1 Convex Functions

1.1 Definition and Basic Properties

Let C be a convex subset of $\mathbb{R}^n$. A function $f : C \to \mathbb{R}$ is called *convex* if

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

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

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

1.2 Characterizations of Differentiable Convex Functions

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

Propopsition: 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 $f(z) \geq f(x) + \nabla f(x)'(z-x)$, 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 $x, y \in C$, $\alpha \in [0,1]$ and $z = \alpha x + (1-\alpha)y$. We have,

$$f(x) \geq f(z) + \nabla f(z)'(x-z)$$

$$f(y) \geq f(z) + \nabla f(z)'(y-z).$$

Then,

$$\alpha f(x) + (1-\alpha) f(y) \geq f(z)'(\alpha(x-z) + (1-\alpha)(y-z)) = f(z) = f(\alpha x + (1-\alpha)y)$$

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

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

Consider α_1, α_2 with $0 < \alpha_1 < \alpha_2 < 1$. Let $\overline{\alpha} = \frac{\alpha_1}{\alpha_2}$ and $\overline{z} = x + \alpha_2(z - x)$. Then $f(x + \overline{\alpha}(z - x)) \leq \overline{\alpha} f(\overline{z}) + (1 - \overline{\alpha}) f(x)$. So,

$$\frac{f(x + \overline{\alpha}(z - x)) - f(x)}{\overline{\alpha}} \leq f(\overline{z}) - f(x).$$

Therefore,

$$\frac{f(x + \alpha_1(z - x)) - f(x)}{\alpha_1} \leq \frac{f(x + \alpha_2(z - x)) - f(x)}{\alpha_2}.$$

So, $g(\alpha_1) \leq g(\alpha_2)$, that is, g is monotonically increasing.
Then $\nabla f(x)'(z - x) = \lim_{\alpha \downarrow 0} g(\alpha) \leq g(1) = f(z) - f(x)$. So we are done.
The proof for (b) is the same as (a), we just change all inequality to strict inequality. □

As a consequence of the above proposition, we have a useful optimal condition for unconstrained optimization. Recall that in singile variable calculus, if $f : \mathbb{R} \to \mathbb{R}$ is a differentiable convex function with $f'(x^*) = 0$, then x^* minimizes f. Here, we give a similar result.

Proposition: Let C be a nonempty convex set and let $f : \mathbb{R}^n \to \mathbb{R}$ be a convex differentiable function over an open set that contains C. Then $x^* \in C$ minimizes f over C if and only if

$$\nabla f(x^*)'(z - x^*) \geq 0, \forall z \in C.$$

Proof. Suppose $\nabla f(x^*)'(z - x^*) \geq 0, \forall z \in C$, then by Prop 1.1.3(a), we have,

$$f(z) - f(x^*) \geq \nabla f(x^*)'(z - x^*) \geq 0, \forall z \in C.$$

Hence x^* indeed minimizes f over C.
Conversely, suppose x^* minimizes f over C. Suppose on the contrary that $\nabla f(x^*)'(z - x^*) < 0$ for some $z \in C$, then

$$\lim_{\alpha \downarrow 0} \frac{f(x^* + \alpha(z - x^*) - f(x^*)}{\alpha} = \nabla f(x^*)'(z - x^*) < 0.$$

Then for sufficiently small α, we have $f(x^* + \alpha(z - x^*) - f(x^*) < 0$, contradicting the optimality of x^*. □

For twice continuously differentiable functions, we have the following characterization.

Proposition 2.1.5: 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 $x, y \in C$,

$$f(y) = f(x) + \nabla f(x)'(y-x) + \frac{1}{2}(y-x)'\nabla^2 f(x + \alpha(y-x))(y-x)$$

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

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

Hence, f is convex over C.

(b) We have $f(y) > f(x) + \nabla f(x)(y-x)$ 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 $z'\nabla^2 f(x)z < 0$. There exists $\epsilon > 0$ such that $x + \epsilon z \in C$ and $z'\nabla^2 f(x + \alpha\epsilon z)z < 0$ for all $\alpha \in [0,1]$. Then

$$f(x+z) = f(x) + \nabla f(x)'z + z'\nabla^2 f(x + \alpha z)z < f(x) + \nabla f(x)'z.$$

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