**THE CHINESE UNIVERSITY OF HONG KONG**
**Department of Mathematics**
**MATH4230 2024-25**
**Lecture 12**
**February 19, 2025 (Wednesday)**

# 1 Recall

## Convex Function

**Definition 1.** Let $f : X \to \mathbb{R}$, where $X$ is a convex subset of $\mathbb{R}^n$. Then $f$ is **convex** if and only if for any $x, y \in X$ and $\lambda \in (0, 1)$,

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

Also, we discussed the following lemma yesterday.

**Lemma 1.** *Let $X \subseteq \mathbb{R}^n$ be non-empty convex set.*

1. *If $f : X \to \mathbb{R}$ is **differentiable**, then*

$$f(x) \text{ is convex} \iff f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle, \quad \forall x, y \in X$$

2. *If $f : X \to \mathbb{R}$ is **twice differentiable**, then*

$$f(x) \text{ is convex} \iff \nabla^2 f(x) \geq 0, \ \forall x \in X$$

*where $\nabla f^2(x) \geq 0$ denotes the Hessian matrix of $f$ is positive semidefinite.*

# 2 Non-differentiable Convex Functions

**Definition 2.** Let $X$ be a convex set and $f : X \to \mathbb{R}$ be a function. A vector $w \in \mathbb{R}^n$ is called a **subgradient** of $f$ at point $x \in X$ if

$$f(y) \geq f(x) + w^T(y - x), \ \ \forall y \in X$$

We denote $\partial f(x) = \{\text{all subgradient of } f \text{ at } x\}$.

**Definition 3.** The **epigraph** of $f : X \to \mathbb{R}$ is defined by

$$\text{epi}(f) := \{(x, t) : x \in X, \ t \geq f(x)\} \subseteq X \times \mathbb{R} \subseteq \mathbb{R}^n \times \mathbb{R}$$

**Lemma 2.** *Let $X$ be a convex set and $f : X \to \mathbb{R}$ be a function. Then*

$$f \text{ is convex} \iff \text{epi}(f) \text{ is convex.}$$

*Proof.* " $\implies$ " Let $(x_1, t_1), (x_2, t_2) \in \text{epi}(f)$ and $\lambda \in (0, 1)$. Then, by definition, we have

$$t_1 \geq f(x_1) \qquad \text{and} \qquad t_2 \geq f(x_2)$$

This implies that

$$\begin{aligned} \lambda t_1 + (1-\lambda)t_2 &\geq \lambda f(x_1) + (1-\lambda)f(x_2) \geq f(\lambda x_1 + (1-\lambda)x_2) \\ &\implies (\lambda x_1 + (1-\lambda)x_2, \ \lambda t_1 + (1-\lambda)t_2) \in \text{epi}(f) \end{aligned}$$

Thus, $\text{epi}(f)$ is convex.
" $\Longleftarrow$ " Suppose $f$ is not convex, then there exists $x_1, x_2 \in X$ and $\lambda \in (0,1)$ such that

$$f(\lambda x_1 + (1-\lambda)x_2) > \lambda f(x_1) + (1-\lambda) f(x_2)$$

At the same time, $(x_1, t_1), (x_2, t_2) \in \text{epi}(f)$ for $t_1 = f(x_1)$, $t_2 = f(x_2)$.
Then it implies that

$$\begin{aligned} &\implies \lambda t_1 + (1-\lambda)t_2 < f(\lambda x_1 + (1-\lambda)x_2) \\ &\implies (\lambda x_1 + (1-\lambda)x_2,\ \lambda t_1 + (1-\lambda)t_2) \notin \text{epi}(f) \end{aligned}$$

This would lead to $\text{epi}(f)$ is not convex. □

**Proposition 3.** *Let $X \subseteq \mathbb{R}^n$ be convex and $f : X \to \mathbb{R}$ be a function.*

1. *If $\partial f(x) \neq \emptyset$, for any $x \in X$, then $f$ is convex.*
2. *If $f$ is convex, then $\partial f(x) \neq \emptyset$ for any $x \in \text{ri}(X)$.*

*Proof.* 1. Let $x, y \in X$ and $\lambda \in (0,1)$. Define $z := \lambda x + (1-\lambda)y$.
Since the subgradient of $f(z)$ is non-empty, i.e. $\partial f(z) \neq \emptyset$, there exist $w \in \partial f(z)$.
By definition, we have $f(y) \geq f(z) + w^T(y-z)$ for all $y \in X$.
Therefore, we have

$$\begin{cases} f(y) \geq f(z) + w^T(y-z) & (1) \\ f(x) \geq f(z) + w^T(x-z) & (2) \end{cases}$$

Multiplying $(1)$ by $(1-\lambda)$ and $(2)$ by $\lambda$ then sum together yields

$$\begin{aligned} \lambda f(x) + (1-\lambda) f(y) &\geq f(z) + w^T(\lambda(x-z) + (1-\lambda)(y-z)) \\ &= f(z) + w^T[\lambda((1-\lambda)x + (\lambda-1)y) + (1-\lambda)(-\lambda x + (\lambda)y)] \\ &= f(z) + w^T(0) \\ &= f(z) \\ f(\lambda x + (1-\lambda)y) &\leq \lambda f(x) + (1-\lambda) f(y) \end{aligned}$$

This proves that $f$ is convex.

2. If $f$ is convex, then by the previous lemma, the **epigraph** $\text{epi}(f)$ is also convex.
Let $x \in \text{ri}(X)$, then $z := (x, f(x)) \in \text{epi}(f) \subseteq \mathbb{R}^{n+1}$.
Since $z + (0, -\varepsilon) = (x, f(x) - \varepsilon) \notin \text{epi}(f)$, $\forall \varepsilon > 0$, then the point $z$ is lying on the relative boundary of $\text{epi}(f)$.
By the separation theorem, there exist $(w, v) \in \mathbb{R}^n \times \mathbb{R}$ such that

$$\begin{aligned} &\langle (w,v), z \rangle \geq \langle (w,v), (y,t) \rangle, \quad \forall (y,t) \in \text{epi}(f) \\ \iff &w^T x + v f(x) \geq w^T y + vt, \quad \forall y \in X,\ t \geq f(y) \end{aligned} \quad (*)$$

Then, we claim the followings:

- We must have $v \leq 0$, otherwise $\lim\limits_{t \to +\infty} vt = +\infty$ and the inequality fail to satisfy.
- Also, we have $v \neq 0$, otherwise $w^T x \geq w^T y$ for any $y \in X$. This implies

$$w^T x \geq w^T \underbrace{(x + \varepsilon w)}_{\in X} = w^T x + \varepsilon \|w\|^2$$

and contradiction occurs since $\varepsilon \|w\|^2 > 0$.

From the above, we can conclude that $v < 0$.
Since $w^T x + vf(x) \geq w^T y + vt, \;\; \forall y \in X, \, t \geq f(y)$, dividing both sides by $v < 0$ yields:

$$\frac{w^T}{v} x + f(x) \leq \frac{w^T}{v} y + t, \;\; \forall y \in X, \; t \geq f(y)$$

Letting $t = f(y)$ in (*), we obtain

$$f(y) \geq f(x) + \left( \frac{w^T}{v} \right) (y - x), \;\; \forall y \in X$$

which implies that

$$\frac{w}{v} \in \partial f(x), \;\; \forall x \in \mathrm{ri}(X).$$

— End of Lecture 12 —

□