Convex Optimization - AI2101
Assignment - III

Bhuvan Chandra K
AI23BTECH11013

March 26, 2025

Question 3.1

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

(a) Show that

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

for all $x \in [a, b]$.

(b) Show that

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

for all $x \in (a, b)$. Draw a sketch that illustrates this inequality.

(c) Suppose f is differentiable. Use the result in (b) to show that

$$f'(a) \leq \frac{f(b)-f(a)}{b-a} \leq f'(b).$$

Note that these inequalities also follow from:

$$f(b) \geq f(a) + f'(a)(b-a), \quad f(a) \geq f(b) + f'(b)(a-b).$$

(d) Suppose f is twice differentiable. Use the result in (c) to show that $f''(a) \geq 0$ and $f''(b) \geq 0$.

Solution:

(a) Let us consider

$$\frac{b-x}{b-a} = \lambda$$

then from this

$$1 - \lambda = \frac{x-a}{b-a}$$

$$(\lambda a + (1-\lambda)b) = \frac{ab-ax}{b-a} + \frac{bx-ab}{b-a} = x$$

We can rewrite the LHS as

$$f(\lambda a + (1-\lambda)b)$$

RHS as

$$\lambda f(a) + (1-\lambda) f(b)$$

Since $0 \leq \lambda \leq 1$, $a, b \in$ dom f
Using the convex function property on the above constraints,

$$f(\lambda a + (1-\lambda)b) \leq \lambda f(a) + (1-\lambda) f(b)$$

We have LHS $\leq$ RHS

$$\therefore f(x) \leq \left(\frac{b-x}{b-a}\right) f(a) + \left(\frac{x-a}{b-a}\right) f(b)$$

(b) From the inequality in part (a)

$$f(x) \leq \left(\frac{b-x}{b-a}\right) f(a) + \left(\frac{x-a}{b-a}\right) f(b)$$

Subtracting $f(a)$ on both sides of inequality from (a),

$$f(x) - f(a) \leq \left(\frac{x-a}{b-a}\right) f(b) + f(a) \left(\frac{b-x}{b-a} - 1\right)$$

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

$$\frac{f(x) - f(a)}{x-a} \leq \frac{f(b) - f(a)}{b-a} \quad \rightarrow (1)$$

Subtracting $f(b)$ on both sides of inequality from (a),

$$f(x) - f(b) \leq \left(\frac{b-x}{b-a}\right) f(a) + f(b) \left(\frac{x-a}{b-a} - 1\right)$$

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

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

Cancelling out -1 on both sides,

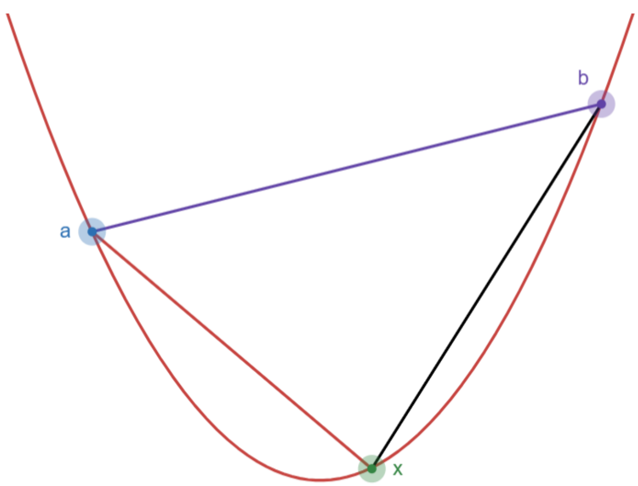
$$\left(\frac{b-x}{b-a}\right) (f(b) - f(a)) \leq f(b) - f(x)$$

$$\frac{f(b) - f(a)}{b-a} \leq \frac{f(b) - f(x)}{b-x} \quad \rightarrow (2)$$

From (1) and (2) we get:

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

This inequality implies that the slope of the line joining $(a, f(a))$ to $(b, f(b))$ will be greater than that of the line joining $(a, f(a))$ to $(x, f(x))$ and lesser than that of the line joining $(x, f(x))$ to $(b, f(b))$.



(c) From the 1st inequality in (b),

$$\frac{f(x)-f(a)}{x-a} \leq \frac{f(b)-f(a)}{b-a}, \quad x \in (a,b)$$

As x approaches a^+, the LHS closes in on $f'(a)$. i.e.,

$$\lim_{x\to a^+} \frac{f(x)-f(a)}{x-a} = f'(a)$$

RHS remains the same,

$$f'(a) \leq \frac{f(b)-f(a)}{b-a} \quad \rightarrow (1)$$

From the second inequality in (b),

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

As x approaches b^-, the RHS closes in on $f'(b)$. i.e.,

$$\lim_{x\to b^-} \frac{f(b)-f(x)}{b-x} = f'(b)$$

LHS remains the same,

$$\frac{f(b)-f(a)}{b-a} \leq f'(b) \quad \rightarrow (2)$$

From (1) and (2), we get:

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

(d) From (c), we have

$$f'(b) \geq f'(a)$$

So,

$$\frac{f'(b)-f'(a)}{b-a} \geq 0$$

As $b \to a^+$,

$$\lim_{b\to a^+} \frac{f(b)-f(a)}{b-a} = f'(a) \geq 0$$

Similarly,

$$\frac{f(b)-f(a)}{b-a} \to 0$$

As $a \to b^-$,

$$\lim_{a\to b^-} \frac{f(b)-f(a)}{b-a} = f'(b) \geq 0$$

Thus, $f'(a) \geq 0$ and $f'(b) \geq 0$.

Question 3.4

Show that a continuous function $f : \mathbb{R}^n \to \mathbb{R}$ is convex if and only if for every line segment, its average value on the segment is less than or equal to the average of its values at the endpoints of the segment: For every $x, y \in \mathbb{R}^n$,

$$\int_0^1 f(x+\lambda(y-x))d\lambda \leq \frac{f(x)+f(y)}{2}.$$

Solution:

We need to prove

$$f \text{ is convex } \iff \forall x, y \in \mathbb{R}, \int_0^1 f(x + \lambda(y - x))d\lambda \leq \frac{f(x) + f(y)}{2}$$
$$\text{Statement I} \iff \text{Statement II}$$

Since f is convex,From Jensen's inequality we get

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

On this inequality, we do integration on both sides with respect to λ on limits $[0, 1] \rightarrow 0 \leq \lambda \leq 1$

$$\int_0^1 f(x + \lambda(y - x))d\lambda \leq \int_0^1 (\lambda f(y) + (1 - \lambda)f(x))d\lambda$$
$$\int_0^1 \lambda f(y)d\lambda + \int_0^1 (1 - \lambda)f(x)d\lambda = f(y)\left[\frac{1}{2} - 0\right] + f(x)\left[1 - \frac{1}{2} - 0\right]$$
$$= \frac{f(x) + f(y)}{2}$$

Hence,

$$\int_0^1 f(x + \lambda(y - x))d\lambda \leq \frac{f(x) + f(y)}{2}$$

Thus, **Statement I** $\Rightarrow$ **Statement II**.
(ii) **Statement II** $\Rightarrow$ **Statement I**
We have, for every $x, y \in \mathbb{R}^n$,

$$\int f(x + \lambda(y - x))d\lambda \leq \frac{f(x) + f(y)}{2}$$

If the function is not convex, then there exist $x, y \in \mathbb{R}^n$ and $\lambda_0 \in (0, 1)$ such that

$$f(y + \lambda_0(x - y)) > \lambda_0 f(x) + (1 - \lambda_0)f(y).$$

Let $\beta_0 = \lambda_0 x + (1 - \lambda_0)y$. Then,

$$f(\beta_0) > \lambda_0 f(x) + (1 - \lambda_0)f(y).$$

Since f is continuous and $f(\beta)$ is above the line joining x and y, there exist a, b on the chord x, y such that

$$f(\beta) \geq \lambda f(x) + (1 - \lambda)f(y), \quad \forall \lambda \in (0, 1), \forall \beta \in [a, b].$$

In this case, the line joining x and y intersects the curve at points a, b. Thus, where $\beta = \lambda a + (1 - \lambda)b$,

$$f(\lambda a + (1 - \lambda)b) \geq \lambda f(a) + (1 - \lambda)f(b), \quad \forall \lambda \in (0, 1).$$
$$\int_0^1 f(\lambda a + (1 - \lambda)b)d\lambda \geq \frac{f(a) + f(b)}{2}, \quad \text{(contradiction)}$$

This contradicts the original assumption.

$$\text{State}\overset{\neg}{\text{m}}\text{ent I} \quad \Rightarrow \quad \text{State}\overset{\neg}{\text{m}}\text{ent II}$$
$$\text{Statement II} \quad \Rightarrow \quad \text{Statement I}$$

Thus,

$$\text{Statement I} \iff \text{Statement II.}$$
$$\therefore f \text{ is convex } \iff \forall x, y \in \mathbb{R}, \int_0^1 f(x + \lambda(y - x))d\lambda \leq \frac{f(x) + f(y)}{2}$$

Question 3.7

Suppose $f : \mathbb{R}^n \to \mathbb{R}$ is convex with $\operatorname{dom} f = \mathbb{R}^n$, and bounded above on $\mathbb{R}^n$. Show that f is constant.

Solution:

We prove this using a contradiction.
Let f be not a constant, so there exist x, y with $f(x) < f(y)$. Let us define

$$F(t) = f(x + t(y - x)), \quad (F : \mathbb{R} \to \mathbb{R})$$

Since f is convex,
then $F(t) = f(x + t(y - x))$ is also convex if $\operatorname{dom} = \mathbb{R}^n$, $x, y \in \operatorname{dom}$.
Because,

$$F(\theta t_1 + (1-\theta)t_2) = f(\theta(x + t_1(y - x)) + (1-\theta)(x + t_2(y - x)))$$

$$F(\theta t_1 + (1-\theta)t_2) \leq \theta f(x + t_1(y - x)) + (1-\theta) f(x + t_2(y - x))$$

$$F(\theta t_1 + (1-\theta)t_2) \leq \theta F(t_1) + (1-\theta) F(t_2)$$

Thus, F is convex. Here,

$$F(0) = f(x), \quad F(1) = f(y)$$

So, $F(0) < F(1)$. Since F is convex,

$$F\left(\frac{t-1}{t} \cdot 0 + \frac{1}{t} \cdot t\right) \leq \frac{t-1}{t} F(0) + \frac{1}{t} F(t)$$

$$F(1) \leq \frac{t-1}{t} F(0) + \frac{1}{t} F(t)$$

$$F(t) \geq t(F(1) - F(0)) + F(0)$$

If f is bounded, F should also be bounded.
But as $t \to \infty$, $F(t)$ gets unbounded (approaches infinity).
This is a contradiction to the assumption. Hence, the function f is a constant.

Question 3.11

Monotone mappings. A function $\psi : \mathbb{R}^n \to \mathbb{R}^n$ is called monotone if for all $x, y \in \operatorname{dom} \psi$,

$$(\psi(x) - \psi(y))^T (x - y) \geq 0.$$

(Note that 'monotone' as defined here is not the same as the definition given in §3.6.1. Both definitions are widely used.) Suppose $f : \mathbb{R}^n \to \mathbb{R}$ is a differentiable convex function. Show that its gradient ∇f is monotone. Is the converse true, i.e., is every monotone mapping the gradient of a convex function?

Solution:

Given $\psi : \mathbb{R}^n \to \mathbb{R}^n$ is a monotone function, it satisfies

$$(\psi(x) - \psi(y))^T (x - y) \geq 0, \quad \forall x, y \in \operatorname{dom}$$

Since f is convex, we have

$$f(x) \geq f(y) + \nabla f(y)^T(x - y) \quad (1)$$

By swapping x, y, we also get

$$f(y) \geq f(x) + \nabla f(x)^T(y - x) \quad (2)$$

Adding (1) and (2),

$$0 \geq \nabla f(y)^T(x - y) + \nabla f(x)^T(y - x)$$

Thus,

$$[\nabla f(x) - \nabla f(y)]^T(x - y) \geq 0$$

This shows that ∇f is monotone.
Let us consider $\psi(x, y) = (-y, x)$. Let D = $(\psi(x, y) - \psi(u, v))^T(x, y) - (u, v)$

$$D = \psi(x, y)^T(x, y) - \psi(x, y)^T(u, v) - \psi(u, v)^T(x, y) + \psi(u, v)^T(u, v)$$

$$D = (-xy + ym) - (-uy + xv)(-vn + uy) + (-vu + uv) = 0$$

Since it is non-negative, $\psi(m, y) = (y, v)$ is monotone.
Here, $\psi = \nabla f$.

$$\frac{\partial f}{\partial x} = -y, \quad \frac{\partial f}{\partial y} = x$$

$$\frac{\partial^2 f}{\partial y \partial x} = 1, \quad \frac{\partial^2 f}{\partial x \partial y} = 1$$

Since ,

$$\frac{\partial^2 f}{\partial y \partial x} \neq \frac{\partial^2 f}{\partial x \partial y}$$

there doesn't exist a convex function f such that

$$\nabla f = \psi(m, y) = (-y, x)$$

Thus, the converse is false.

Question 3.13

Kullback-Leibler divergence and the information inequality. Let D_{KL} be the Kullback-Leibler divergence, as defined in (3.17). Prove the information inequality:

$$D_{KL}(u, v) \geq 0 \quad \text{for all } u, v \in \mathbb{R}^n_{++}.$$

Also show that $D_{KL}(u, v) = 0$ if and only if $u = v$.
Hint. The Kullback-Leibler divergence can be expressed as

$$D_{KL}(u, v) = f(u) - f(v) - \nabla f(v)^T(u - v),$$

where

$$f(v) = \sum_{i=1}^{n} v_i \log v_i$$

is the negative entropy of v.

Solution:

The function negative entropy of v is defined as:

$$f(v) = \sum_{i=1}^{n} v_i \log v_i, \quad v \in \mathbb{R}^n_{++}$$

The Hessian matrix of $f(v)$ is:

$$\frac{\partial^2 f}{\partial v_i \partial v_j} = \begin{cases} \frac{1}{v_i}, & i = j \\ 0, & i \neq j \end{cases}$$

The Hessian matrix is diagonal with elements:

$$H(f(v)) = \text{diag}\left(\frac{1}{v_1}, \ldots, \frac{1}{v_n}\right)$$

Since each v_i is positive, the Hessian is positive definite.
Since the Hessian is positive definite, $f(v)$ is convex.
So, from the convexity condition, we get:

$$f(u) \geq f(v) + \nabla f(v)^T (u - v), \quad \forall u, v \in \mathbb{R}^n_{++}$$

Substituting f in the expression:

$$\sum_{i=1}^{n} u_i \log u_i \geq \sum_{i=1}^{n} v_i \log v_i + \sum_{i=1}^{n} (1 + \log v_i)(u_i - v_i)$$

i.e.,

$$\sum_{i=1}^{n} u_i (\log u_i - \log v_i) + \sum_{i=1}^{n} (u_i - v_i) \geq 0$$

$$\sum_{i=1}^{n} u_i \log\left(\frac{u_i}{v_i}\right) + \sum_{i=1}^{n} (u_i - v_i) \geq 0$$

For $f(v)$, we get:

$$D_{KL}(u, v) = f(u) - f(v) - \nabla f(v)^T (u - v)$$

We get $D_{KL}(u, v) \geq 0$ for all $u, v \in \mathbb{R}^n_{++}$ (proved from the first expression).
For the equality case,
i) If $u = v$, then

$$D_{KL}(u, v) = 0 - 0 = 0$$

ii) If $D_{KL}(u, v) = 0$, then

$$\sum_{i=1}^{n} u_i \log\left(\frac{u_i}{v_i}\right) = \sum_{i=1}^{n} (u_i - v_i)$$

This is only possible when $u_i = v_i \ \forall i$, so $u = v$.

$$\therefore D_{KL}(u, v) = 0 \text{ if and only if } u = v.$$