VE485, Optimization in Machine Learning (Summer 2020) Homework: 2. Convex Function

2. Convex Function

Lecturer: Xiaolin Huang `xiaolinhuang@sjtu.edu.cn`

Student: Chongdan Pan `pandddddda@sjtu.edu.cn`

Problem 1

Definition of convexity. Suppose $f : \mathbb{R} \to \mathbb{R}$ is convex, and $a,\ b \in \textbf{dom} f$ with $a < b$.

1. Show that
$$f(x) \leqslant \frac{b-x}{b-a} f(a) + \frac{x-a}{b-a} f(b)$$
for all $x \in [a, b]$.
2. Show that
$$\frac{f(x)-f(a)}{x-a} \leqslant \frac{f(b)-f(a)}{b-a} \leqslant \frac{f(b)-f(x)}{b-x}$$
for all $x \in (a, b)$. Draw a sketch that illustrates this inequality.
3. Suppose f is differentiable. Use the result in 2 to show that
$$f'(a) \leqslant \frac{f(b)-f(a)}{b-a} \leqslant f'(b).$$
Note that these inequalities also follow from (3.2) in the textbook :
$$f(b) \geqslant f(a) + f'(a)(b-a), \qquad f(a) \geqslant f(b) + f'(b)(a-b).$$
4. Suppose f is twice differentiable. Use the result in 3 to show that $f''(a) \geqslant 0$ and $f''(b) \geqslant 0$.

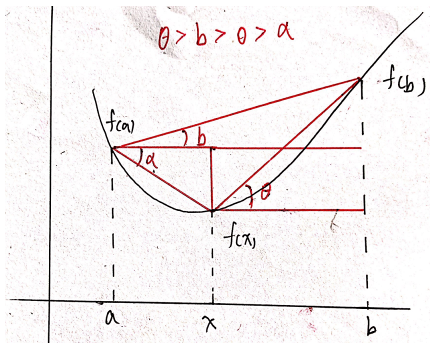
Answer

1. **(a) Since** $f(x)$ **is convex, then** $f(\theta a + (1-\theta)b) \leq \theta f(a) + (1-\theta) f(b), \forall \theta$
Let $x = \theta a + (1-\theta)b$**, then** $\theta = \frac{b-x}{b-a}, 1-\theta = \frac{x-a}{b-a}$
Then $f(x) \leq \frac{b-x}{b-a} f(a) + \frac{x-a}{b-a} f(b)$
2. **(b)** $f(a) \geq f(x) + \nabla f(x)^{\mathrm{T}}(a-x)$
$\nabla f(a)^{\mathrm{T}} \leq \frac{f(x)-f(a)}{x-a} \leq \nabla f(x)^{\mathrm{T}}$
Similarly, $\nabla f(x)^{\mathrm{T}} \leq \frac{f(b)-f(x)}{b-x} \leq \nabla f(b)^{\mathrm{T}}$
Then $\frac{f(x)-f(a)}{x-a} \leq \frac{f(b)-f(x)}{b-x}$

$xf(a) + bf(x) - bf(a) \leq xf(b) + af(x) - af(b)$
Plus $bf(b)$ **on both sides, we get** $bf(b) + xf(a) - bf(a) - xf(b) \leq bf(b) + af(x) - af(b) - bf(x)$
$(b-x)(f(b) - f(a) \leq (b-a)(f(b) - f(x))$
Then $\frac{f(b)-f(a)}{b-a} \leq \frac{f(b)-f(x)}{b-x}$

Similarly, we can plus $af(a)$ **on both sides, we get** $af(a) + bf(x) - bf(a) - af(x) \leq af(a) + xf(b) - af(b) - xf(a)$
$(a-b)(f(a) - f(x)) \leq (a-x)(f(a) - f(b))$
Then $\frac{f(x)-f(a)}{x-a} \leq \frac{f(b)-f(a)}{b-a}$

Therefore $\frac{f(x)-f(a)}{x-a} \leq \frac{f(b)-f(a)}{b-a} \leq \frac{f(b)-f(x)}{b-x}$



3. **(c) In (b), we have shown that** $f(a)' \leq \frac{f(x)-f(a)}{x-a}$ **and** $\frac{f(b)-f(x)}{b-x} \leq f(b)'$
   Therefore, $f(a)' \leq \frac{f(b)-f(a)}{b-a} \leq f(b)'$
4. **(d) According to (c),** $\forall \triangle x > 0, f(a)' \leq f(a + \triangle x)'$
   $f(a)'' = \lim_{\triangle x \to 0} \frac{f(a+\triangle x)' - f(a)'}{\triangle x} \geq 0$
   Therefore $f(a)'' \geq 0, f(b)'' \geq 0$

Problem 2

Composition with an affine function. Show that the following functions $f : \mathbb{R}^n \to \mathbb{R}$ are convex.

1. $f(x) = \|Ax - b\|$, where $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, and $\| \cdot \|$ is a norm on $\mathbb{R}^m$.
2. $f(x) = -\left(\det(A_0 + x_1 A_1 + \cdots + x_n A_n)\right)^{1/m}$, on $\{x \mid A_0 + x_1 A_1 + \cdots + x_n A_n \succ 0\}$, where $A_i \in \mathbf{S}^m$.
3. $f(X) = \mathbf{tr}\left(A_0 + x_1 A_1 + \cdots + x_n A_n\right)^{-1}$, on $\{x \mid A_0 + x_1 A_1 + \cdots + x_n A_n \succ 0\}$, where $A_i \in \mathbf{S}^m$. (Use the fact that $\mathbf{tr}(X^{-1})$ is convex on $\mathbf{S}^m_{++}$; see exercise 3.18 in the text book.)

Answer:

1. **(a) According to definition, the norm must satisfy triangular inequality:**
   $f(a + b) \leq f(a) + f(b)$
   Let $a = \theta x, b = (1-\theta)y$, **we can get** $f(\theta x + (1-\theta)y) \leq f(\theta x) + f((1-\theta)y) = \theta f(x) + (1-\theta)f(y)$
   So norm is a convex function
   It's clear that $g(x) = Ax - b$ **is an affine function.**
   Therefore $f(x)$**, the composition of norm and an affine function, is convex.**

2. **(b) Let $h(X) = -(\det X)^{1/m}$, by restricting it to a line we get $H(t) = h(X + tV)$**
   $H(t) = -(\det(X + tV))^{1/m} = -(\det X)^{1/m}(\det(1 + X^{-1/2}tVX^{-1/2}))^{1/m}$
   $H(t) = -(\det(X + tV))^{1/m} = -(\det X)^{1/m}(\prod_{i=1}^{m}(1 + t\lambda_i))^{1/m}$ **where λ_i is the eigenvalue of $X^{-1/2}VX^{-1/2}$**
   Then, both $H(t)$ and $h(X)$ are convex functions.

   Let $g(x) = A_0 + x_1A_1 + \cdots + x_nA_n$, then $g(x) \preceq 0, g(x) \in \mathbf{dom}f$
   Then $h(X) = -(\det X)^{1/m}$ is convex, and $g(x)$ is an affine transformation
   Therefore $f(x)$, the composition of a convex function and an affine function, is convex.

3. **(c) $\mathbf{tr}(X^{-1}) = \sum_{i=1}^{m}\frac{1}{\lambda_i}$, where λ_i is the eigenvalue of X**
   Then $Tr(X^{-1})$ is convex.
   Let $g(x) = A_0 + x_1A_1 + \cdots + x_nA_n$, then $g(x) \preceq 0, g(x) \in \mathbf{dom}f$
   Since $h(X) = \mathbf{tr}(X) = X^{-1}$ is convex, and $g(x)$ is an affine transformation
   $f(x)$, the composition of a convex function and an affine function, is convex.

Problem 3

Young's inequality. Let $f : \mathbb{R} \to \mathbb{R}$ be an increasing function, with $f(0) = 0$, and let g be its inverse. Define F and G as

$$F(x) = \int_0^x f(a)\,da, \qquad G(y) = \int_0^y g(a)\,da.$$

Show that F and G are conjugates, which then leads to the Young's inequality,

$$xy \leqslant F(x) + G(y).$$

Answer:
Obviously, $F(x), G(y)$ is convex and increasing, with $F(0) = G(0) = 0$.
$F^*(y) = yx - F(x)$ **where** $y = F(x)' = f(x)$
Since $g(y)$ is the inverse function of $f(x)$, then $x = g(y)$
Therefore $F^*(y) = yg(y) - F(x)$
Since $F(x) = yg(y) - G(y)$
Then $F^*(y) = G(y)$, which means F and G are conjugates.
According to Fenchel inequality, $F(x) + G(y) \geq xy$