1.3 Convex Functions

In this course, we will consider extended-real-valued functions, which take values in $\overline{\mathbb{R}} := (-\infty, \infty]$, with the convention that $a + \infty = \infty \; \forall a \in \mathbb{R}$, $\infty + \infty = \infty$, and $t \cdot \infty = \infty \; \forall t > 0$.

1.3.1 Convex Functions

Definition:(Convex Functions) Let C be a convex subset of $\mathbb{R}^n$. A function $f : C \to \overline{\mathbb{R}}$ is called *convex* on C if

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)y, \forall x, y \in C, \forall \lambda \in [0,1].$$

A function is called *stricly convex* if the inequality above is strict for all $x, y \in C$ with $x \neq y$, and all $\lambda \in (0,1)$. A function is called *concave* if $(-f)$ is convex.

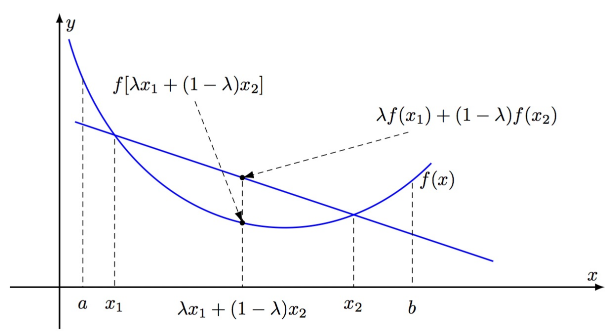

Figure 1: Convex Function

Definition:(Level Sets) For a function $f : C \to \mathbb{R}$, we define the *level sets* of f to be $\{x \mid f(x) \leq \lambda\}$.

If a function is convex, then all its level sets are also convex (Exercise). However, the convexity of all level sets of a function does not necessarily imply the convexity of the function itself.

Examples of Convex Functions
The following functions are convex:

(a) $f(x) := \langle a, x \rangle + b$ for $x \in \mathbb{R}^n$, where $a \in \mathbb{R}^n$ and $b \in \mathbb{R}$.

(b) $g(x) := ||x||$ for $x \in \mathbb{R}^n$.

(c) $h(x) := x^2$ for $x \in \mathbb{R}$.

(d) $F(x) := \frac{1}{2}x^T Ax$ for $x \in \mathbb{R}^n$, where A is a $n \times n$ symmetric positive semidefinite matrix. (i.e. $x^T Ax \geq 0$ for all $x \in \mathbb{R}^n$)

Definition:(Epigraph and Effective Domain)
The *epigraph* of a function $f : X \to [-\infty, \infty]$, where $X \subset \mathbb{R}^n$, is given by

$$\text{epi}f = \{(x, w)|\ x \in X,\ w \in \mathbb{R},\ f(x) \leqslant w\}.$$

The *effective domain* of f is given by

$$\text{dom}f = \{x|\ f(x) < \infty\}.$$

Note that dom$f$ is just the projection of epi$f$ on $\mathbb{R}^n$.

Definition:(Proper Function)
A function f is *proper* if $f(x) < \infty$ for at least one $x \in X$. f is *improper* if it is not proper. By considering epi$f$, it means that epi$f$ is not empty and does not conatin any vertical line.

Theorem:(Jensen inequality)
A function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is convex if and only if for any $\lambda_i \geq 0$ with $\sum \lambda_i = 1$ and for any elements $x_i \in \mathbb{R}^n$, it holds that

$$f\big(\sum \lambda_i x_i\big) \leq \sum \lambda_i f(x_i)$$

Proof. It suffices to prove that any convex function satisfies the Jensen inequality. We will prove this by induction.
The case $m = 1, 2$ are simple. So suppose the inequality holds for all $k \leq m$.
Suppose $\lambda_i \geq 0$ satisfies $\sum_{i=1}^{m+1} \lambda_i = 1$. Then $\sum_{i=1}^{m} \lambda_i = 1 - \lambda_{m+1}$.
If $\lambda_{m+1} = 1$, then $\lambda_i = 0$ for all i. Then the inequality holds.
So suppose $\lambda_{m+1} < 1$. Then

$$\sum_{i=1}^{m} \frac{\lambda_i}{1 - \lambda_{m+1}} = 1$$

and

$$
\begin{aligned}
f\big(\sum_{i=1}^{m+1} \lambda_i x_i\big) &= f\left((1-\lambda_{m+1})\sum_{i=1}^{m} \frac{\lambda_i}{1-\lambda_{m+1}} x_i + \lambda_{m+1} x_{m+1}\right) \\
&\leq (1-\lambda_{m+1}) f\left(\sum_{i=1}^{m} \frac{\lambda_i}{1-\lambda_{m+1}} x_i\right) + \lambda_{m+1} f(x_{m+1}) \\
&\leq (1-\lambda_{m+1}) \sum_{i=1}^{m} \frac{\lambda_i}{1-\lambda_{m+1}} f(x_i) + \lambda_{m+1} x_{m+1} \\
&= \sum_{i=1}^{m+1} \lambda_i f(x_i)
\end{aligned}
$$

□

The following gives a geometric characterization of convexity.

Proposition: A function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is convex if and only if epi$f \subset \mathbb{R}^{n+1}$ is convex.

Proof. Assume f is convex. Let $(x_1, t_1), (x_2, t_2) \in$ epi$f$ and $\lambda \in [0, 1]$. Then

$$f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda)f(x_2) \leq \lambda t_1 + (1-\lambda)t_2$$

Hence $(\lambda(x_1, t_1) + (1-\lambda)(x_2, t_2) \in$ epi$f$.
Conversely, suppose epi$f$ is convex. Let $x_1, x_2 \in$ dom$f$ and $\lambda \in [0, 1]$.
Since epi$f$ is convex, $\lambda(x_1, f(x_1)) + (1-\lambda)(x_2, f(x_2)) \in$ epi$f$. Then

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

Therefore, f is convex. □

Definition:(Closed function) If the epigragh of a function $f : X \to \overline{\mathbb{R}}$ is closed, we say that f is a *closed* function.

For example, the indicator funtion δ_X is convex if and only if X is convex, is closed if and only if X is closed, where

$$\delta_X(x) := \begin{cases} 0 & x \in X \\ \infty & \text{otherwise} \end{cases}$$

In fact, closedness is related to the concept of lower semicontinuity.
Recall that a function f is called *lower semicontinuous* at $x \in X$ if

$$f(x) \leq \liminf_{k\to\infty} f(x_k)$$

for every sequence $\{x_k\} \subset X$ with $x \to x_k$. f is *lower semicontinuous* if it is lower semicontinuous at each $x \in X$. f is *upper semicontinuous* if $-f$ is lower semicontinuous.

Proposition: Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be a function, then the following are equivalent:

(i) The level set $V_\gamma = \{x | f(x) \leq \gamma\}$ is closed for every γ.

(ii) f is lower semicontinuous.

(iii) epi$f$ is closed.

Proof. If $f(x) = \infty$ for all x, then the result holds. So assume $f(x) < \infty$ for some $x \in \mathbb{R}^n$. Therefore, epi$f$ is nonempty and there exists level sets of f that are nonempty.
(i) $\implies$ (ii). Assume V_γ is closed for every γ. Suppose f is not lower semicontinuous, that is

$$f(x) > \liminf_{k \to \infty} f(x_k)$$

for some x and sequence $\{x_k\}$ converging to x. Let γ satisfies

$$f(x) > \gamma > \liminf_{k \to \infty} f(x_k).$$

Hence, there exists a subsequence $\{x_{k_i}\}$ such that $f(x_{k_i}) \leq \gamma$ for all i. So, $\{x_{k_i}\} \subset V_\gamma$. But V_γ is closed, x also belongs to V_γ. Therefore, $f(x) \leq \gamma$, contradiction.
(ii) $\implies$ (iii). Assume f is lower semicontinuous. Let (x, w) be the limit of $\{(x_k, w_k)\} \subset \text{epi}(f)$. We have $f(x_k) \leq w_k$ for all k. Since f is lower semicontinuous, taking limit we have,

$$f(x) \leq \liminf_{k \to \infty} f(x_k) \leq w.$$

Hence $(x, w) \in$ epi$f$ and so epi$f$ is closed.
(iii) $\implies$ (i). Assume epi$f$ is closed. Let $\{x_k\}$ be a sequence in V_γ converging to x for some γ. We have $f(x_k) \leq \gamma$, so $(x_k, \gamma) \in$ epi$f$ for each k. Since epi$f$ is closed and $(x_k, \gamma) \to (x, \gamma)$, we have $(x, \gamma) \in$ epi$f$, that is $f(x) \leq \gamma$. Hence $x \in V_\gamma$ and V_γ is closed. □