SYS 6003: Optimization	Fall 2015
Lecture 6	
Instructor: Quanquan Gu	*Date: Sep 12^{th}*

The following theorem provides a sufficient and necessary condition to verify a function is convex.

Theorem 1 (First order condition for convex functions) *Suppose $f : \mathbb{R}^d \to \mathbb{R}$ is a continuously differentiable function over its convex domain $\mathbf{dom} f$, then $f(\mathbf{x})$ is convex if and only if*

$$f(\mathbf{y}) \geq f(\mathbf{x}) + \nabla f(\mathbf{x})^\top (\mathbf{y} - \mathbf{x}) \tag{1}$$

for all $\mathbf{x}, \mathbf{y} \in \mathbf{dom} f$.

Proof: We first prove the forward direction "⇒"
Suppose that $f(\mathbf{x})$ is convex, then for any $\mathbf{x}, \mathbf{y} \in \mathbf{dom} f$ and any $\alpha \in [0, 1]$, we have $f(\alpha\mathbf{y} + (1-\alpha)\mathbf{x}) \leq \alpha f(\mathbf{y}) + (1-\alpha) f(\mathbf{x})$. Rearranging this inequality leads to

$$f(\mathbf{y}) \geq \frac{f(\alpha\mathbf{y} + (1-\alpha)\mathbf{x}) - (1-\alpha) f(\mathbf{x})}{\alpha} = f(\mathbf{x}) + \frac{f(\alpha\mathbf{y} + (1-\alpha)\mathbf{x}) - f(\mathbf{x})}{\alpha}. \tag{2}$$

By Taylor expansion, we have

$$f(\mathbf{x} + \alpha(\mathbf{y} - \mathbf{x})) = f(\mathbf{x}) + \alpha \nabla f(\mathbf{x})^\top (\mathbf{y} - \mathbf{x}) + o(\alpha).$$

where $o(\alpha)$ means $\lim_{t\to 0} o(\alpha)/\alpha = 0$. Also note that $\mathbf{x} + \alpha(\mathbf{y} - \mathbf{x}) = \alpha\mathbf{y} + (1-\alpha)\mathbf{x}$. Thus, it follows from (2) that

$$f(\mathbf{y}) \geq f(\mathbf{x}) + \frac{\alpha \nabla f(\mathbf{x})^\top (\mathbf{y} - \mathbf{x}) + o(\alpha)}{\alpha} = f(\mathbf{x}) + \nabla f(\mathbf{x})^\top (\mathbf{y} - \mathbf{x}) + \frac{o(\alpha)}{\alpha},$$

which immediately leads to (1) by taking $\alpha \to 0$.

Now we prove the backward direction "⇐":
We want to show that, for any $\mathbf{x}, \mathbf{y} \in \mathbf{dom} f$ and any $\alpha \in [0, 1]$,

$$f(\alpha\mathbf{x} + (1-\alpha)\mathbf{y}) \leq \alpha f(\mathbf{x}) + (1-\alpha) f(\mathbf{y}).$$

Let $\mathbf{z} \equiv \alpha\mathbf{x} + (1-\alpha)\mathbf{y}$, since $\mathbf{dom} f$ is a convex set, we have $\mathbf{z} \in \mathbf{dom} f$. Since $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbf{dom} f$, we have

$$f(\mathbf{x}) \geq f(\mathbf{z}) + \nabla f(\mathbf{z})^\top (\mathbf{x} - \mathbf{z}) \tag{3}$$

$$f(\mathbf{y}) \geq f(\mathbf{z}) + \nabla f(\mathbf{z})^\top (\mathbf{y} - \mathbf{z}). \tag{4}$$

Now multiply inequality (3) by α and inequality (4) by $(1-\alpha)$ to obtain:

$$\alpha f(\mathbf{x}) + (1-\alpha) f(\mathbf{y}) \geq \alpha f(\mathbf{z}) + \alpha \nabla f(\mathbf{z})^\top (\mathbf{x} - \mathbf{z}) + (1-\alpha) f(\mathbf{z}) + (1-\alpha) \nabla f(\mathbf{z})^\top (\mathbf{y} - \mathbf{z})$$

And we are left with the right hand side equal to:

$$\begin{aligned} f(\mathbf{z}) + \nabla f(\mathbf{z})^\top(\alpha\mathbf{x} + \alpha\mathbf{z} + (1+\alpha)(\mathbf{y}-\mathbf{z})) &= f(\mathbf{z}) + \nabla f(\mathbf{z})^\top(\alpha\mathbf{x} + (1-\alpha)\mathbf{y} - \mathbf{z}) \\ &= f(\mathbf{z}) + \nabla f(\mathbf{z})^\top(\mathbf{z}-\mathbf{z}) \\ &= f(\mathbf{z}). \end{aligned}$$

And since we know $f(\mathbf{z}) = f(\alpha\mathbf{x}+(1-\alpha)\mathbf{y})$, we can conclude this is indeed a convex function.
■

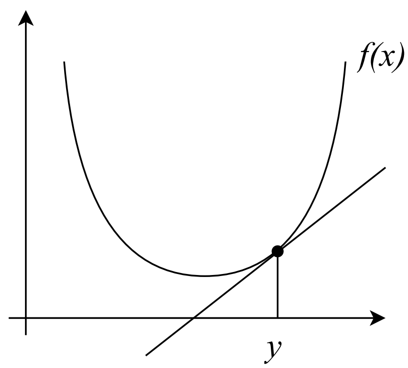

Figure 1: Illustrating the 1^{st} Order Condition for Convex Functions

In order to prove that a function is convex, we can use the definition. But sometimes that can be tedious. In the following, we will introduce second order sufficient and necessary condition for convex functions, which provides an easy way to prove a function is convex.

Theorem 2 (Second order condition for convex functions) *Suppose $f : \mathbb{R}^d \to \mathbb{R}$ is twice continuously differentiable over its convex domain* $\mathbf{dom}f$, *then f is convex if and only if,*

$$\nabla^2 f(\mathbf{x}) \succeq 0 \textit{ for all } \mathbf{x} \in \mathbf{dom}f$$

Proof: By the mean value theorem, we have:

$$f(\mathbf{y}) = f(\mathbf{x}) + \nabla f(\mathbf{x})^\top(\mathbf{y}-\mathbf{x}) + \frac{1}{2}(\mathbf{y}-\mathbf{x})^\top\nabla^2 f(\mathbf{z})(\mathbf{y}-\mathbf{x}), \tag{5}$$

where $\mathbf{z} = \alpha\mathbf{x} + (1-\alpha)\mathbf{y}$, $\alpha \in [0,1]$. Note that since $\mathbf{dom}f$ is convex, we have $\mathbf{z} \in \mathbf{dom}f$.

We first prove the forward direction "⇒":
Since f is convex, by the first order condition, we have for any $\mathbf{x}, \mathbf{y} \in \mathbf{dom}f$,

$$f(\mathbf{y}) \geq f(\mathbf{x}) + \nabla f(\mathbf{x})^\top(\mathbf{y}-\mathbf{x}). \tag{6}$$

Therefore, by combining (5) and (6), we have

$$\frac{1}{2}(\mathbf{y}-\mathbf{x})^\top\nabla^2 f(\mathbf{z})(\mathbf{y}-\mathbf{x}) \geq 0.$$

Let $\mathbf{y} \to \mathbf{x}$, then $\mathbf{z} \to \mathbf{x}$. By the continuity of $\nabla^2 f(\mathbf{x})$, we then have

$$(\mathbf{y} - \mathbf{x})^\top \nabla^2 f(\mathbf{x})(\mathbf{y} - \mathbf{x}) \geq 0.$$

Due to the arbitrariness of $\mathbf{y}$ and $\mathbf{x}$, it follows that $\nabla^2 f(\mathbf{x}) \succeq 0$.

We now prove the backward direction "$\Leftarrow$":

Consider any $\mathbf{x}, \mathbf{y} \in \mathbf{dom} f$ where $\alpha \in [0, 1]$. Let $\mathbf{z} = \alpha\mathbf{x} + (1-\alpha)\mathbf{y}$, since $\mathbf{dom} f$ is a convex set, we have $\mathbf{z} \in \mathbf{dom} f$. Since $\nabla^2 f(\mathbf{x}) \succeq 0$ for all $\mathbf{x} \in \mathbf{dom} f$, we have $\nabla^2 f(\mathbf{z}) \succeq 0$. For any $\mathbf{y} \in \mathbf{dom} f$, by the definition of positive semidefinite, we then know

$$(\mathbf{y} - \mathbf{x})^\top \nabla^2 f(\mathbf{z})(\mathbf{y} - \mathbf{x}) \geq 0. \tag{7}$$

Therefore, by combining (5) and (7), we have $f(\mathbf{y}) \geq f(\mathbf{x}) + \nabla f(\mathbf{x})^\top (\mathbf{y} - \mathbf{x})$. By the first-order condition for convex functions, f is convex. ■

Now we will illustrate the application of second-order condition for convex functions with several examples.

Example 1 (Quadratic Function) $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\top \mathbf{P}\mathbf{x} + \mathbf{q}^\top \mathbf{x} + r$, *where* $\mathbf{P} \in \mathbb{R}^{d\times d}, \mathbf{P} \succeq 0, \mathbf{q} \in \mathbb{R}^d$, $r \in \mathbb{R}$, $\mathbf{x} \in \mathbb{R}^d$.

$f(\mathbf{x})$ *is convex, since*

$$\nabla^2 f(\mathbf{x}) = \mathbf{P} \succeq 0.$$

Example 2 (Loss function of Linear Regression)

$$f(\mathbf{x}) = \frac{1}{2}\|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2^2,$$

where $\mathbf{A} \in \mathbb{R}^{n\times d}, \mathbf{x} \in \mathbb{R}^d$, $\mathbf{b} \in \mathbb{R}^n$. $f(\mathbf{x})$ *is convex, since*

$$\nabla^2 f(\mathbf{x}) = \mathbf{A}^\top \mathbf{A} \succeq 0,$$