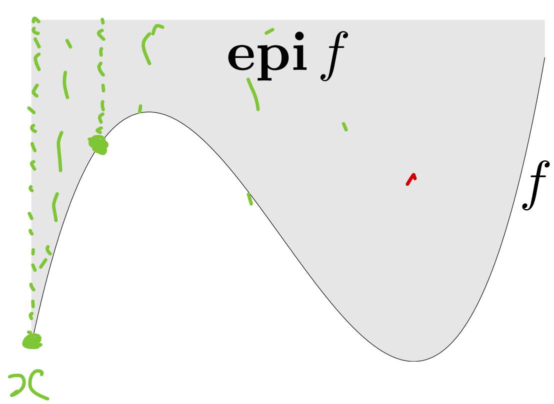
α-**sublevel set** of $f : \mathbf{R}^n \to \mathbf{R}$:

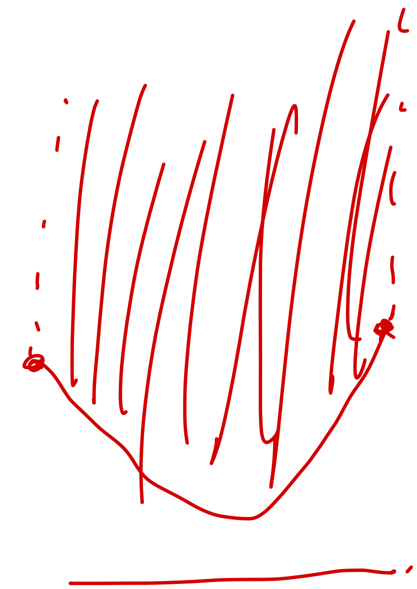
$$C_\alpha = \{x \in \mathbf{dom}\, f \mid f(x) \leq \alpha\}$$

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

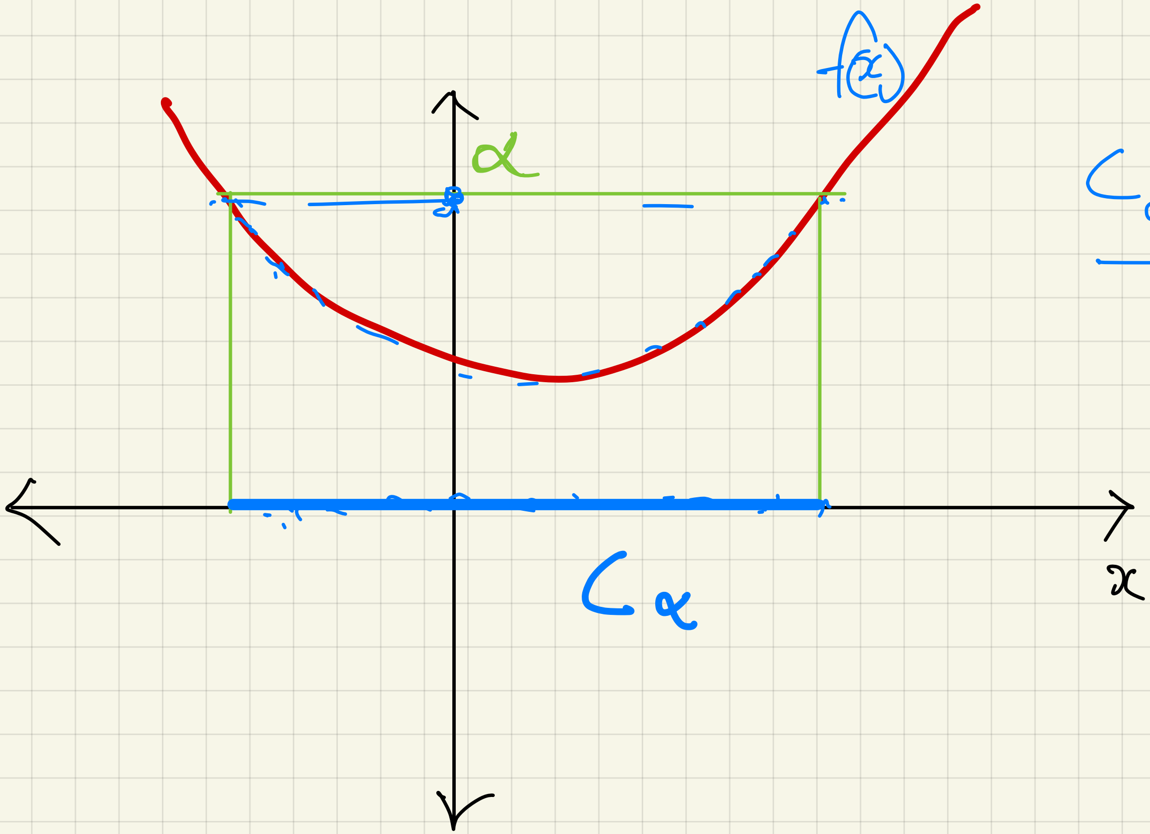
epigraph of $f : \mathbf{R}^n \to \mathbf{R}$:

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

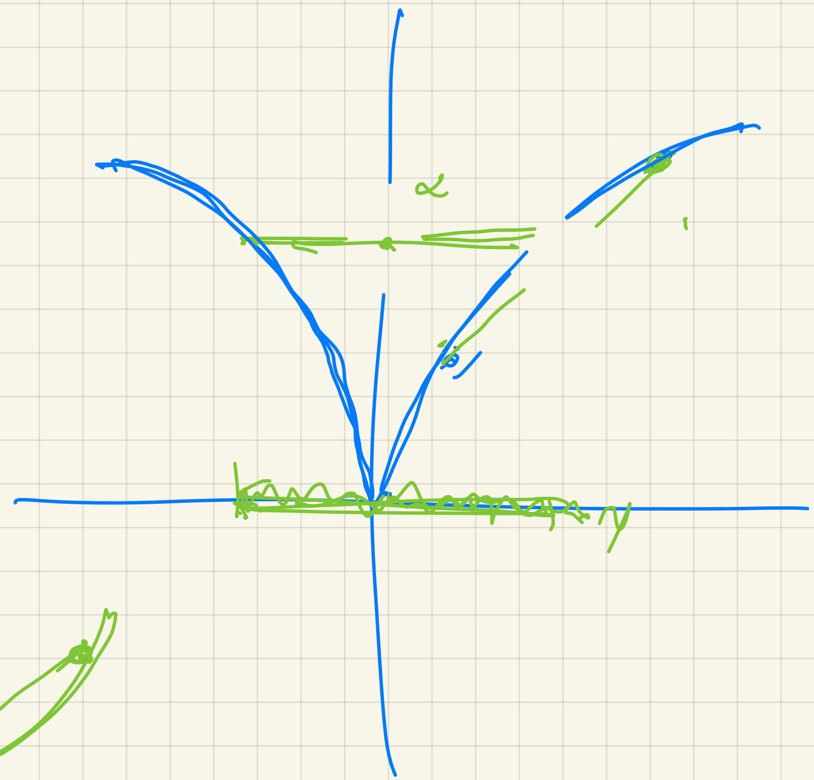


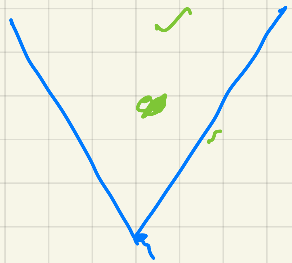


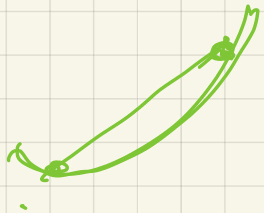
f is convex if and only if $\mathbf{epi}\, f$ is a convex set



$$\underline{C_\alpha} = \{ x \in \mathrm{dom} f,\ f(x) \le \alpha \}$$







Convex

not differentiable.

1^{st} and 2^{nd} order Conditions (RECAP)

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

f is differentiable

f is not differentiable

① 1^{st} 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 2^{nd} order condition.

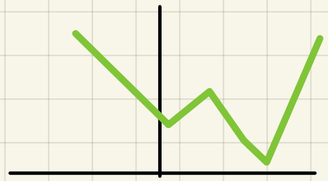
f is twice differentiable with convex domain then f is convex iff

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

* f can be convex or non convex.

* Convexity is checked using other methods.

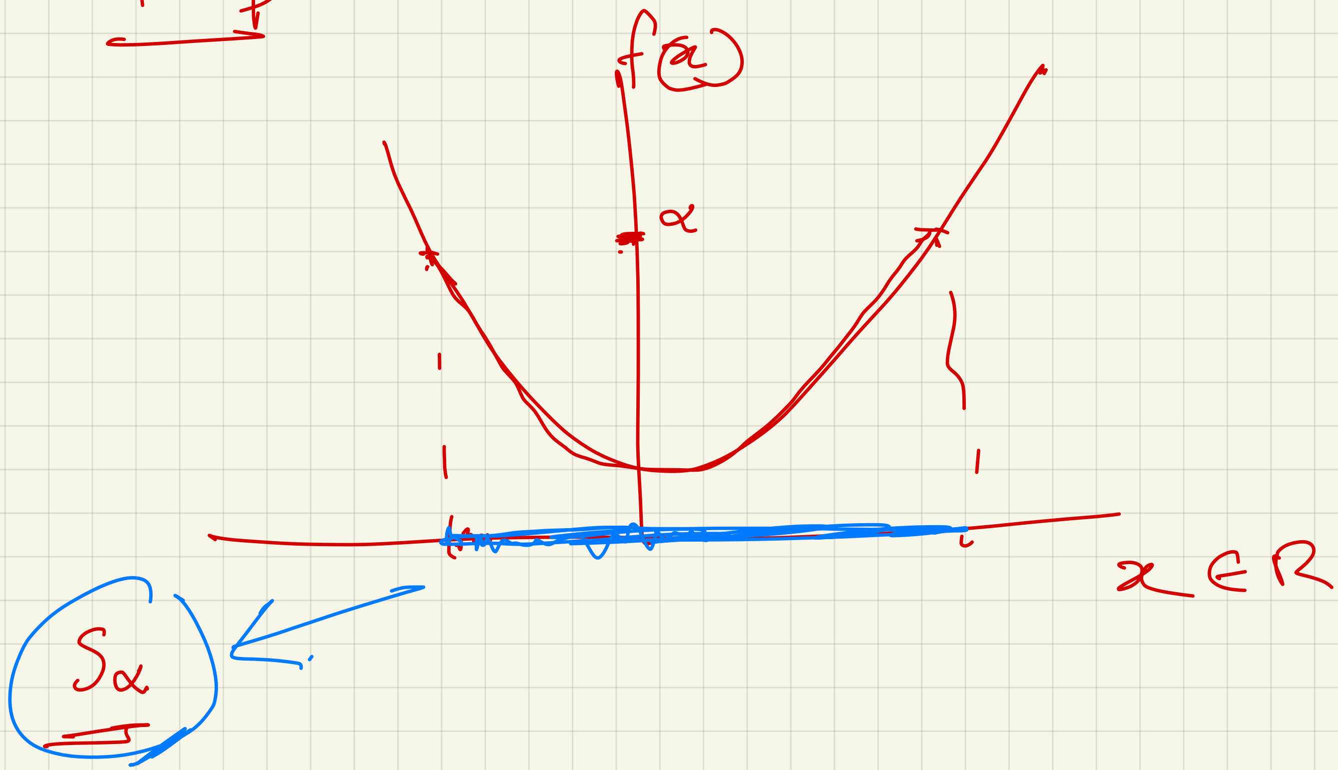
Convex

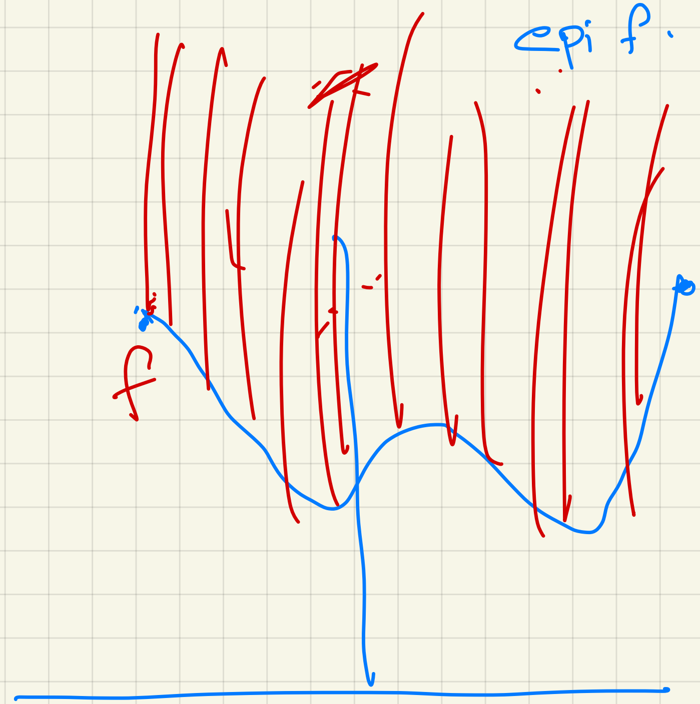


Non Convex

Recap

α-sublevel set.





Jensen's inequality

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

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

$f(Ex) \leq E(f(x))$

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

$$\mathbf{prob}(z = x) = \theta, \qquad \mathbf{prob}(z = y) = 1 - \theta$$

$$f(\theta_1 \underline{x}_1 + \theta_2 \underline{x}_2 + \cdots \theta_n \underline{x}_n) \le \theta_1 f(x_1) + \theta_2 f(x_2) \cdots + \theta_n f(x_n)$$

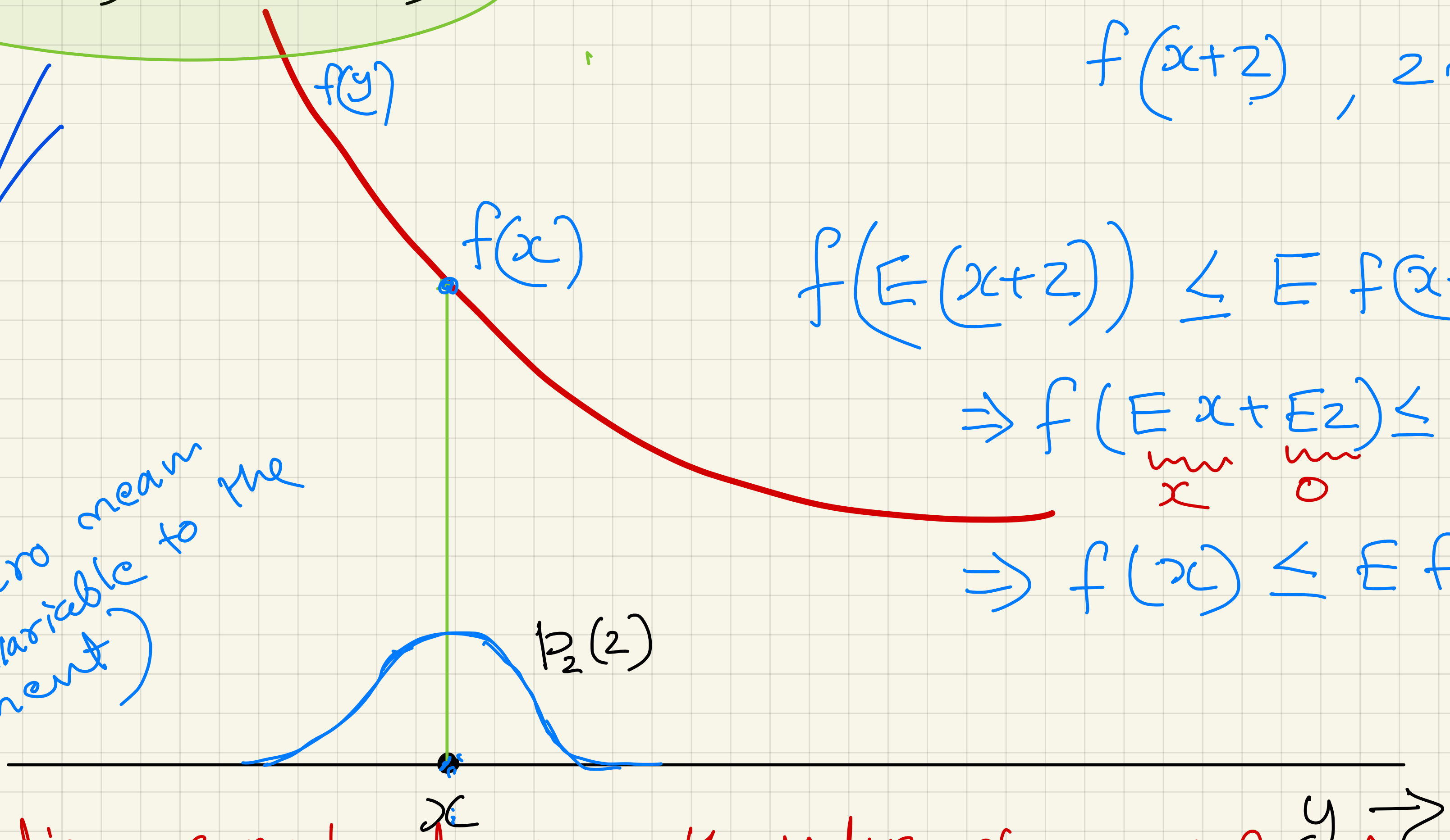
$\sum \theta_i = 1$

$\theta_i \in [0, 1]$

$$f\left(\int_{\mathrm{dom} f} \underline{x}\, p(\underline{x})\, d\underline{x}\right) \le \int_{\mathrm{dom} f} f(\underline{x})\, p(\underline{x})\, d\underline{x}$$

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

convex function.



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

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

$$\Rightarrow f(\underbrace{E x}_{x} + \underbrace{E z}_{0}) \le E f(x+z)$$

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

(adding a zero mean random variable to the argument)

Randomization cannot decrease the value of a cvx 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

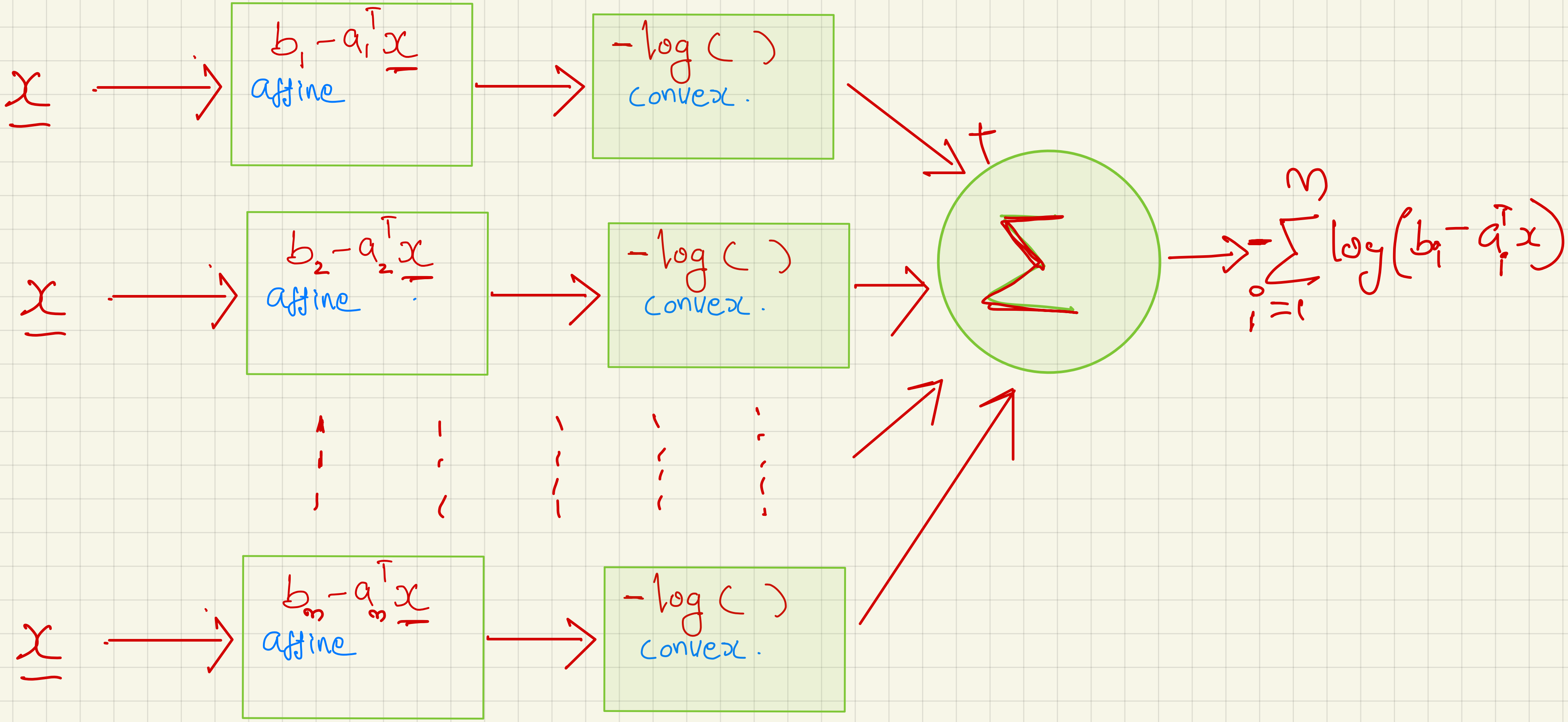
examples

- log barrier for linear inequalities

$$f(x) = -\sum_{i=1}^{m} \log(b_i - a_i^T x), \qquad \mathbf{dom}\, f = \{x \mid a_i^T x < b_i, i = 1, \ldots, m\}$$

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

$\| Ax - c \|^2$



$\underline{x}$
$b_1 - a_1^T \underline{x}$
affine
$-\log(\ )$
convex.
$\underline{x}$
$b_2 - a_2^T \underline{x}$
affine
$-\log(\ )$
convex.
$\underline{x}$
$b_m - a_m^T \underline{x}$
affine
$-\log(\ )$
convex.
+
Σ
$-\sum_{i=1}^{m} \log(b_i - a_i^T x)$

Pointwise maximum

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

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

examples

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

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

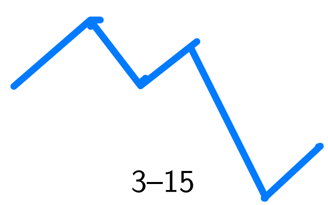
$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix}$$

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

proof:

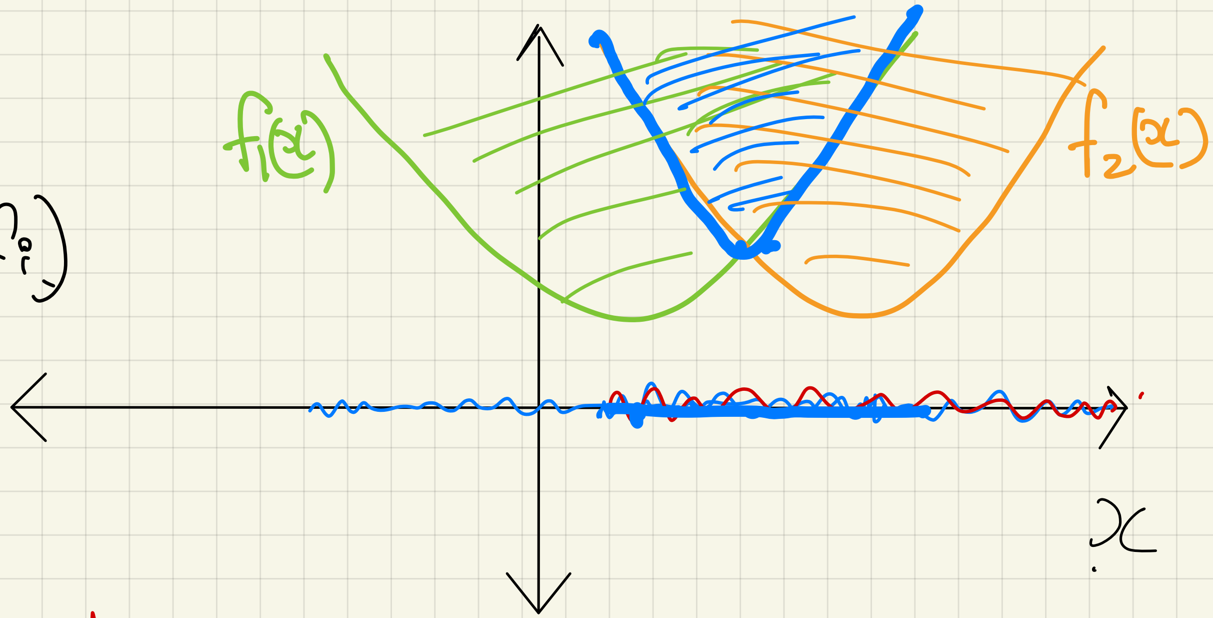
$$f(x) = \max\{x_{i_1} + x_{i_2} + \cdots + x_{i_r} \mid 1 \le i_1 < i_2 < \cdots < i_r \le n\}$$

NOTE: All piecewise-linear function are not convex



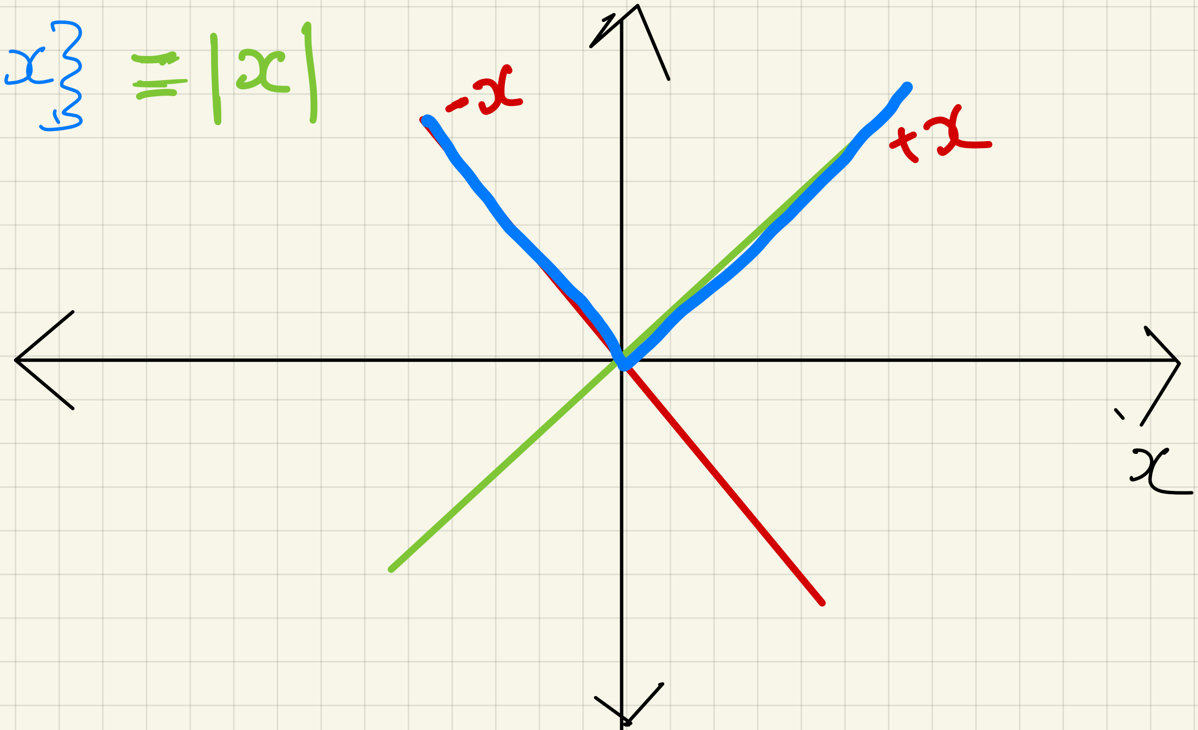
$$f(x) = \max(f_1(x), f_2(x))$$

$$\operatorname{epi}(\max_i f_i) = \bigcap_i \operatorname{epi}(f_i)$$



piece-wise linear function.

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



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)

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

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

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

$$\max\{\overset{2+3}{a+b}, \overset{1+6}{c+d}\} = 7$$

$$\leq \max\{\underset{2,\ 1}{a, c}\} + \max\{\underset{3,\ 6}{b, d}\}$$

$$2 + 6 = 8$$

Sum of r largest componet $\equiv \max_{i} \underline{a}_i^T \underline{x}$

$\boxed{r=3}$

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

Choose ${}^nC_r$ different $\underline{a}_i$'s by varying the positions of '1's

$${}^{50}C_3$$

Pointwise supremum

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

$\mathcal{A}$ 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

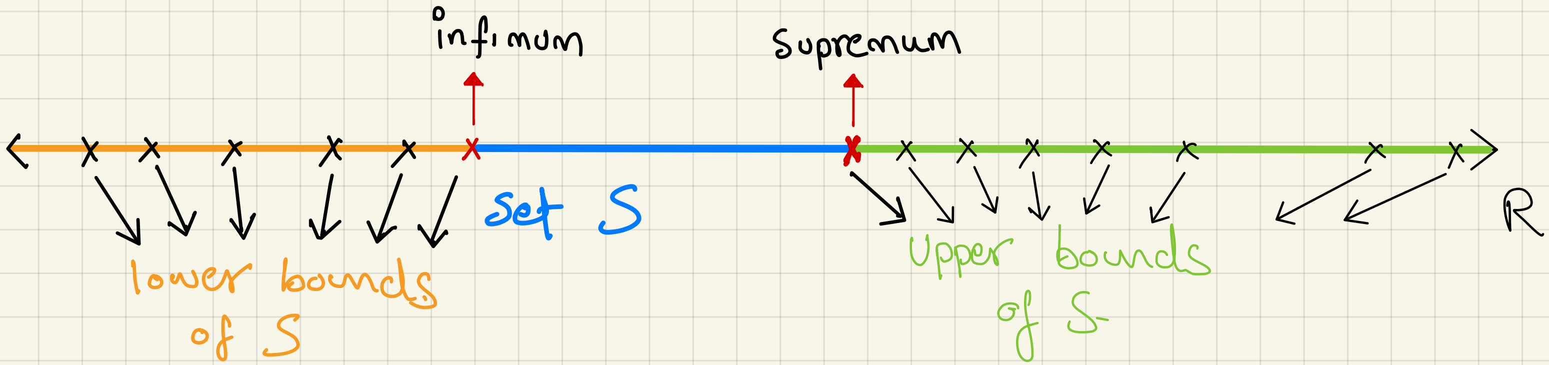
$S \subseteq \mathbb{R}$

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

eg ①: $S = [1\ 5)$

$\max S = ?$

$\sup S = 5$

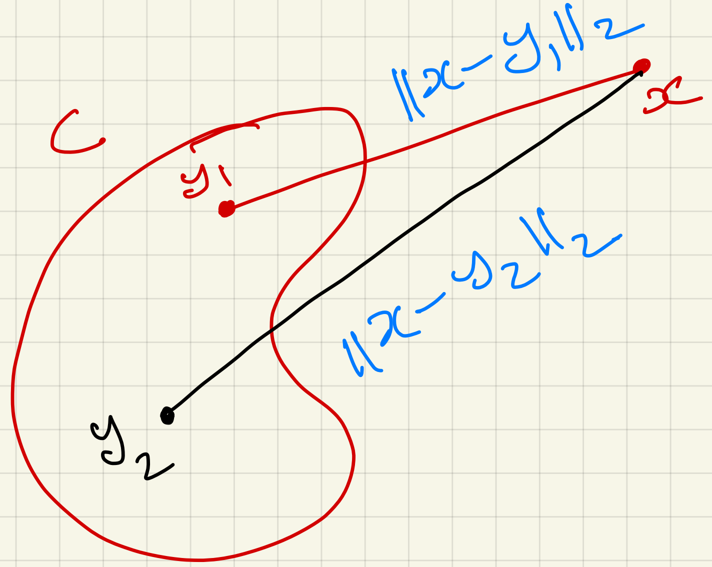


eg. ②: $S = [1, 5]$

$\max S = 5$

$\sup S = 5$

Infemum: largest lower bound.



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

$$\left.\begin{array}{l} \underline{y}_1^T X y_1 \\[2ex] \underline{y}_2^T X y_2 \\[2ex] \vdots \end{array}\right\} \longrightarrow \underbrace{\text{linear function}}_{\downarrow \ \text{convex}} \text{ in } X$$

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

Composition with scalar functions

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

$f(Ax+b)$; h ; g

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

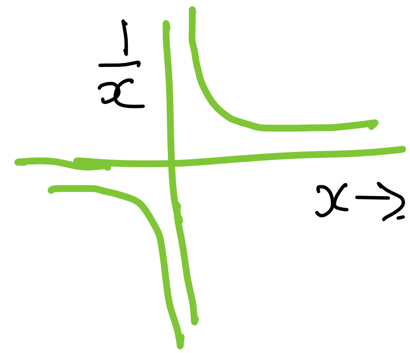
f is convex if $\begin{array}{l} g \text{ convex, } h \text{ convex, } \tilde{h} \text{ nondecreasing} \\ g \text{ concave, } h \text{ convex, } \tilde{h} \text{ nonincreasing} \end{array}$

h	g
$\exp$	$g(x)$
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 ↓
h is cvx

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

$h(y) = 1/y$

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

$g : R^n \to R$
$h : R \to R$

Proof $(n=1)$

$Ax+b$

$$f'(x) = h'(g(x))\,\overset{A}{g'(x)}$$

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

for $f(x)$ to be convex.

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

$+$

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

g is convex

h is nondecreasing $\uparrow$

$-$

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

g is concave

h is non increasing $\downarrow$

$\dot{f}(Ax+b)$

$h() \to f()$

$g() = A() + b$

h cvx, g cvx $\} \Rightarrow$ ~~cvx~~

$\nabla h()^T g''$

$\nabla h_1() g_1'' + \nabla h_2() g_2'' + \cdots$

Vector composition

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

$\|Ax+b\| \to$ cvx

$f''(Ax+b) AA^T$

$$f(x) = h(g(x)) = h(g_1(x), g_2(x), \ldots, g_k(x))$$

$h() = \|\ \|$ cvx

$g() = A() + b$ cvx

f is convex if $\begin{array}{l} g_i \text{ convex, } h \text{ convex, } \tilde{h} \text{ nondecreasing in each argument} \\ g_i \text{ concave, } h \text{ convex, } \tilde{h} \text{ nonincreasing in each argument} \end{array}$

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

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

≤ 0

① $\to$ ≤ 0 ≥ 0 ≤ 0

≥ 0

② $\to$ ≥ 0 ≥ 0 ≥ 0

$g \to$ concave

$h \to$ concave.

$h \to g$

examples

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

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

log-sum-exp $\to$ cvx

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

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

- $f(x, y) = x^T A x + 2x^T B y + y^T C y$ with

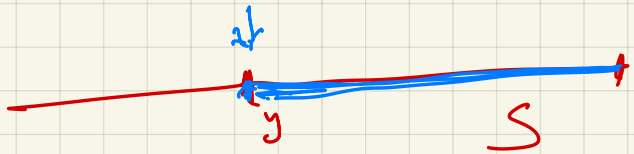
$$\begin{bmatrix} A & B \\ B^T & C \end{bmatrix} \succeq 0, \qquad C \succ 0$$

  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: $\mathbf{dist}(x, S) = \inf_{y \in S} \|x - y\|$ is convex if S is convex

Proof



Let $\epsilon > 0$

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

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

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

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

$3 \le 2 + \epsilon$

$a \le b + \epsilon$ $\therefore$

$a \le b$

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

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

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

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

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

$$(C^{-1})^T = C^{-1}$$

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

Since C is symmetric.

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

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

Schur complement

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

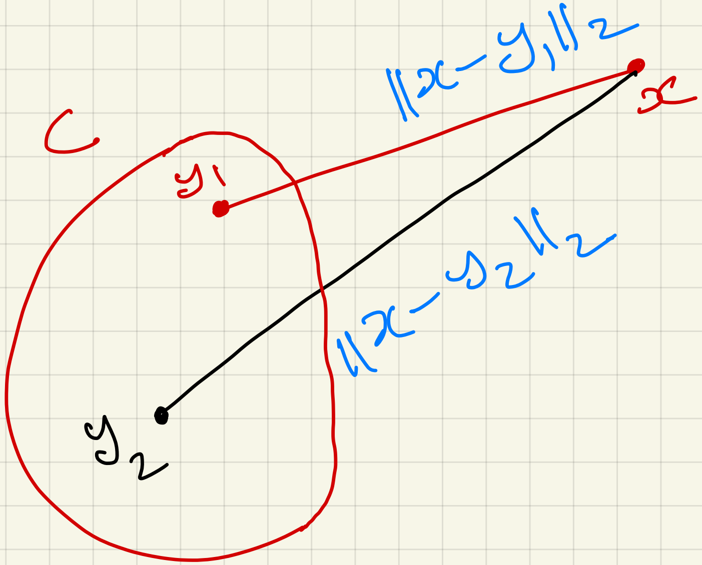
$$X = \begin{bmatrix} A & B \\ B^T & C \end{bmatrix},$$

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

$\begin{bmatrix} x_1/t \\ x_2/t \\ t \end{bmatrix}$

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

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

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

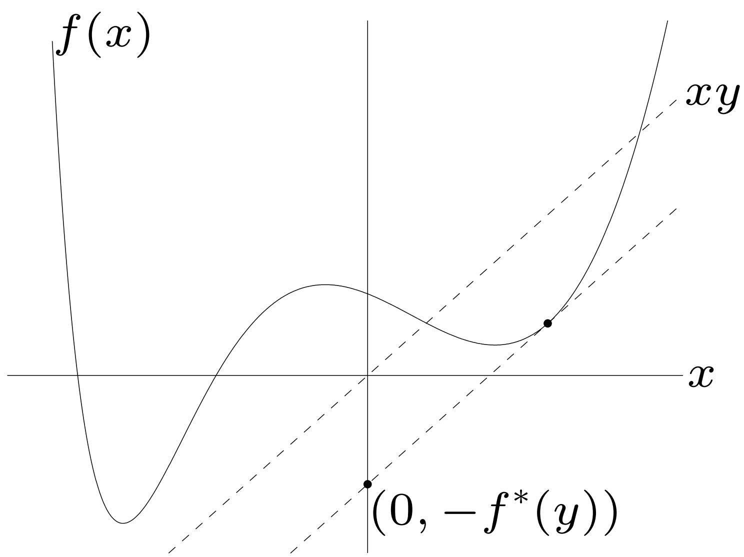
$$g(x) = (c^T x + d) f\left((Ax+b)/(c^T x + d)\right)$$

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

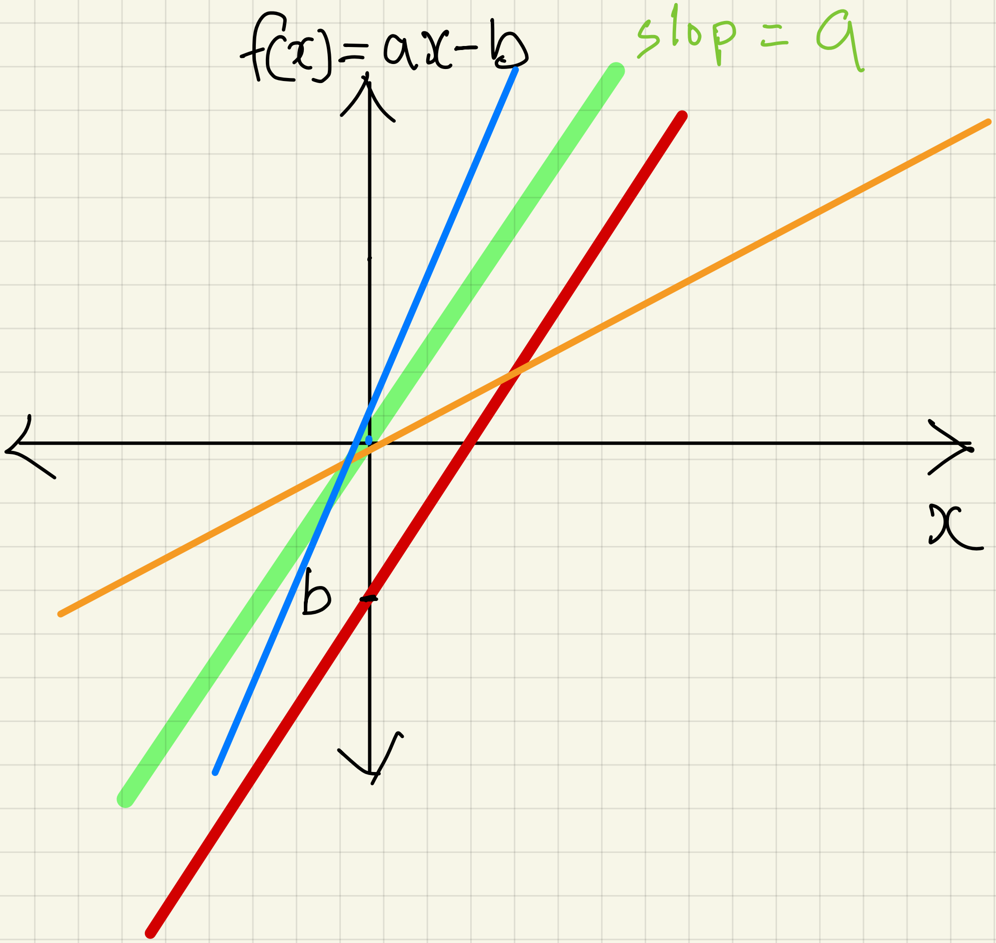
The conjugate function

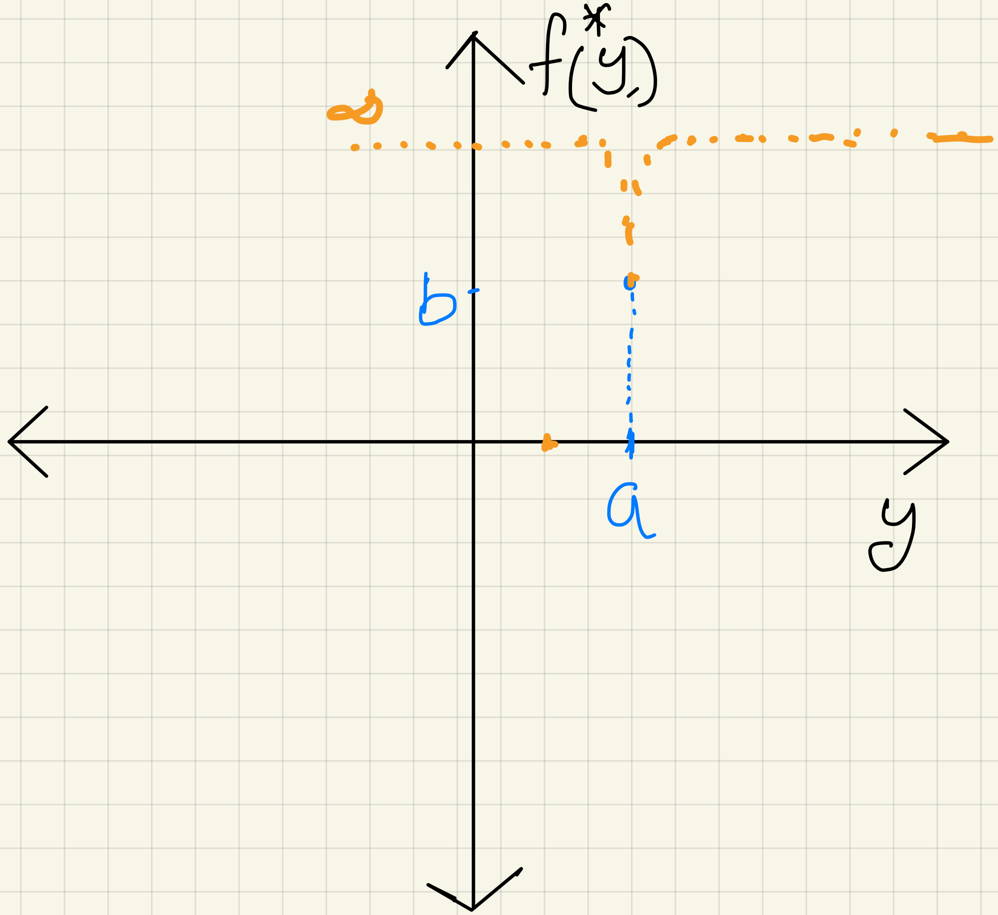
the **conjugate** of a function f is

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



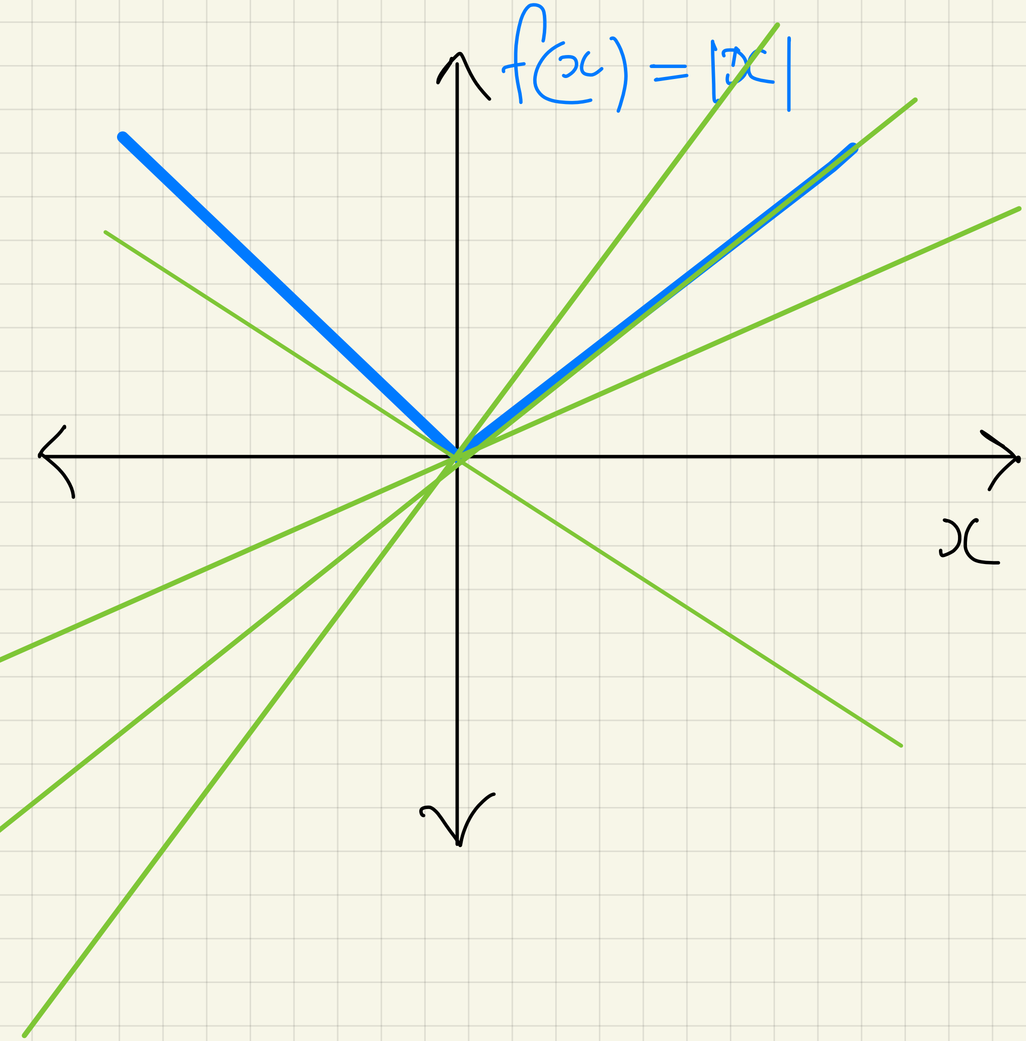
- f^* is convex (even if f is not)
- will be useful in chapter 5



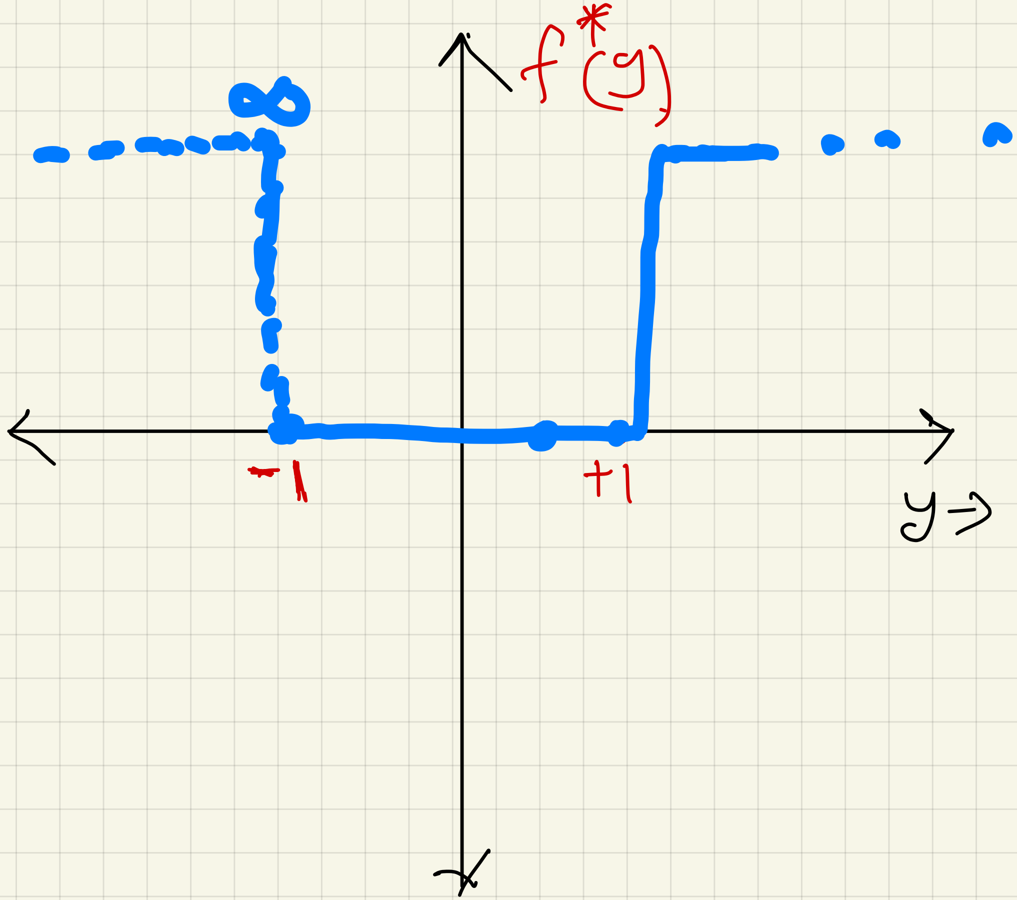


$$f^*(y) = \sup_{x\in\mathbb{R}} \left(yx - (ax-b) \right)$$

$$= \sup_{x\in\mathbb{R}} \left((y-a)x + b \right) = \begin{cases} b & \text{if } y=a \\ \infty & \text{if } y\neq a \end{cases}$$



f(x) = |x|
x



f*(y)
∞
-1
+1
y →

examples

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

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

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

$g'(x^*) = y + \frac{1}{x^*} = 0$

$x^* = (-1/y)$

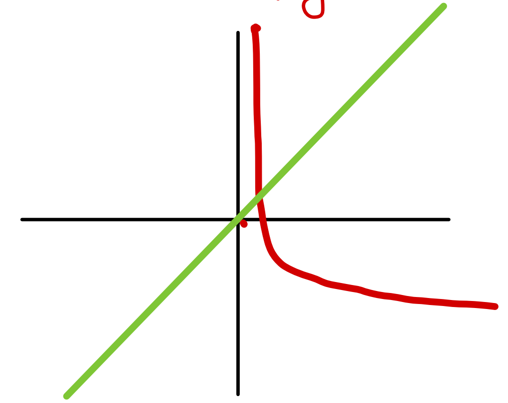
$g(x^*) = -1 + \log\left(\frac{-1}{y}\right)$

$= -1 - \log(-y)$

- strictly convex quadratic $f(x) = (1/2)x^TQx$ with $Q \in \mathbf{S}^n_{++}$

$$f^*(y) = \sup_x(\overbrace{y^Tx - (1/2)x^TQx}^{g(x)})$$

$$= \frac{1}{2}y^TQ^{-1}y$$



$g'(x) = y - Qx^* = 0$

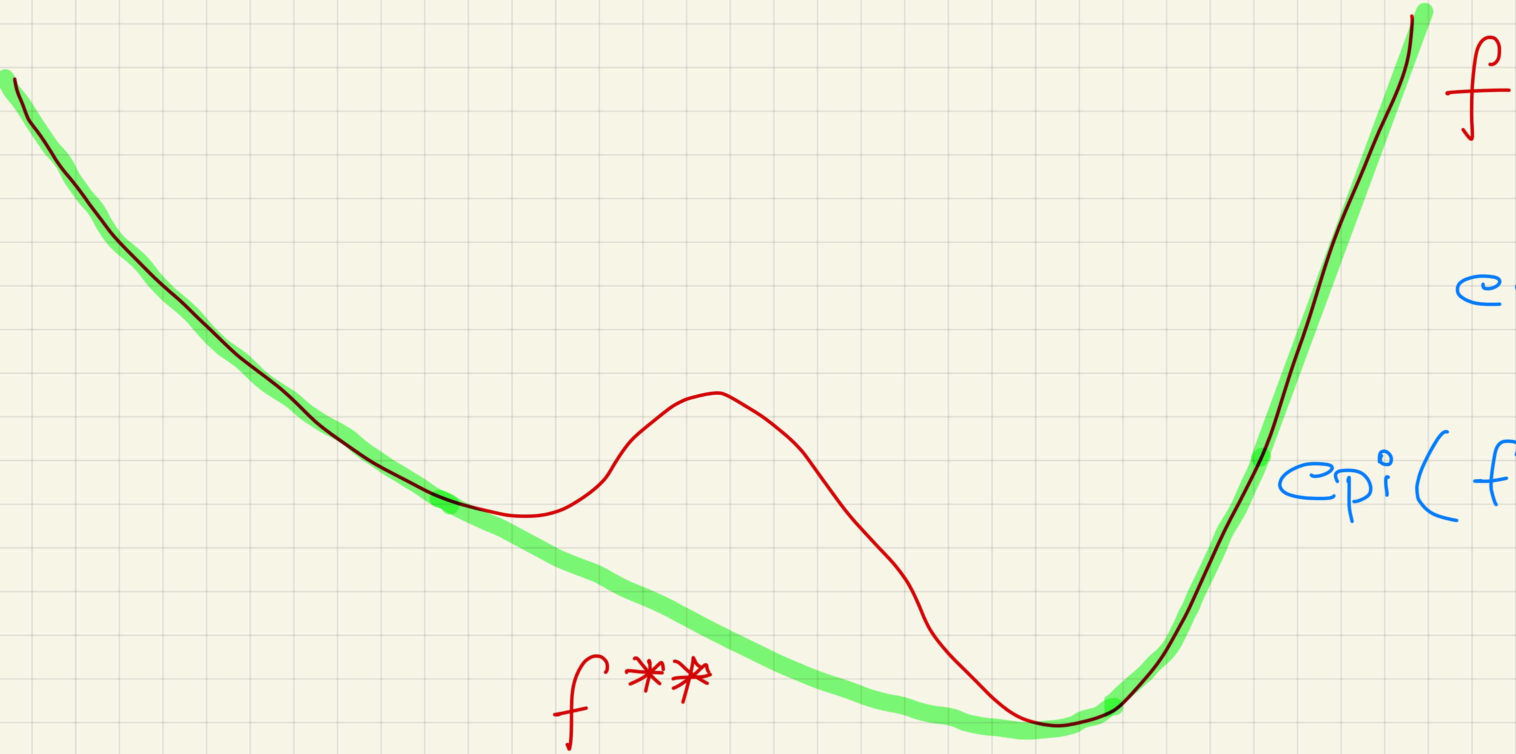
$\Rightarrow x^* = Q^{-1}y$

$g(x^*) = y^TQ^{-1}y - \frac{1}{2}(Q^{-1}y)^TQ(Q^{-1}y)$

$= y^TQ^{-1}y - \frac{1}{2}y^T\cancel{Q^{-1}}\cancel{Q}Q^{-1}y$

$= \frac{1}{2}y^TQ^{-1}y$

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



f^{**} is the convex envelop (f^{env}) of f

$\mathrm{epi}(f^{env}) = \mathrm{conv}(\mathrm{epi}\ f)$

convex hull.

if f is convex

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

(Note: f should be closed.)

Quasiconvex functions

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

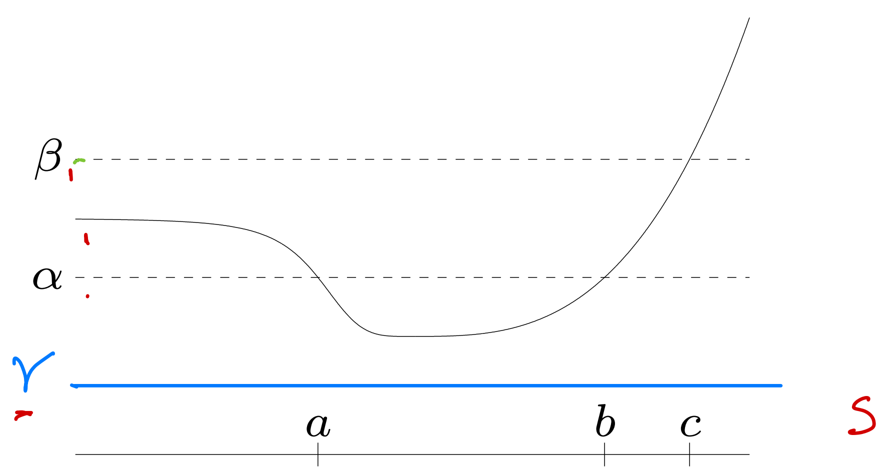
$$S_\alpha = \{x \in \mathbf{dom}\, f \mid f(x) \leq \alpha\}$$

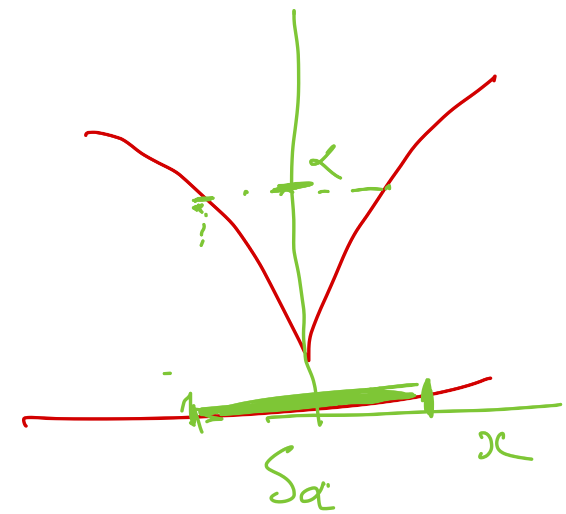
are convex for all α

$S_\beta = (-\infty, c]$ ✓

$S_\alpha = [a, b]$ ✓

$S_\gamma = \phi$ –

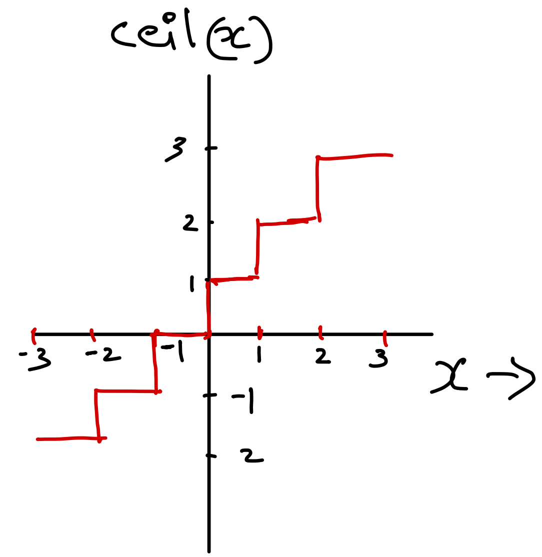




- f is quasiconcave if $-f$ is quasiconvex or α-super level set is convex
- f is quasilinear if it is quasiconvex and quasiconcave

$S_\alpha^{sup} = \{x \in \mathrm{dom}\, f \mid f(x) \geq \alpha\}$

Examples



- $\sqrt{|x|}$ is quasiconvex on **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

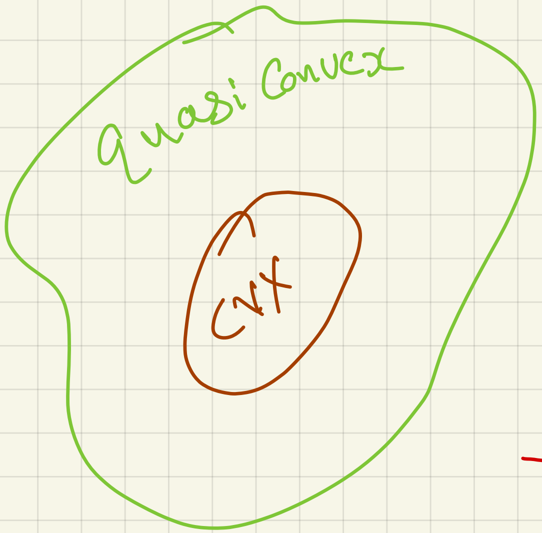
$$f(x) = \frac{a^T x + b}{c^T x + d}, \qquad \mathbf{dom}\, f = \{x \mid c^T x + d > 0\}$$

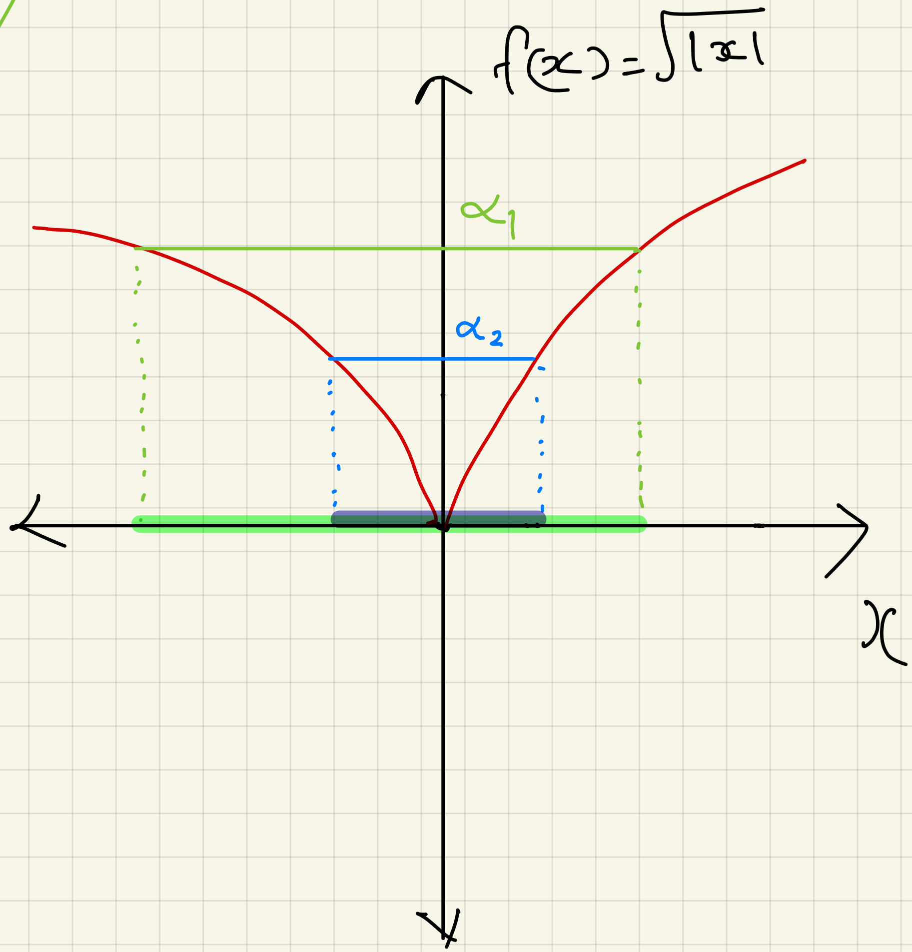
  is quasilinear

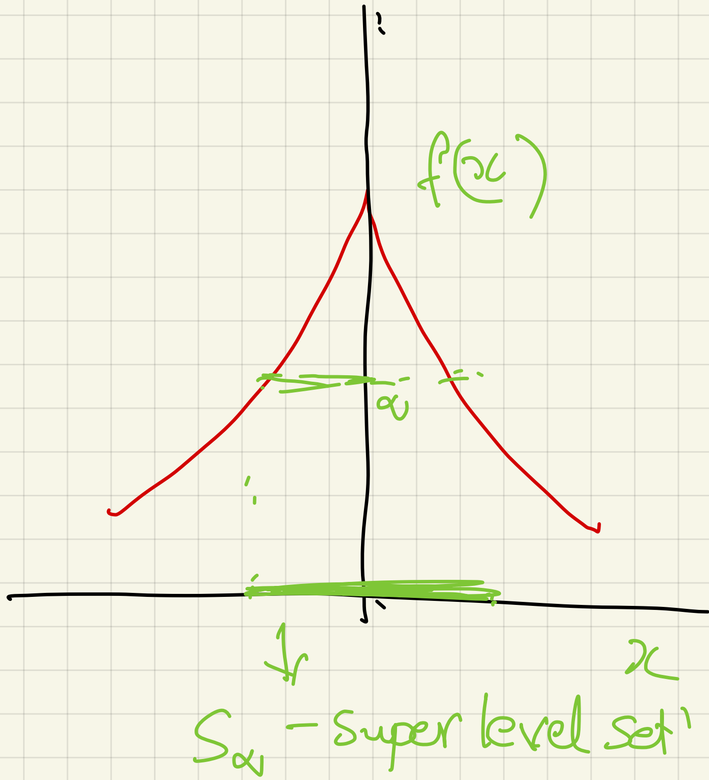
- distance ratio

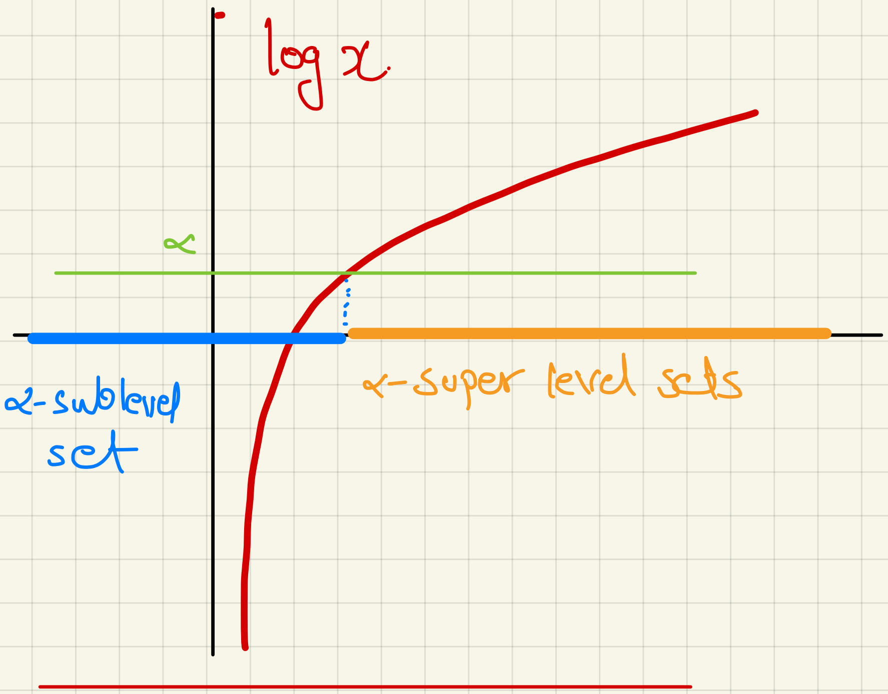
$$f(x) = \frac{\|x - a\|_2}{\|x - b\|_2}, \qquad \mathbf{dom}\, f = \{x \mid \|x - a\|_2 \leq \|x - b\|_2\}$$

  is quasiconvex


quasi convex
cvx


$f(x) = \sqrt{|x|}$
α_1
α_2
x


$f(x)$
α
x
S_α - super level set



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

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

Let $\underline{x} \in S_\alpha^{sup} \Rightarrow x_1 x_2 \geq \alpha$

$\quad \underline{y} \in S_\alpha^{sup} \Rightarrow y_1 y_2 \geq \alpha$

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

$\theta \in [0, 1]$

$$(\theta x_1 + (1-\theta) y_1)(\theta x_2 + (1-\theta) y_2)$$

$$= \theta^2 x_1 x_2 + (1-\theta)^2 y_1 y_2 + \theta(1-\theta)(x_1 y_2 + x_2 y_1)$$

$$= \theta^2 x_1 x_2 + (1-\theta)^2 y_1 y_2 + \theta(1-\theta)\left(\frac{x_1}{x_2} x_2 y_2 + \frac{x_2}{x_1} x_1 y_1\right)$$

$$\geq \alpha\theta^2 + \alpha(1-\theta)^2 + \theta(1-\theta)\alpha\left(\frac{x_1}{x_2} + \frac{x_2}{x_1}\right)$$

$$= \alpha\theta^2 + \alpha(1-\theta)^2 + \theta(1-\theta)\alpha\left(\underbrace{\left(\sqrt{\frac{x_1}{x_2}} - \sqrt{\frac{x_2}{x_1}}\right)^2}_{>0} + 2\right)$$

$$\geq \alpha\theta^2 + \alpha(1-\theta)^2 + \theta(1-\theta)\alpha(2)$$

$$= \alpha(\theta + 1 - \theta)^2 = \alpha.$$

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

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

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

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

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

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

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 $\{x \mid \|x - a\|_2 \le \|x - b\|_2\}$. To see this, we consider the α-sublevel set of f, with $\alpha \le 1$ since $f(x) \le 1$ on the halfspace $\{x \mid \|x - a\|_2 \le \|x - b\|_2\}$. This sublevel set is the set of points satisfying

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

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

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

$B(x_c, r)$

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

internal rate of return

$r = 8\%$

$x \xrightarrow{\text{1 year}} x + (.08)x$

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

0 1 2 3

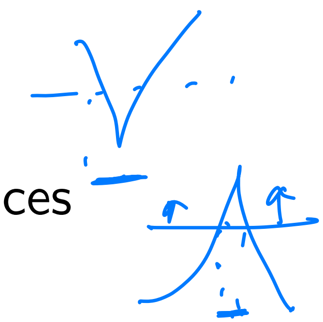
-10000 2000 3000 1000

$$\mathrm{PV}(x, r) = \sum_{i=0}^{n} (1+r)^{-i} x_i$$

$$-10000 + \frac{2000}{1+r} + \frac{3000}{(1+r)^2} + \cdots$$

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

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



IRR is quasiconcave: superlevel set is intersection of open halfspaces

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

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

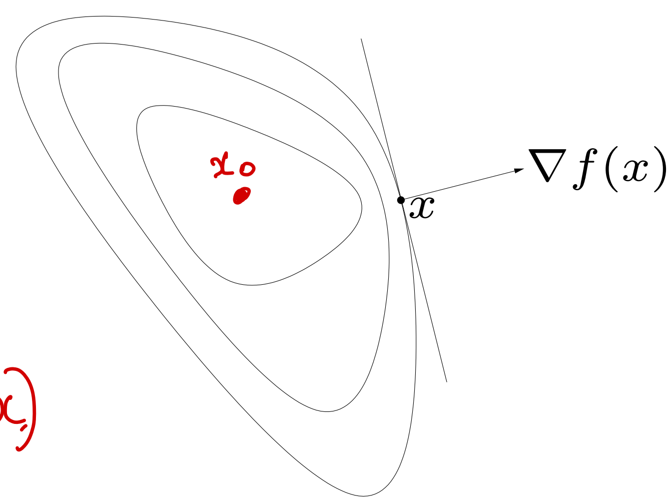
Properties

modified Jensen inequality: for quasiconvex f

$$0 \le \theta \le 1 \quad \Longrightarrow \quad f(\theta x + (1-\theta)y) \le \max\{f(x), f(y)\}$$

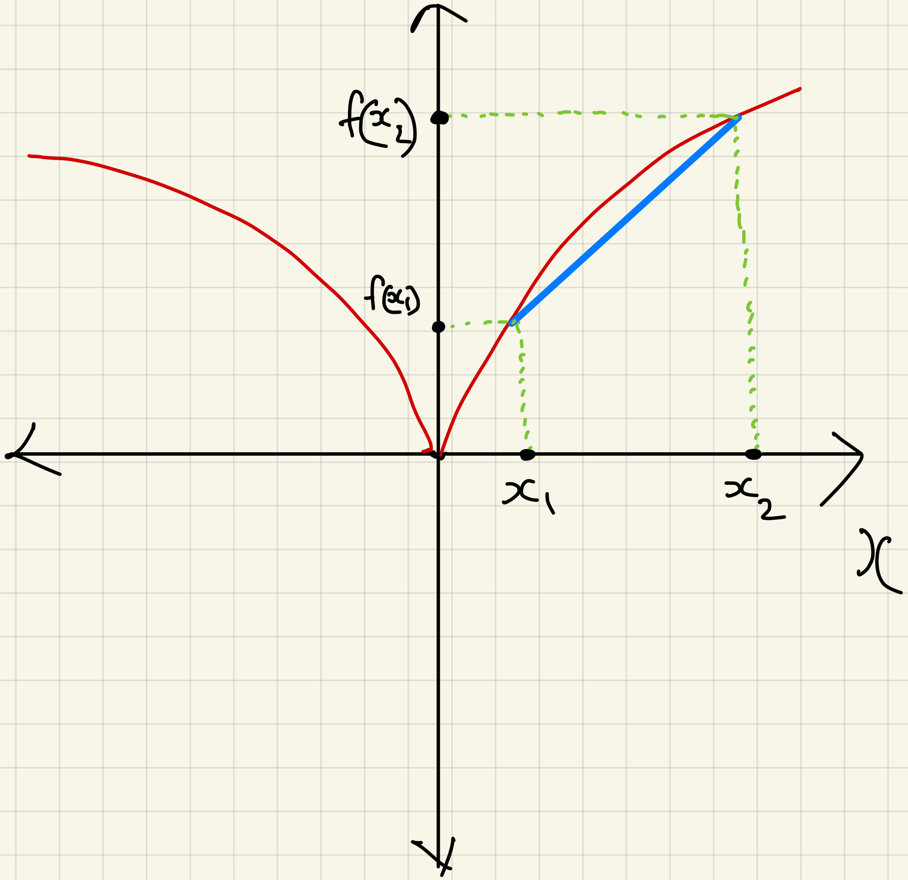
first-order condition: differentiable f with cvx domain is quasiconvex iff

$$f(y) \le f(x) \quad \Longrightarrow \quad \nabla f(x)^T (y - x) \le 0$$



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

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

$f(x) = x^a,$

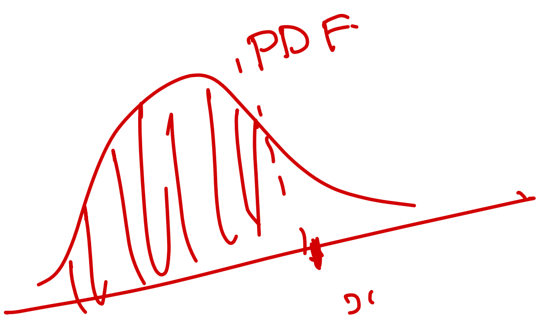
$\log f(x) = a \log x$

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

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

$K e^{-(\quad)}$ $\quad -\frac{1}{2}(x-\bar{x})^T \Sigma^{-1} (x-\bar{x})$

- 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.

$-\infty$ $\quad x$

log-concave function.

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

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

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

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

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 $x \in \mathbf{dom}\, f$

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

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

is log-concave (not easy to show)

$f_1, f_2 \rightarrow$ log-concave

$\left.\begin{array}{l}\log f_1 \\ \log f_2\end{array}\right\} \rightarrow$ concave.

$g = f_1 \cdot f_2$

$\log g = \underbrace{\log f_1 + \log f_2}$.

sum of 2 concave functions is concave

$\Rightarrow g$ is log concave

Further reading: properties of log-concave function.

$$g(x) = \log f(x)$$

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

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

$g(x)$ is log concave $\iff$ $\nabla^2 g(x) \preceq 0$

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

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

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

  is log-concave

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

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

  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

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

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

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

proper cone.

for x, $y \in \mathbf{dom}\, f$, $0 \le \theta \le 1$

example $f : \mathbf{S}^m \to \mathbf{S}^m$, $f(X) = X^2$ 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.*,

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

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

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

$$\left.\begin{array}{l} f : S^n \to S^n \\ f(X) = X^2 \end{array}\right\} \Rightarrow S^n_+ \text{ - convex.}$$

PROOF

$$z^T X^T X z = (Xz)^T (Xz)$$

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

$$X, Y \in S^n$$

$$\Rightarrow h(\theta X + (1-\theta)Y) \le \theta h(X) + (1-\theta) h(Y) \to \text{defenition of convexity}, \quad \theta \in [0,1]$$

$$\underline{z}^T (\theta \underline{X} + (1-\theta)Y)^2 \underline{z} \le \theta \underline{z}^T X^2 \underline{z} + (1-\theta) \underline{z}^T Y^2 \underline{z}$$

$$\Rightarrow \theta \underline{z}^T X^2 \underline{z} + (1-\theta) \underline{z}^T Y^2 \underline{z} - z^T (\theta \underline{X} + (1-\theta)Y)^2 \underline{z} \ge 0$$

$$\Rightarrow \underline{z}^T \underbrace{\left[\theta X^2 + (1-\theta) Y^2 - (\theta \underline{X} + (1-\theta)Y)^2\right]}_{PSD} \underline{z} \ge 0$$

$$\Rightarrow \quad \theta X^2 + (1-\theta) Y^2 - (\theta \underline{X} + (1-\theta) Y)^2 \succeq_{S_+^n} 0$$

$$\Rightarrow \quad (\theta \underline{X} + (1-\theta) Y)^2 \preceq_{S_+^n} \theta X^2 + (1-\theta) Y^2$$

$\Rightarrow$ $f(X) = X^2$ is S_+^n - convex //