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 strictly convex if and only if the above inequality is strict for $x \neq z$.

**Proof.** ($\Leftarrow$) Let $x, y \in C$, $\alpha \in [0,1]$ and $z = \alpha x + (1-\alpha)y$. 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,

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

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

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

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 $\langle \nabla f(x), (z-x) \rangle = \lim_{\alpha \downarrow 0} g(\alpha) \leq g(1) = f(z) - f(x)$. So we are done.

The proof for (b) is the same as (a), we just change all inequality to strict inequality. \hfill \Box

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 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 $C$.  

\hfill 1
(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 $C$. \qed

1.4 Relative Interior

Consider $I = [0, 1] \subset \mathbb{R}$. Then the interior of $I$ is (0,1). However, if we consider $I$ as a subset in $\mathbb{R}^2$, then the interior of $I$ is empty. This motivates the following definition.

Definition: (Relative Interior) Let $C \subset \mathbb{R}^n$. We say that $x$ is a relative interior point of $C$ if $x \in B(x; \epsilon) \cap \text{aff}(C) \subset C$, for some $\epsilon > 0$. The set of all relative interior point of $C$ is called the relative interior of $C$, and is denoted by $\text{ri}(C)$. The relative boundary of $C$ is equal to $\text{cl}(C) \setminus \text{ri}(C)$.

Lemma: Let $\Delta_m$ be an m-simplex in $\mathbb{R}^n$ with $m \geq 1$. Then $\text{ri}(\Delta_m) \neq \emptyset$.

Proof. Let $x_0, ..., x_m$ be the vertices of $\Delta_m$. Let

$$\bar{x} := \frac{1}{m+1} \sum_{i=0}^{m} x_i$$

Note that $V := \text{span}\{x_1 - x_0, ..., x_m - x_0\}$ is the m-dimensional subspace parallel to $\text{aff}(\Delta_m) = \text{aff}\{x_0, ..., x_m\}$. Hence for all $x \in V$, there exists unique $\lambda_i$ such that

$$x = \sum_{i=1}^{m} \lambda_i (x_i - x_0)$$

2
Let $\lambda_0 := -\sum_{i=1}^{m} \lambda_i$, then $(\lambda_0, ..., \lambda_m) \in \mathbb{R}^{m+1}$ and

$$x = \sum_{i=0}^{m} \lambda_i x_i, \text{ with } \sum_{i=0}^{m} \lambda_i = 0$$

Let $L : V \rightarrow \mathbb{R}^{m+1}$ be the mapping that sends $x$ to $(\lambda_0, ..., \lambda_m)$. It is easy to check that $L$ is linear and thus continuous. Hence there exists $\delta$ such that $||L(u)|| < 1$ if $||u|| < \delta$.

Let $x \in (\bar{x} + B(0, \delta)) \cap \text{aff}(\Delta_m)$ Then, $x = \bar{x} + u$, where $||u|| < \delta$.

Since $x, \bar{x} \in \text{aff}(\Delta_m)$ and $u = x - \bar{x}$, $u \in V$. Hence $||L(u)|| < \frac{1}{m+1}$.

Suppose $L(u) = (\mu_0, ..., \mu_m)$, then $u = \sum_{i=0}^{m} \mu_i x_i$ and $x = \sum_{i=0}^{m} (\frac{1}{m+1} + \mu_i) x_i$.

Since $\sum_{i=0}^{m} \mu_i = 0$, $\sum_{i=0}^{m} (\frac{1}{m+1} + \mu_i) = 1$. Therefore, $x \in \Delta_m$.

Thus $(\bar{x} + B(0; \delta)) \cap \text{aff}(\Delta_m) \subseteq \Delta_m$, so $\bar{x} \in \text{ri}(\Delta_m)$.

**Proposition:** Let $C$ be a nonempty convex set. Then $\text{ri}(C)$ is nonempty.

**Proof.** Let $m$ be the dimension of $C$.

If $m = 0$, then $C$ must be a singleton. Hence $\text{ri}(C) \neq \emptyset$.

Suppose $m \geq 1$. We first show that there exists $m+1$ affinely independent elements $x_0, ..., x_m \in C$.

Let $\{x_0, ..., x_k\}$ be a maximal affinely independent set in $C$.

Consider $K := \text{aff}(\{x_0, ..., x_k\})$. $K \subseteq \text{aff}(C)$ since $\{x_0, ..., x_m\} \subseteq C$.

Suppose $y \in C$ but $y \notin K$. Then, $\{x_0, ..., x_k, y\}$ is also affinely independent, which is a contradiction. Therefore $C \subseteq K$ and hence $\text{aff}(C) \subseteq K$. Then

$$k = \dim(K) = \dim(\text{aff}(C)) = m$$

Therefore, there exists $m+1$ affinely independent elements $x_0, ..., x_m \in C$.

Let $\Delta_m$ be the $m$-simplex formed by $\{x_0, ..., x_m\}$. By above, $\text{aff}(\Delta_m) = \text{aff}(C)$.

Since $\text{ri}(\Delta_m)$ is not empty, it follows that $\text{ri}(C)$ is also nonempty. □