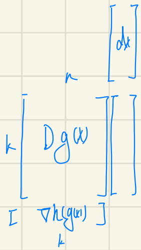
Lec 4: Convex functions

- vector composition

$h: \mathbb{R}^k \to \mathbb{R} \quad g_i: \mathbb{R}^n \to \mathbb{R}$

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



$\nabla f = \nabla h(g(x))^T \cdot Dg(x)$

$$f''(x) = g'(x)^T \nabla^2 h(g(x)) \cdot g'(x) + \underbrace{\nabla h(g(x))^T \cdot g''(x)}_{\sum_i \frac{\partial h}{\partial g_i} \cdot g_i''(x)}$$

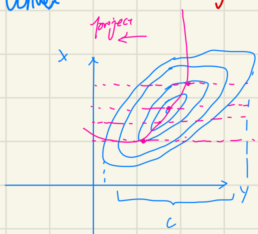
f is convex if
- g_i is convex, h is convex, $\tilde{h}$ non decreasing in each argument
- g_i is concave, h is convex, $\tilde{h}$ non increasing in each argument

- minimization

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

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

is convex



eg. Schur complement

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

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

minimizing over y gives $g(x) = \inf_y f(x,y)$

$\nabla f_y(x,y) = 2B^T x + 2Cy := 0 \quad y = -C^{-1}B^T x$

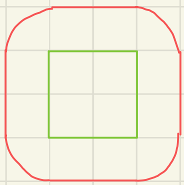
$$g(x) = x^T A x - 2 x^T B C^{-1} B^T x + x^T B C^{-T} C C^{-1} B^T x$$
$$= x^T (A - x^T B C^{-1} B^T) x$$

g is convex in x ; hence schur complement $A - BC^{-1}B^T \succeq 0$

eg. distance to a convex set

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

$\|x-y\|_2 = \| [1,-1] \cdot \begin{bmatrix} x \\ y \end{bmatrix} \|_2$ is jointly cvx in x,y



the level curves smooth out

- perspective function

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

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

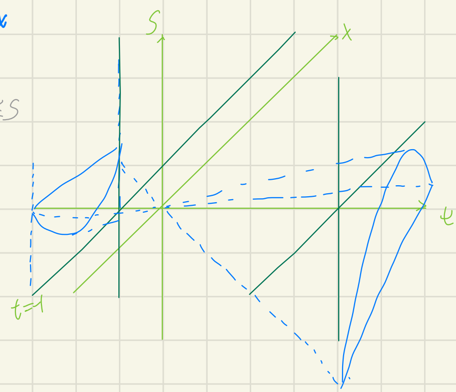
g is jointly cvx in (x,t) if f is cvx

$(x,t,s) \in \text{epi } g \iff g(x,t) = t \cdot f(x/t) \le s$

$\iff f(x/t) \le s/t$

$\iff (x/t, s/t) \in \text{epi } f$

$(x/t, s/t)$ is a cvx set



eg1 $f(x) = x^T x$ is cvx

$t \cdot f(x/t) = x^T x / t$ is cvx for $t > 0$

eg2. negative logarithm

$f(x) = -\log(x)$ is cvx

relative entropy $g(x,t) = t \cdot \log t - t \cdot \log(x)$ is cvx on $\mathbb{R}^2_{++}$

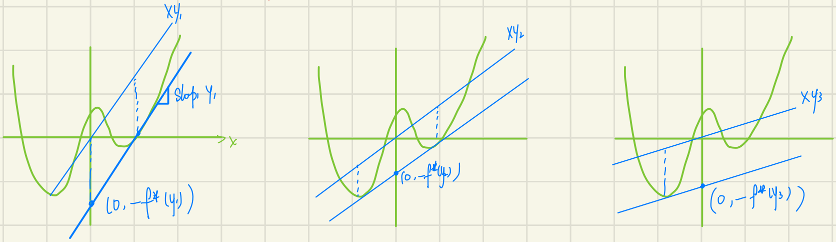
eg3. if f is cvx, then

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

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

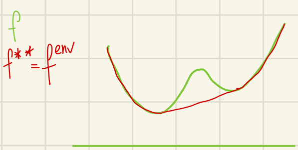
· Conjugate function

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



$f^*(y) = \sup_{x \in \operatorname{dom} f} (y^T x - f(x))$ is supremum over affine functions of y

$f^*(y)$ is convex



$\operatorname{epi}(f^{env}) = \operatorname{convexHull}(\operatorname{epi} f)$

eg. negative logarithm

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

$$f^*(y) = \sup_{x>0} (xy + \log(x)) = \begin{cases} +\infty & y \geq 0 \\ -1 + \log(-1/y) & y < 0 \end{cases}$$

when $y<0$; $xy + \log x$ is a cvx in x

$$\nabla_x (xy + \log x) = y + 1/x := 0 \quad x = 1/y$$

eg strictly convex quadratic

$$f(x) = \tfrac{1}{2} x^T Q x \qquad Q \in S^n_{++}$$

$$f^*(y) = \sup_x \left(y^T x - \tfrac{1}{2} x^T Q x \right)$$

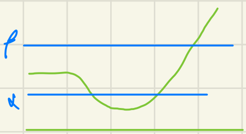
$$= \tfrac{1}{2} y^T Q^{-1} y$$

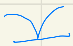
- Quasiconvex function

  $f: \mathbb{R}^n \to \mathbb{R}$ is quasiconvex if dom f is convex and the sublevel set

  $$S_\alpha = \{ x \in \operatorname{dom} f \mid f(x) \le \alpha \}$$

  is cvx for all α



eg $\sqrt{|x|}$ is qcvx on $\mathbb{R}$ 

eg. $\operatorname{ceil}(x) = \inf \{ z \in \mathbb{Z} \mid z \ge x \}$ (quasilinear) (qcvx and qcnv)

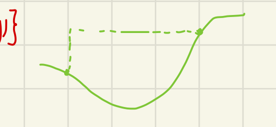
eg $f(x_1, x_2) = x_1 x_2 = x^T Q x \qquad Q = \begin{bmatrix} 0 & \tfrac{1}{2} \\ \tfrac{1}{2} & 0 \end{bmatrix}$ is quasiconcave

eg. linear fractional $f(x) = \dfrac{a^T x + b}{c^T x + d}$ dom $f = \{ x \mid c^T x + d > 0 \}$ is qcvx

eg. distance ratio $f(x) = \dfrac{\|x - a\|_2}{\|x - b\|_2}$ dom $f = \{ x \mid \|x - a\|_2 \le \|x - b\|_2 \}$ is qcvx

- Modified Jensen's Inequality for quasiconvex

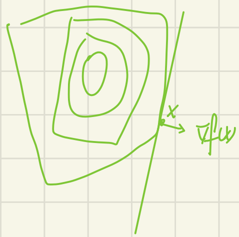
  $$f(\theta x + (1-\theta) y) \le \max\{ f(x), f(y) \}$$



- First order condition for qcvx

differentiable f with cvx domain is qcvx iff

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



- Log-convex and log-concave

powers: x^a is on $\mathbb{R}_{++}$ is $\begin{cases} \text{log cvx for } a \le 0 \\ \text{log cnv for } a \ge 0 \end{cases}$

many probability distributions are log-concave

eg. normal $\quad p(x; u, \Sigma) = \frac{1}{\sqrt{(2\pi)^n \cdot |\Sigma|}} \cdot \exp\left[-\frac{1}{2}(x-u)^T \cdot \Sigma^{-1} \cdot (x-u)\right]$

cumulative Gaussian distribution function Φ is log-cnv.

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

- log-concave function

twice-differentiable f with cvx domain is log-concave iff

$\nabla^2 f(x) \preceq \frac{\nabla f(x) \cdot \nabla f(x)^T}{f(x)}$

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

for concave: Hessian is neg semi def

for log-cnv: Hessian is allowed 1 pos eigen value

production: product of 2 log-cnv is log-cnv integration: if $f: \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ is log-cnv, then

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

is log-cnv (not easy to show

convolution preserves log-concavity

$$(f*g)(x) = \int f(x-y)\, g(y)\, dy$$ is log-concave if f and g are concave

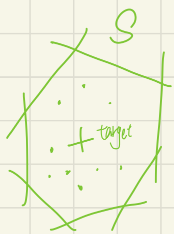
if $C \subseteq \mathbb{R}^n$ is cvx and y is a random var with log-concave pdf.

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

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

eg. yield function

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



adjust manufacturing to hit the target

S is acceptable manufacturing

bias $w \sim p(w)$ is log-cnv

S is cvx

- Y is log-cnv (can maximize it)
- yield regions $\{x \mid Y(x) \geq \alpha\}$ are cvx