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.

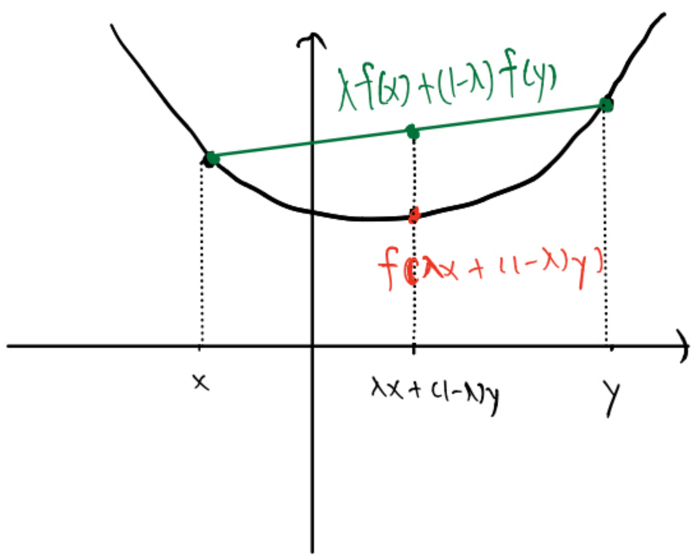

Figure 1: Convex Function

1.3.2 Characterizations of Differentiable Convex Functions

We now give some characterizations of convexity for once or twice differentiable functions.

Proposition: Let C be a nonempty convex open set. Let $f : \mathbb{R}^n \to \mathbb{R}$ be differentiable over an open set that contains C.

(a) f is convex if and only if $f(z) \geq f(x) + \langle \nabla f(x), (z-x) \rangle$, for all $x, z \in C$.

(b) f is stricly convex if and only if the above inequality is strict for $x \neq z$.

Proof. ($\Longleftarrow$) Let $x, y \in C$, $\lambda \in [0,1]$ and $z = \lambda x + (1-\lambda)y$. We have,

$$f(x) \geq f(z) + \langle \nabla f(z), (x-z) \rangle$$

$$f(y) \geq f(z) + \langle \nabla f(z), (y-z) \rangle.$$

Then,

$$\lambda f(x) + (1-\lambda) f(y) \geq f(z) + \langle \nabla f(z), \lambda(x-z) + (1-\lambda)(y-z) \rangle = f(z) = f(\lambda x + (1-\lambda) y)$$

Hence f is convex.
Conversely, suppose f is convex. For $x \neq z$, define $g : (0,1] \to \mathbb{R}$ by

$$g(t) = \frac{f(x + t(z-x)) - f(x)}{t}.$$

Consider t_1, t_2 with $0 < t_1 < t_2 < 1$. Let $\bar{t} = \frac{t_1}{t_2}$ and $\bar{z} = x + t_2(z-x)$. Then $f(x + \bar{t}(\bar{z} - x)) \leq \bar{t} f(\bar{z}) + (1-\bar{t}) f(x)$. So,

$$\frac{f(x + \bar{t}(\bar{z}-x)) - f(x)}{\bar{t}} \leq f(\bar{z}) - f(x).$$

Therefore,

$$\frac{f(x + t_1(z-x)) - f(x)}{t_1} \leq \frac{f(x + t_2(z-x)) - f(x)}{t_2}.$$

So, $g(t_1) \leq g(t_2)$, that is, g is monotonically increasing.
Then $\langle \nabla f(x), (z-x) \rangle = \lim_{t \downarrow 0} g(t) \leq g(1) = f(z) - f(x)$. So we are done.
The proof for (b) is the same as (a), we just change all inequality to strict inequality. □

For twice differentiable functions, we have the following characterization.
Proposition: Let C be a nonempty convex set $\subset \mathbb{R}^n$ and $f : \mathbb{R}^n \to \mathbb{R}$ be twice continuously differentiable over an open set that contains C. Then:

(a) If $\nabla^2 f(x)$ is positive semidefinite for all $x \in C$, then f is convex over C.

(b) If $\nabla^2 f(x)$ is positive definite for all $x \in C$, then f is strictly convex over C.

(c) If C is open and f is convex over C, then $\nabla^2 f(x)$ is positive semidefinite for all $x \in C$.

Proof. (a) For all $x, y \in C$,

$$f(y) = f(x) + \langle \nabla f(x), (y-x) \rangle + \frac{1}{2}(y-x)^T \nabla^2 f(x + \alpha(y-x))(y-x)$$

for some $\alpha \in [0, 1]$. Since $\nabla^2 f$ is positive semidefinite, we have

$$f(y) \geq f(x) + \langle \nabla f(x), (y-x) \rangle, \forall x, y \in C.$$

Hence, f is convex over C.
(b) We have $f(y) > f(x) + \langle \nabla f(x), (y-x) \rangle$ for all $x.y \in C$ with $x \neq y$ since $\nabla^2 f$ is positive definite.
(c) Assume there exist $x \in C$ and $z \in \mathbb{R}^n$ such that $z^T \nabla^2 f(x) z < 0$. For z with sufficiently small norm, we have $x + z \in C$ and $z^T \nabla^2 f(x + \alpha z) z < 0$ for all $\alpha \in [0, 1]$. Then

$$f(x+z) = f(x) + \langle \nabla f(x), z \rangle + z^T \nabla^2 f(x + \alpha z) z < f(x) + \langle \nabla f(x), z \rangle.$$

This contradicts the convexity of f over C. Hence, $\nabla^2 f$ is indeed positive semidefinite over C. □