conjugate functions

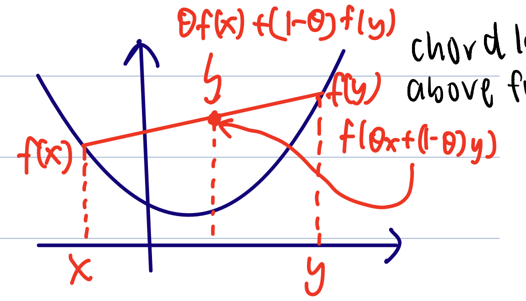
quasi-convex

DSA3102: Lecture 4 (Reading Sections 3.3 & 3.4)

Convex Function: $f: \mathbb{R}^n \to \mathbb{R}$ is convex if dom f is a convex set & $\forall x, y \in$ dom f

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

for all $\theta \in [0,1]$.



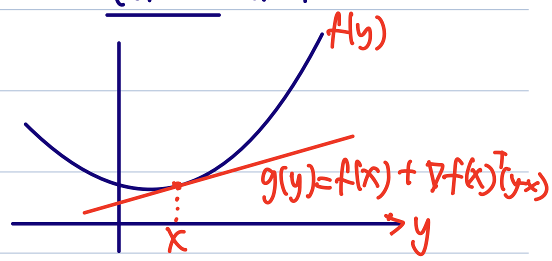
chord lie above function

First-order condition: Assume $f: \mathbb{R}^n \to \mathbb{R}$ is convex and diff^ble. Then $\forall x, y \in$ dom f

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

e.g. $|x|$ not differentiable at 0

tangent if one-dimensional



Second-order condition: Assume $f: \mathbb{R}^n \to \mathbb{R}$ is convex and twice diff^ble. Then $\forall x \in$ dom f,

Hessian

$$\nabla^2 f(x) \succeq 0$$

i.e., $\nabla^2 f(x)$ is positive semi-definite.

Conjugate Functions

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

need not be convex

sweep through all possible y and create a new function

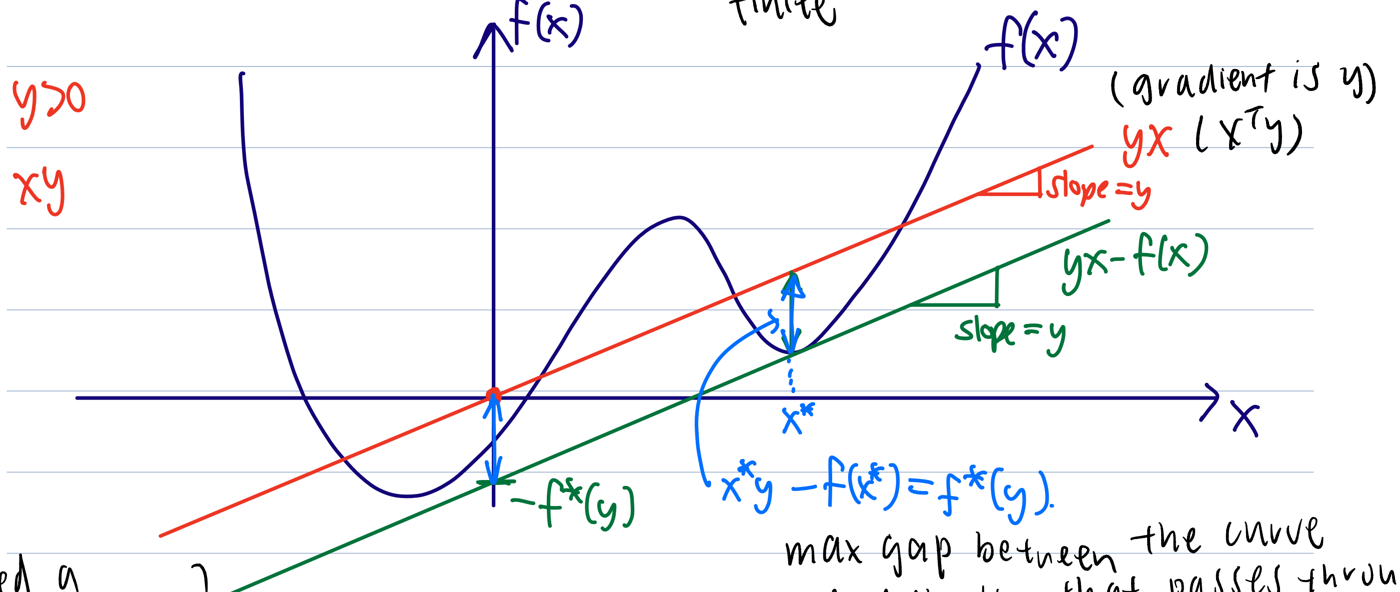
Def: The <u>conjugate function</u> of f is $f^*: \mathbb{R}^n \to \mathbb{R}$ defined as

$$f^*(y) = \sup_{x \in \operatorname{dom} f} \{ \underbrace{x^T y - f(x)}_{g(x)} \}$$

$g(x)$ and take maximum

$$\operatorname{dom} f^* = \{ y \in \mathbb{R}^n : \underbrace{\sup_{x \in \operatorname{dom} f} x^T y - f(x) < \infty}_{\text{finite}} \}.$$

maximise gap between curve and red line



max gap between the curve and the line that passes through origin with slope y

Created a convex function?

Rmk: $f^*(y) = \sup_{x \in \operatorname{dom} f} \{ \boxed{x^T y - f(x)} \}$ → affine in y

Even if f is not convex, f^* is convex.

– $f^*(y)$ is the supremum of a bunch of <u>affine</u> and hence <u>convex</u> functions ($x^T y - f(x)$ is affine in y).

---

linear function

Ex: $f(x) = ax + b \quad x \in \mathbb{R}$ affine function.

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

$y \in \mathbb{R}$

domain contains values of y that does NOT give infinity

$\boxed{y \text{ can choose}}$ $\quad = \{ \sup_{x \in \mathbb{R}} x(y-a) \} - b$

not of interest (only maximising over x so b can be brought out)

$\nearrow$ sup $(x(a-a)) = 0$

$$f^*(y) = \begin{cases} -b & y = a \\ +\infty & y \neq a \end{cases}$$

f^* is finite only in this singleton

In other words $\text{dom } f^* = \{a\}$ & $f^*(a) = -b$.

Ex: Negative logarithm.

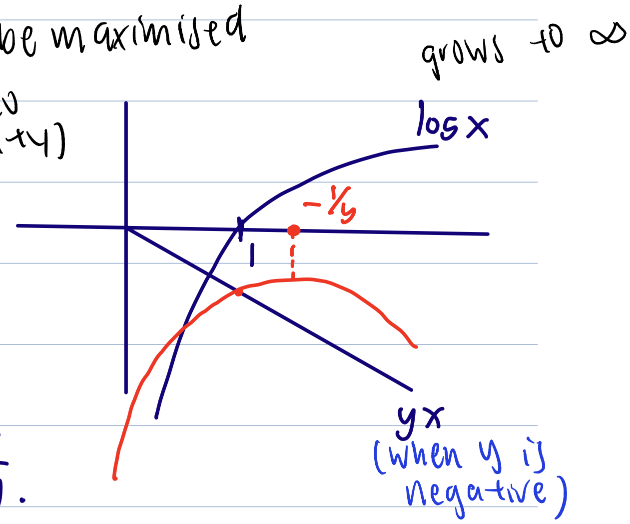
$f(x) = -\log x \qquad x \in \mathbb{R}_{++}$

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

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

$g(x)$ to be maximised

$y=0 \Rightarrow g(x) = \log x$ (goes to infinity)



a) $y \geq 0$: $\sup_{x>0} g(x) = +\infty$.

$\hookrightarrow$ not inside domain of f^*

xy decreasing but $\log x$ is increasing $\Downarrow$ turning point

b) $y < 0$: $g'(x) = 0$

$$\Rightarrow y + \frac{1}{x} = 0 \Rightarrow x^* = -\frac{1}{y}.$$

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

In conclusion, the conjugate of f is

$$f^*(y) = -\log(-y) - 1$$

need -ve y in order for argument of log to be +ve

& $\text{dom } f^* = -\mathbb{R}_{++}$ negative numbers

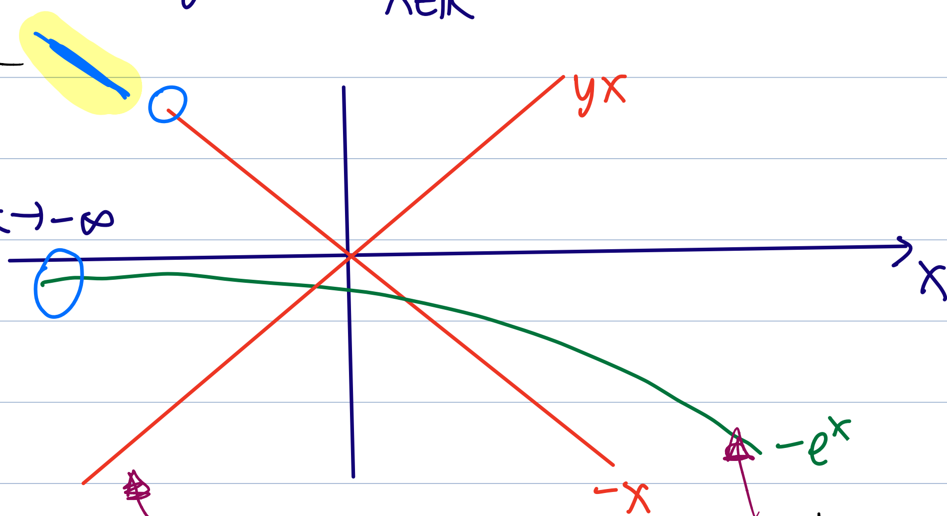
Ex: Exponential

$f(x) = e^x \qquad x \in \mathbb{R} \qquad \text{dom } f = \mathbb{R}$.

Consider $f^*(y) = \sup_{x\in\mathbb{R}} \{xy - e^x\}$.

a) $y<0$ $\quad y=-1 \quad \sup_{x\in\mathbb{R}} -x-e^x = +\infty$

supremum is +ve (attained when x goes to $-\infty$)



(one is increasing, one is decreasing, so we get maximum point)

b) $y>0 \quad \sup_{x\in\mathbb{R}} \underbrace{xy-e^x}_{g(x)} \qquad g'(x)=0$

$$\Downarrow$$

$$y=e^x \Rightarrow x^*=\log y$$

$f^*(y) = \sup_{x\in\mathbb{R}} xy - e^x = \boxed{y\log y - y}$ put in the x^* that we deem to be optimal

this is convex

c) $y=0 \qquad f^*(y) = \sup_{x\in\mathbb{R}} -e^x = 0.$

$$f^*(y) = \begin{cases} 0 & y=0 \\ y\log y - y & y>0 \end{cases} \qquad \operatorname{dom} f^* = \mathbb{R}_+.$$

$= y\log y - y, \; y\geq 0$ (with the interpretation that $0\log 0 = 0$)

justified because of continuity

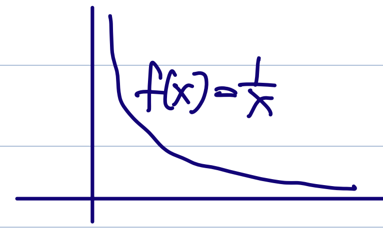
$h(y) = y\log y$

$h(0) = \lim_{y\to 0^+} h(y) = 0$ (L'Hospital's rule)

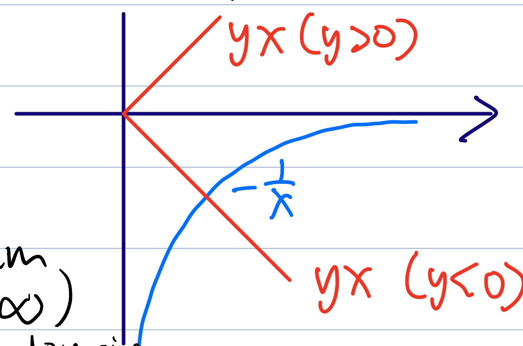
Inverse

$f(x) = 1/x$ $\qquad$ $\text{dom } f = \mathbb{R}_{++}$

+ve quadrant only



$f^*(y) = \sup_{x>0} \; xy - \frac{1}{x}.$



a) $y>0$: Then $g(x) = xy - \frac{1}{x}$ is unbounded above. (supremum is $+\infty$) $\to$ not in domain

b) $y<0$: $\quad g(x) = xy - \frac{1}{x} \Rightarrow g'(x) = y + \frac{1}{x^2} = 0$

$x^* = (-y)^{-1/2}$

$f^*(y) = y(-y)^{-1/2} - (-y)^{1/2} = -2(-y)^{1/2}$

c) $y=0$: $\quad g(x) = -1/x$, sup is attained as $x \to +\infty$

$\to$ no need to really be attained unlike max

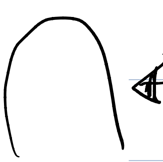
$f^*(y) = 0$

$f^*(y) = -2(-y)^{1/2} \qquad y \le 0 \qquad \text{dom } f^* = -\mathbb{R}_+.$

Ex: $f(x) = \frac{1}{2}x^TQx$ $\quad Q \in S^n_{++}$ positive semi-definite symmetric matrix

Consider $g(x) = \underbrace{x^Ty}_{\text{linear}} - \frac{1}{2}x^TQx.$

$g(x)$ is bounded above as a f^n of x for every $y \in \mathbb{R}^n$.



(plus linear term still give this shape $\Rightarrow$ there is a maximum

$\nabla g(x) = y - Qx = 0$

inverse exist because +ve definite (eigenvalues strictly positive)

$$x^* = Q^{-1} y.$$

$$f^*(y) = (Q^{-1}y)^T y - \frac{1}{2}(Q^{-1}y)^T Q (Q^{-1}y)$$
$$= \frac{1}{2} y^T Q^{-1} y.$$

$\operatorname{dom} f^* = \mathbb{R}^n.$

Log-sum-exp function

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

$x = (x_1, \dots, x_n)^T$ length n vector

Fix $y \in \mathbb{R}^n$

Consider $g(x) = y^T x - \log\left(\sum_{i=1}^{n} e^{x_i}\right)$

$$[\nabla g(x)]_i = \frac{\partial g}{\partial x_i} = y_i - \frac{e^{x_i}}{\sum_{j=1}^{n} e^{x_j}} = 0.$$

The optimal x has components that satisfy

non-linear (cannot solve for X^* from here)

$$\frac{e^{x_i}}{\sum_{j=1}^{n} e^{x_j}} = y_i \qquad \forall i \in [n].$$

$C = \sum_{j=1}^{n} e^{x_j}$ → normalization

These n equations admit a sol^n in x if and only if y is a positive vector ($y_i > 0$) and $\sum_{i=1}^{n} y_i = 1$.

In other words, y must be a prob. vector. (where all components are +ve)

$\forall i \in [n]$, $\quad y_i = \frac{e^{x_i}}{C} \quad$ $C = \sum_{j=1}^{n} e^{x_j}$ → $Cy_i = e^{x_i}$

$\log C + \log y_i = x_i$

Substitute this expression into $g(x) = \boxed{x^T y} - \log \sum_i e^{x_i}$ to get

$$f^*(y) = \boxed{\sum_{i=1}^{n} x_i \frac{e^{x_i}}{C}} - \log \sum_i e^{x_i}$$

$$= \sum_{i=1}^{n} x_i \frac{e^{x_i}}{C} - \log C$$

sum over all the y_i is 1 because it is a probability vector

$$= \sum_{i=1}^{n} (\cancel{\log C} + \log y_i)\, y_i - \cancel{\log C}$$

$$= \sum_{i=1}^{n} y_i \log y_i \qquad \mathrm{dom}\, f^* = \{y \in \mathbb{R}^n : y_i \geq 0, \sum_{i=1}^{n} y_i = 1\}$$

negative entropy.

---

find sparse vectors etc.

$\underline{\ell_1\text{-norm:}}$ $f(x) = \|x\|_1 \qquad x \in \mathbb{R}^n$

$f(x) = \sum_i |x_i|$ (sum of absolute values of x_i)

$$f^*(y) = \sup_{x \in \mathbb{R}^n} y^T x - \sum_{i=1}^{n} |x_i|$$

$$= \sup_{x \in \mathbb{R}^n} \sum_{i=1}^{n} (y_i x_i - |x_i|)$$

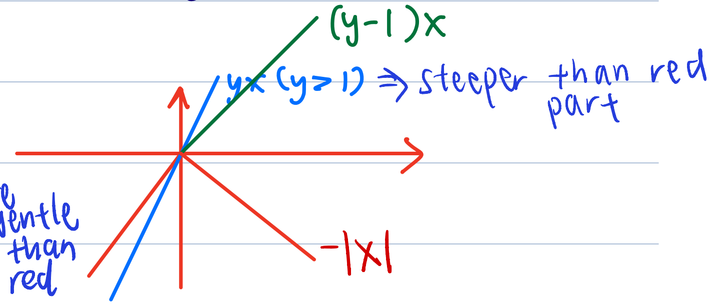
one dimensional case

Special case $n=1$: $g(x) = \underline{yx} - |x|$

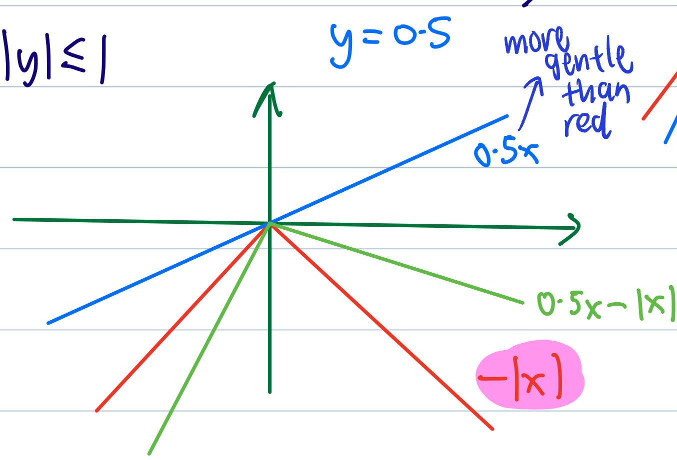
$$f^*(y) = \begin{cases} 0 & |y| \leq 1 \\ +\infty & |y| > 1 \end{cases}$$

sum on RHS? (supremum attained at ∞)

a) $|y|>1$, $g(x)$ is unbounded above.



b) $|y|\leq 1$



$f^*(y)=0$

$\forall\ |y|\leq 1$

$= \max_i |y_i|$ (maximum absolute value of y_i)

$$f^*(y)=\begin{cases} 0 & \text{if } \|y\|_\infty \leq 1 \\ +\infty & \text{if } \|y\|_\infty > 1 \end{cases}$$

Properties of the conjugate

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

f^* is convex.

(max value) bigger than

$\forall x\in \text{dom} f$, $\quad f^*(y)\geq x^Ty-f(x)$

$\Rightarrow f^*(y)+f(x)\geq x^Ty.$ (Fenchel's inequality).

Scaling and composition with affine function

$a>0$, $b\in\mathbb{R}$ $\qquad g(x)=af(x)+b.$ $\qquad$ domain of f is $\text{dom} f$

f has conjugate f^*

finding conjugate of g

Claim: $g^*(y) = a f^*\left(\frac{y}{a}\right) - b$ $\quad$ $\operatorname{dom} g^* = \{y : \frac{y}{a} \in \operatorname{dom} f^*\}$

$f^*(y') = \sup_x y'x - f(x)$

Pf: $g^*(y) = \sup_{x \in \operatorname{dom} f} (xy - g(x))$

a is a constant

↳ can be brought outside

$= \sup_x (xy - af(x) - b)$

$= \left(a \sup_x \left(x\frac{y}{a} - f(x)\right)\right) - b.$

$= a f^*\left(\frac{y}{a}\right) - b$

maximising over x so x disappear

Sum of independent functions.

$f(u,v) = f_1(u) + f_2(v).$

f_i : functions with conjugates f_i^*.

Claim: $f^*(y,z) = f_1^*(y) + f_2^*(z).$

where $f_1^*(y)$, $f_2^*(z)$ are the conjugates of f_1 & f_2 resp.

indep functions

Application: $f(x) = \|x\|_1 = |x_1| + |x_2|$

from 1 dimensional case to N dimensional case for l1 function

---

double conjugate

Thm: Let $f : \mathbb{R}^n \to \mathbb{R}$ be any function. Then

i) $f(x) \geq f^{**}(x) \quad \forall x \in \mathbb{R}^n$

ii) If f is closed (epi f is closed) and convex, then

$f^{**}(x) = f(x) \quad \forall x \in \mathbb{R}^n.$ $\quad$ use strict seperating hyperplane Thm

Pf: i) $f^*(y) = \sup_x x^Ty - f(x)$

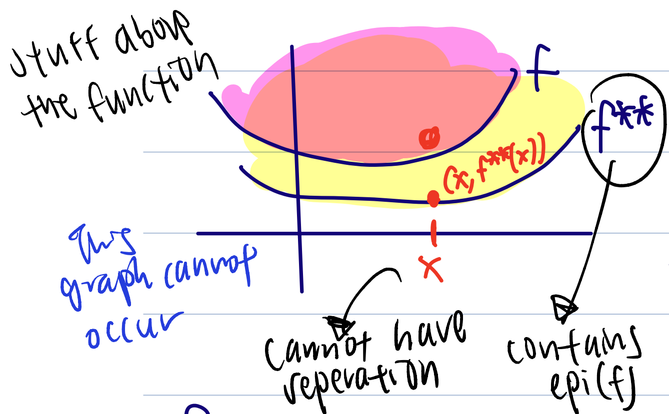
$f^*(y) \geq x^Ty - f(x) \qquad \forall x \in \mathbb{R}^n$

$$\Rightarrow \quad f(x) \ge x^T y - f^*(y) \quad \forall x, y$$

We may maximize the RHS over y which yields

$$f(x) \ge \underbrace{\sup_y x^T y - f^*(y) = (f^*)^*(x)}_{\text{conjugate of the conjugate}} =: f^{**}(x). \quad ////$$

ii) By (i), $f \ge f^{**} \Rightarrow \mathrm{epi}(f^{**}) \supseteq \mathrm{epi}(f)$



Suffices for us to show that

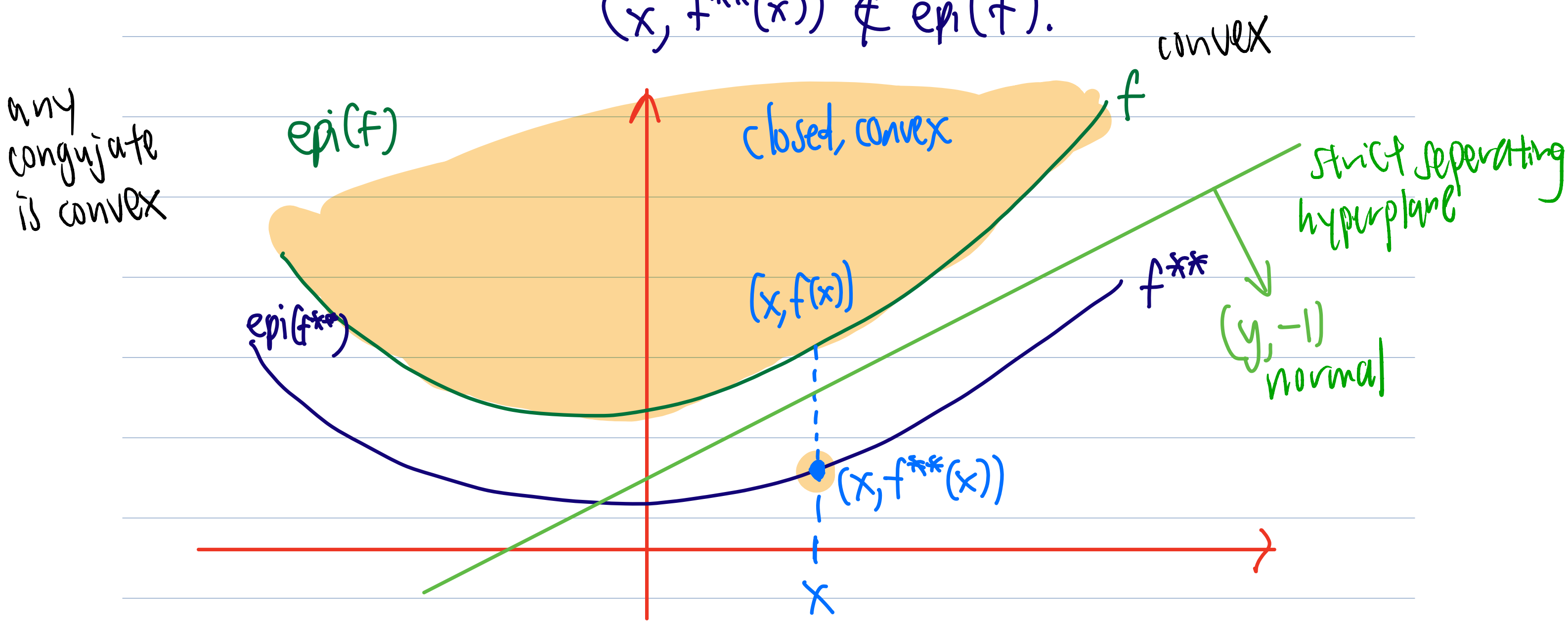
$$\mathrm{epi}(f^{**}) \subseteq \mathrm{epi}(f)$$

Suffices for us to show that

$$\forall x \in \mathrm{dom}(f^{**}), \quad (x, f^{**}(x)) \in \mathrm{epi}(f).$$

Suppose, to the contrary, $\exists\, x \in \mathrm{dom}(f^{**})$ s.t.

$$(x, f^{**}(x)) \notin \mathrm{epi}(f).$$



By the strict separating hyperplane thm (applied to $\underbrace{\underbrace{\{(x, f^{**}(x))\}}_{\text{singleton}} \;\&\; \mathrm{epi}(f)}_{\text{strict seperation}}$ $\longrightarrow$ one must be bounded, both must be closed direction is all that matters

$\exists\ (y, -1)$ s.t.

$$y^Tz - s < c < y^Tx - f^{**}(x)$$

$\forall\ (z,s) \in epi(f) \Rightarrow s \geq f(z)$

We may take $s = f(z)$

$$y^Tz - f(z) < c < y^Tx - f^{**}(x)$$

But the LHS holds $\forall\ z \in \mathrm{dom}\, f$. Hence, we may maximize over $z \in \mathrm{dom}\, f$

$$\sup_z \{ y^Tz - f(z) \} < c < y^Tx - f^{**}(x)$$

$$\Downarrow$$

definition of congugate

$$f^*(y) < c < y^Tx - f^{**}(x)$$

$$\Downarrow$$

contradicts (so seperation cannot happen)

$$f^*(y) + f^{**}(x) < y^Tx$$ $\Rightarrow\Leftarrow$ Fenchel's ineq.

Fenchel's inequality : $g^*(y) + g(x) \geq x^Ty$

Take $g = f^*$

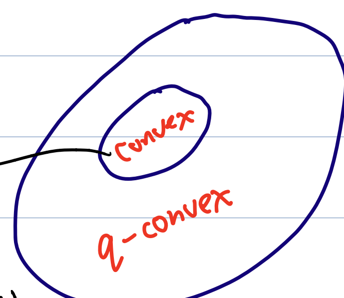
Quasiconvex functions

$f: \mathbb{R}^n \to \mathbb{R}$ is quasiconvex if $\mathrm{dom}\, f$ is convex & the sublevel sets

$$S_\alpha = \{ x \in \mathrm{dom}\, f : f(x) \leq \alpha \}$$

are convex $(\alpha \in \mathbb{R})$

if set is convex, its sub level sets are convex



f : Quasiconcave if $-f$ quasiconvex.

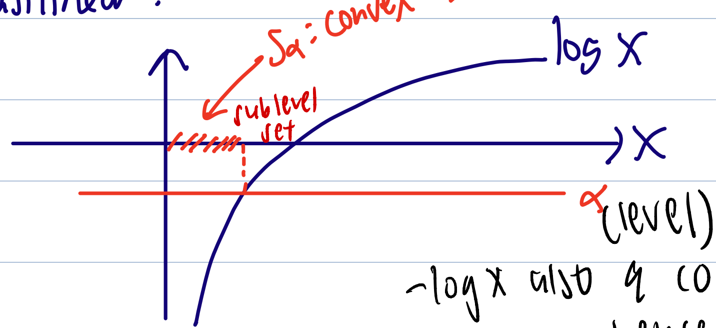
Both q-convex & q-concave $\equiv$ q-linear.

Ex: $\log x$ on $\mathbb{R}_{++}$: concave

But $\log x$ is quasilinear. (both q-concave and q-convex)

S_α : convex $\Rightarrow$ $\log x$ is q-convex.

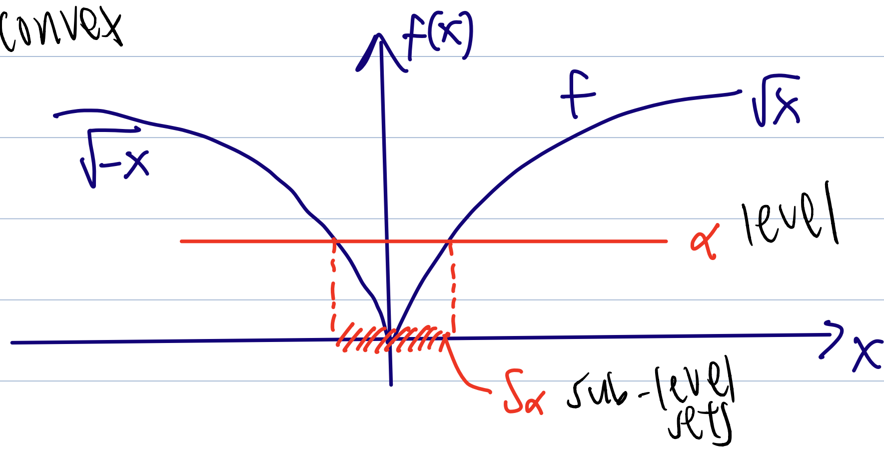
all intervals on $\mathbb{R}$ are convex sets



$-\log x$ also q convex hence whole thing is q-linear

Eg: $f(x) = \begin{cases} \sqrt{x} & x \geq 0 \\ \sqrt{-x} & x < 0 \end{cases}$ quasiconvex.

f neither concave nor convex but it is q-convex



Ex: $f(x_1, x_2) = x_1 x_2$ dom $f = \mathbb{R}_+^2$

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

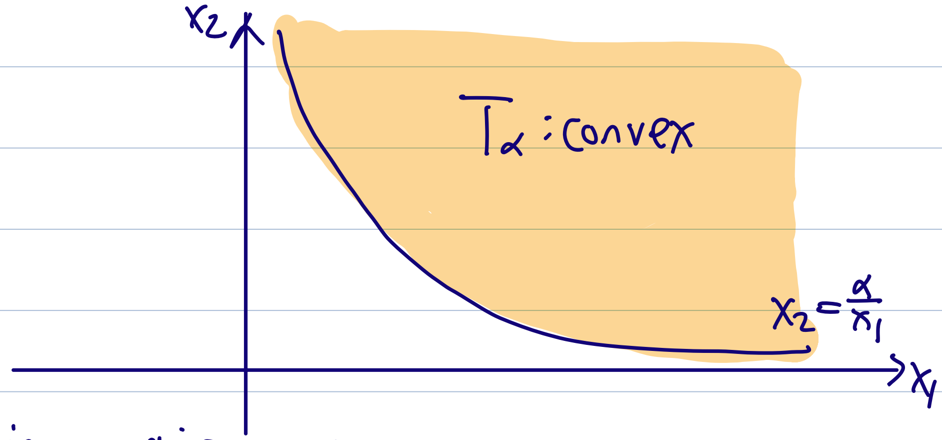
indefinite $\quad$ f is neither convex nor concave.

(neither +ve semi definite nor -ve)

Superlevel sets of f $\quad$ α must be +ve

$T_\alpha = \{(x_1, x_2) : x_1 x_2 \geq \alpha\}$

$x_2 \geq \frac{\alpha}{x_1}$



Hence f is quasiconcave.

Ex: $f(x) = \max\{i \in [n] : x_i \neq 0\}$.
max of index

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

length

$x = (1, 0, 0, 5, 6, 0, 0, 0)$ $\quad$ $n = 8$

$f(x) = 5$

$x = (1, 0, 0, 5, 6, 7, 11, 0)$ $\quad$ non-linear function

$f(x) = 7$

check that sub-level set is convex

sublevel set

Claim: f is quasiconvex.

$(f(x) \leq \alpha) \iff x_i = 0 \ \forall\, i = \lfloor \alpha \rfloor + 1, \dots, n$

$f(x) \leq 5.5 \iff x_i = 0 \ \forall\, i = 6, \dots, 8$ ✓

$f(x) \leq 4.9 \iff x_i = 0 \ \forall\, i = 5, \dots, 8$ ✗

$\Big\}\ f(x) = 5.$

$S_\alpha = \{x : f(x) \leq \alpha\} = \{x : x_i = 0 \ \forall\, i = \lfloor \alpha \rfloor + 1, \dots, n\}$

subspace. $\quad$ convex.

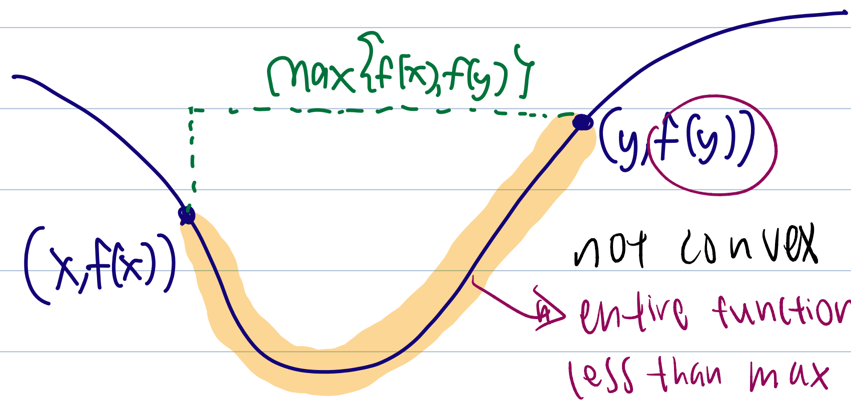
in 3 dimension: 3rd dimension is 0, then you are on x-y plane which is a subspace

Subspaces are convex

$\Rightarrow f$: quasiconvex ////

Thm: f is quasiconvex iff dom f is convex & $\forall x, y \in \text{dom} f$ & $\theta \in [0,1]$

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



not convex

$\Rightarrow$ entire function were is less than max of f values

Thm: TFAE $\quad S_\alpha = \{x : f(x) \leq \alpha\}$

1) S_α is convex $\forall \alpha \in \mathbb{R}$ original definition of q-convex
2) $\forall x, y \in \text{dom} f$ & $\theta \in [0,1]$

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

Pf: $(2 \Rightarrow 1)$

Take $x, y \in S_\alpha$ $\quad$ (WTS: $\theta x + (1-\theta)y \in S_\alpha$)

$\Rightarrow f(x) \leq \alpha$, $f(y) \leq \alpha$

$\Rightarrow \max\{f(x), f(y)\} \leq \alpha$

Using statement 2, we have

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

convex combi belong in S_α

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

$\Rightarrow$ Since $\theta \in [0,1]$ is arbitrary, S_α is convex. ////.

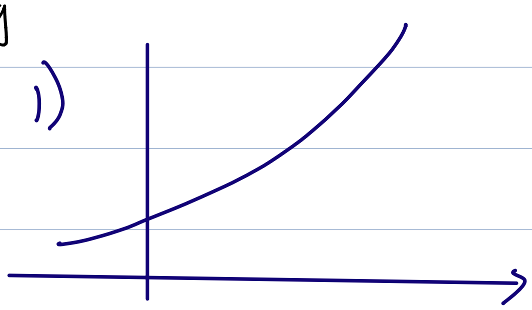
Quasiconvex functions on $\mathbb{R}$.

Fact: A continuous function $f: \mathbb{R} \to \mathbb{R}$ is quasiconvex iff at least one of the following is true

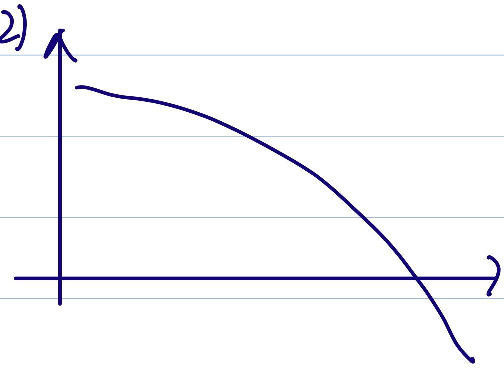
1) it is non-decreasing $\neq$ increasing } monotone
2) it is non-increasing
3) $\exists\, c \in \mathrm{dom} f$ s.t. $\forall t \leq c$ f is non-increasing & $\forall t \geq c$, f is non-decreasing

midpoint/turning point

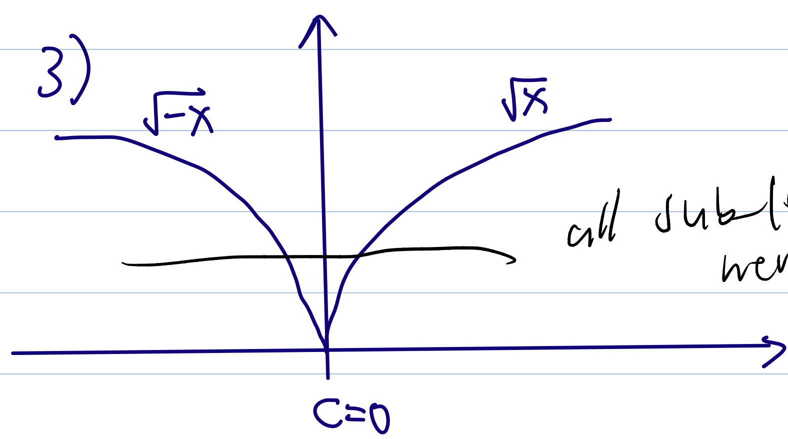
1)



2)



3)



all sublevel sets were are convex

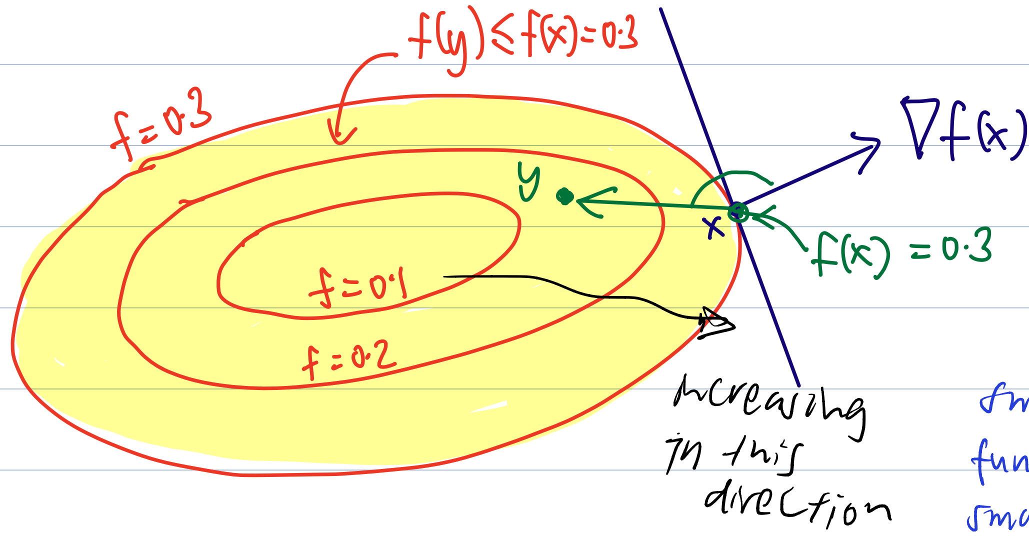
Differentiable Q-convex f : diff^{ble}

$f: \mathbb{R}^n \to \mathbb{R}$ quasiconvex iff dom f is convex &
$\forall x, y \in \text{dom} f$

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

taylor approx

cf: convex $\forall x, y \in \text{dom} f$ $f(y) \ge f(x) + \nabla f(x)^T (y-x)$



makes obtuse angle with $y-x$

such that function has smaller values than that of a particular point x

Second-order condition

$f: \mathbb{R}^n \to \mathbb{R}$ quasiconvex, twice diff^{ble}

$\forall y \in \mathbb{R}^n$ $\nabla f(x)^T y = 0 \Rightarrow y^T \nabla^2 f(x) y \ge 0.$

only this if convex case

↳ y belongs to null space of gradient

Conic combi of convex functions is convex

Properties: 1) If $f_1, \dots, f_n$ quasiconvex, $w_i \ge 0$

$$f = \max\{w_1 f_1, \dots, w_n f_n\} \text{ quasiconvex.}$$

2) Partial Minization

$f(x,y)$: quasiconvex in (x,y) C is a convex set.

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

create new q convex function from old ones