Homework 1

Wednesday, November 27, 2019 2:27 PM

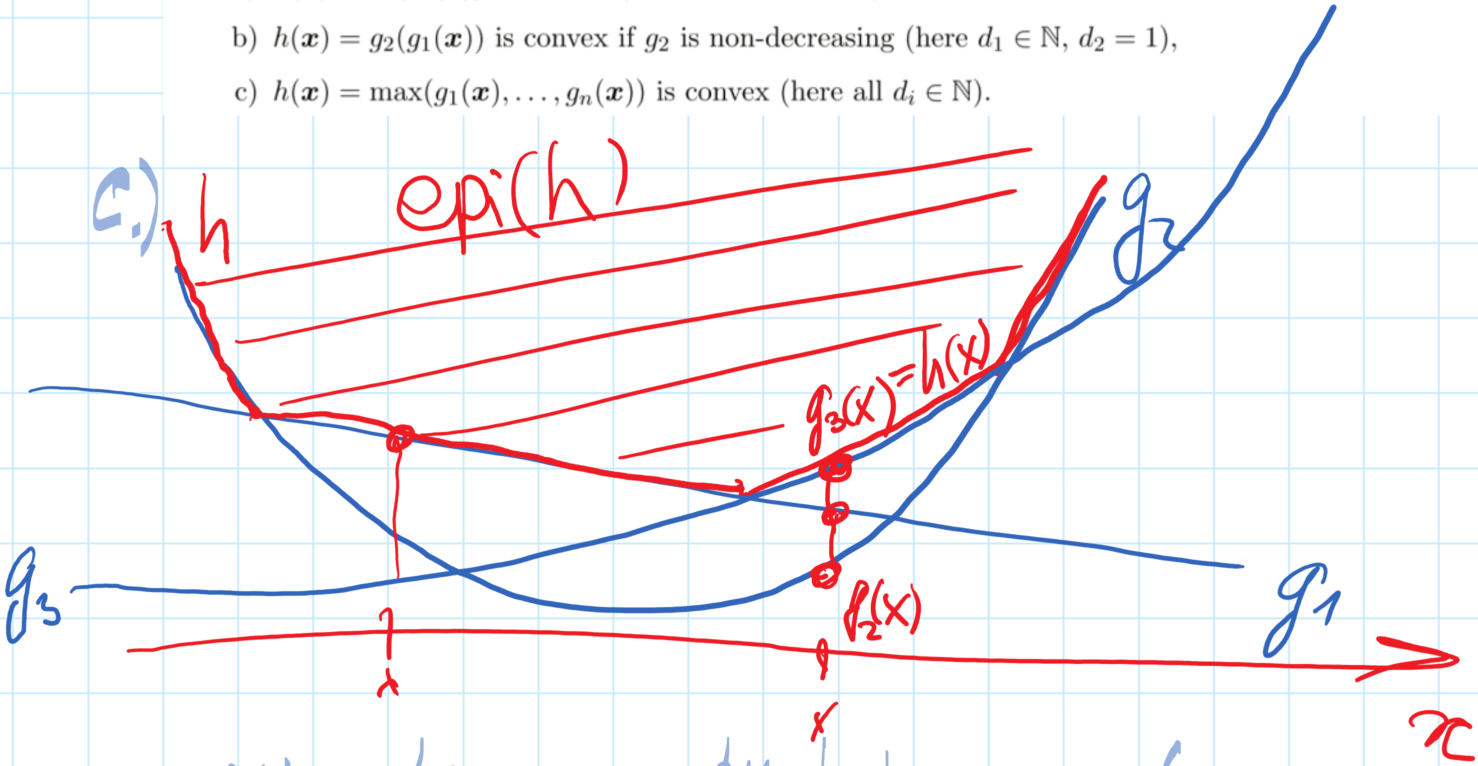
Problem 1: Given n convex functions $g_i : \mathbb{R}^{d_i} \to \mathbb{R}$ for $i \in \{1, \ldots, n\}$, prove or disprove that the function

a) $h(\boldsymbol{x}) = g_2(g_1(\boldsymbol{x}))$ is convex (here $d_1 \in \mathbb{N}$, $d_2 = 1$),

b) $h(\boldsymbol{x}) = g_2(g_1(\boldsymbol{x}))$ is convex if g_2 is non-decreasing (here $d_1 \in \mathbb{N}$, $d_2 = 1$),

c) $h(\boldsymbol{x}) = \max(g_1(\boldsymbol{x}), \ldots, g_n(\boldsymbol{x}))$ is convex (here all $d_i \in \mathbb{N}$).

c.)



$$\mathrm{epi}(h) = \{(x,y) \in \mathbb{R}^{d+1} \mid h(x) \le y\} =$$

$$= \{(x,y) \in \mathbb{R}^{d+1} \mid \max(g_1(x), .., g_n(x)) \le y\}$$

$$= \{(x,y) \in \mathbb{R}^{d+1} \mid g_i(x) \le y \text{ for } i=1..n\}$$

$$= \bigcap_{i=1}^{n} \underbrace{\{(x,y) \in \mathbb{R}^{d+1} \mid g_i(x) \le y\}}$$

g_i is convex $\Rightarrow$ $\mathrm{epi}(g_i)$ is convex

$\Rightarrow \underbrace{\bigcap_{i=1}^{n} \mathrm{epi}(g_i)}_{=\mathrm{epi}(h)}$ is convex

$\Rightarrow h$ is convex

Problem 4: Let $f : \mathbb{R}^2 \to \mathbb{R}$ be the following convex function:

$$f(x_1, x_2) = e^{x_1+x_2} - 5 \cdot \log(x_2)$$

a) Consider the following shaded region S in $\mathbb{R}^2$. Is this region convex? Why?

b) Find the <u>maximizer</u> $\boldsymbol{x}^*$ of f over the shaded region S. Justify your answer.

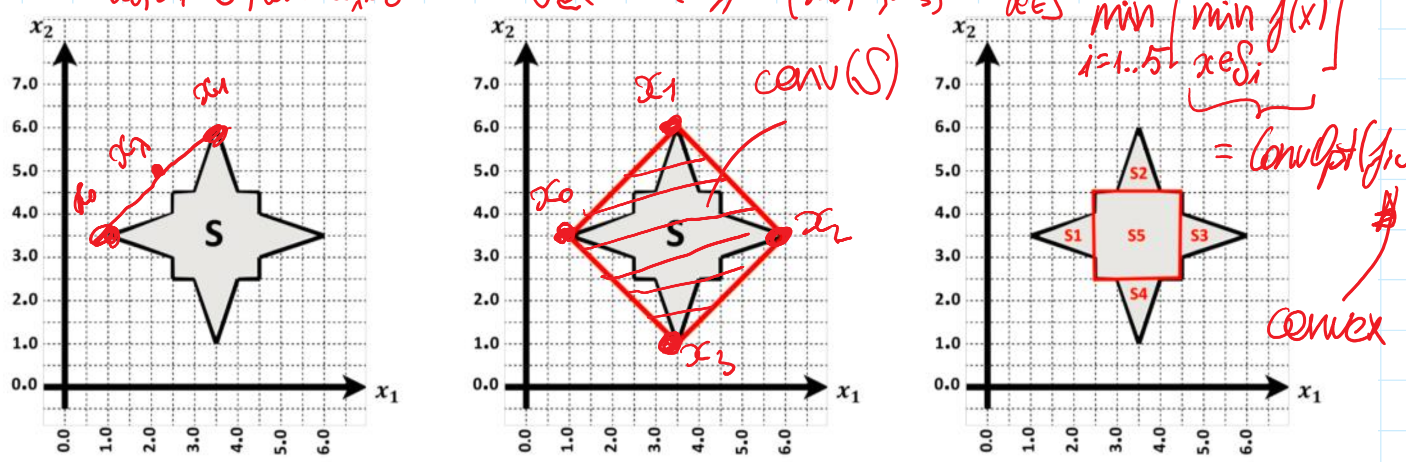
c) Assume that we are given an algorithm `ConvOpt`(f, D) that takes as input a convex function f and convex region D, and returns the <u>minimum</u> of f over D. Using the `ConvOpt` algorithm, how would you find the global <u>minimum</u> of f over the shaded region S?

$x_0, x_1 \in S$, but $x_\lambda \notin S$

$V_e(\mathrm{conv}(S)) = \{x_0, .., x_3\}$

$\min_{x \in S} f(x) = \min_{i=1..5} \big[\underbrace{\min_{x \in S_i} f(x)}_{= \mathrm{ConvOpt}(f, S_i)}\big]$

convex



(a) initial non-convex set (b) convex hull (c) union of convex sets

In-Class 6

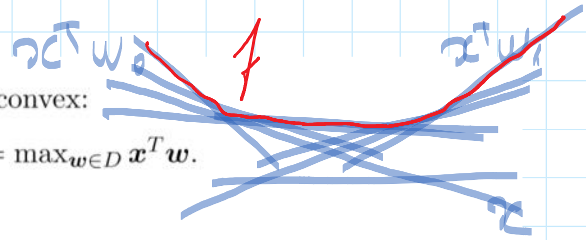
Wednesday, November 27, 2019 2:52 PM

Problem 6: Prove that the following functions $f : \mathbb{R}^d \to \mathbb{R}$ are convex:

a) $D \subset \mathbb{R}^d$ bounded and closed set and f is defined by $f(\boldsymbol{x}) = \max_{\boldsymbol{w}\in D} \boldsymbol{x}^T\boldsymbol{w}$.

b) f is the objective function of logistic regression, that is

$$f(\boldsymbol{w}) = -\ln p(\boldsymbol{y} \mid \boldsymbol{w}, \boldsymbol{X}) = -\sum_{i=1}^{N}\left(y_i \ln \sigma(\boldsymbol{w}^T\boldsymbol{x}_i) + (1-y_i)\ln(1-\sigma(\boldsymbol{w}^T\boldsymbol{x}_i))\right).$$



a) Consider $x_0, x_1 \in \mathbb{R}^d$, $\lambda \in (0,1)$, $x_\lambda = (1-\lambda)x_0 + \lambda x_1$

$$f(x_\lambda) = \max_{w\in D} x_\lambda^T w = \max_{w\in D}\left((1-\lambda)x_0^T w + \lambda x_1^T w\right) =$$

$$= (1-\lambda)x_0^T w^* + \lambda x_1^T w^* \le \max_{w\in D}(1-\lambda)x_0^T w + \max_{w\in D}\lambda x_1^T w$$

w^* = maximizer

$$= (1-\lambda) f(x_0) + \lambda f(x_1)$$

$\le$

$\Rightarrow f$ is convex

Problem 8: Show that a twice differentiable function $f : \cancel{D}\ \mathbb{R}^d \to \mathbb{R}$ with ~~a convex domain $D \subset \mathbb{R}^d$~~ is convex if and only if its Hessian is positive semi-definite on the whole ~~domain D~~ $\mathbb{R}^d$.

For all $x_0, x_1 \in \mathbb{R}^d$ $\underbrace{\text{and some } \lambda \in (0,1)}$ → depends on x_0, x_1

$$(*)\quad f(x_1) = f(x_0) + \nabla f(x_0)^T (x_1 - x_0) + \frac{1}{2}(x_1 - x_0)^T \nabla^2 f(x_\lambda)(x_1 - x_0)$$

„$\Leftarrow$" $\underbrace{\text{from Hessian is PSD}}$ to f is convex

$$\Rightarrow (x_1 - x_0)^T \nabla^2 f(x_\lambda)(x_1 - x_0) \geq 0$$

$$\Rightarrow f(x_1) \geq f(x_0) + \nabla f(x_0)^T (x_1 - x_0)$$

→ 1-st order convexity condition ✓

„$\Rightarrow$" Assume for some $x \in \mathbb{R}^d$ $\nabla^2 f(x)$ is not PSD, that means for some $d \in \mathbb{R}^d$, s.t.

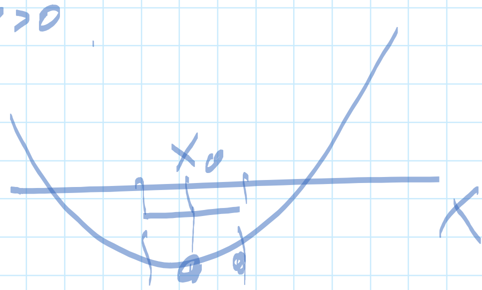
$$d^T \nabla^2 f(x) d < 0$$

Consider $(*)$ with $x_0 = x$, $x_1 = x + \varepsilon d$ for $\varepsilon > 0$.

For small ε we will have

$$d^T \nabla^2 f(x_\lambda) d < 0,$$

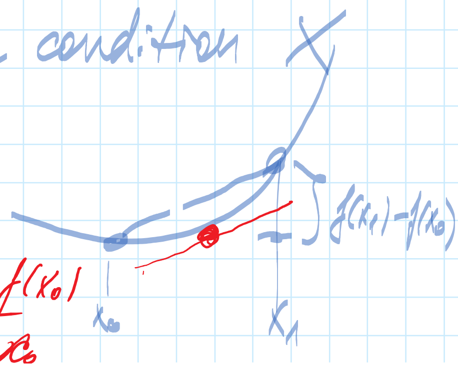
because $\nabla^2 f$ is continuous.



$$(*) \Rightarrow f(x_1) = f(x_0) + \nabla f(x_0)^T (x_1 - x_0) + \frac{1}{2}\varepsilon^2 \underbrace{d^T \nabla f(x_\lambda)^T d}_{<0}$$

$$\Rightarrow f(x_1) < f(x_0) + \nabla f(x_0)^T (x_1 - x_0)$$

→ 1st order convexity condition X



$$\exists x_\lambda : f'(x_\lambda) = \frac{f(x_1) - f(x_0)}{x_1 - x_0}$$

$x_1 - x_6$