SYS 6003: Optimization	Fall 2016

Lecture 7

Instructor: Quanquan Gu — *Date: September 14^{th}*

We continue to illustrate the application of second-order condition for convex functions with more examples.

Example 1 (Quadratic over Linear Function)

$$f(x,y) = \frac{x^2}{y},\ y > 0.$$

$f(x,y)$ is convex over $\mathbb{R} \times (0, +\infty)$. To show this, we first calculate the partial derivatives. The first order partial derivatives are:

$$\frac{\partial f(x,y)}{\partial x} = \frac{2x}{y}, \quad \frac{\partial f(x,y)}{\partial y} = -\frac{x^2}{y^2}.$$

The second order partial derivatives of $f(x,y)$ are:

$$\frac{\partial^2 f(x,y)}{\partial x^2} = \frac{2}{y}, \quad \frac{\partial^2 f(x,y)}{\partial y^2} = \frac{2x^2}{y^3}, \quad \frac{\partial^2 f(x,y)}{\partial x \partial y} = -\frac{2x}{y^2}.$$

Then we can write down the Hessian matrix of $f(x,y)$ as:

$$\nabla^2 f(x,y) = \begin{bmatrix} \frac{2}{y} & -\frac{2x}{y^2} \\ -\frac{2x}{y^2} & \frac{2x^2}{y^3} \end{bmatrix}.$$

Factoring out $2/y^3$, we can achieve:

$$\nabla^2 f(x,y) = \frac{2}{y^3} \begin{bmatrix} y^2 & -xy \\ -xy & x^2 \end{bmatrix}.$$

Note that the matrix can be factorized as the outer product of two vectors, yielding

$$\nabla^2 f(x,y) = \frac{2}{y^3} \begin{pmatrix} y \\ -x \end{pmatrix} (y, -x),$$

where we notice that:

$$\begin{pmatrix} y \\ -x \end{pmatrix} (y, -x) \succeq 0.$$

Therefore we have:

$$\nabla^2 f(x,y) \succeq 0.$$

By the second order condition, we know that $f(x,y)$ is convex.

Example 2 (Log-sum-exponential Function) *$f : \mathbb{R}^d \to \mathbb{R}$ is defined as follows*

$$f(\mathbf{x}) = \log\left[\sum_{i=1}^{d} \exp(x_i)\right]. \tag{1}$$

It is a convex function.

Example 3 (Geometric Mean) *$f : \mathbb{R}^d \to \mathbb{R}$ is defined as follows*

$$f(\mathbf{x}) = \left[\prod_{i=1}^{d} x_i\right]^{1/d}. \tag{2}$$

*It is a **concave** function.*

For convex function, we can show that its local minimum is also a global minimum. In detail, the following theorem shows that, a local minimum of a convex function is also a global minimum.

Theorem 1 (Local Minimum is also a Global Minimum) *Let $f\mathbb{R}^d \to \mathbb{R}$ be convex. If $\mathbf{x}^*$ is a local minimum of f over a convex set $\mathcal{D}$, then $\mathbf{x}^*$ is also a global minimum of f over a convex set $\mathcal{D}$.*

Proof: Since $\mathcal{D}$ is a convex set, for any $\mathbf{y}$, $\mathbf{y} - \mathbf{x}^*$ is a feasible direction. Since $\mathbf{x}^*$ is a local minimum, for any $\mathbf{y} \in \mathcal{D}$, we can choose a small enough $\alpha > 0$, such that

$$f(\mathbf{x}^*) \leq f(\mathbf{x}^* + \alpha(\mathbf{y} - \mathbf{x}^*)). \tag{3}$$

Furthermore, since f is convex, we have

$$f(\mathbf{x}^* + \alpha(\mathbf{y} - \mathbf{x}^*)) = f(\alpha\mathbf{y} + (1-\alpha)\mathbf{x}^*) \leq \alpha f(\mathbf{y}) + (1-\alpha) f(\mathbf{x}^*). \tag{4}$$

Combining (3) and (4), we have

$$f(\mathbf{x}^*) \leq \alpha f(\mathbf{y}) + (1-\alpha) f(\mathbf{x}^*),$$

which implies that $f(\mathbf{x}^*) \leq f(\mathbf{y})$. Since $\mathbf{y}$ is an arbitrary point in $\mathcal{D}$, this immediately proves that $\mathbf{x}^*$ is a global minimum. ■

Theorem 2 (First-order Condition for a Global Minimum) *Let function $f : \mathbb{R}^d \to \mathbb{R}$ be convex and continuously differentiable. $\mathbf{x}^*$ is a global minimum of f over a convex set $\mathcal{D}$ if and only if,*

$$\nabla f(\mathbf{x}^*)^\top(\mathbf{x} - \mathbf{x}^*) \geq 0, \quad \textit{for all} \quad \mathbf{x} \in \mathcal{D}. \tag{5}$$

Proof: "$\Rightarrow$"
Since $\mathbf{x}^*$ is a global minimum, $\mathbf{x}^*$ must also be a local minimum. By the first order necessary condition of a local minimum, we have $\nabla f(\mathbf{x}^*)^\top \mathbf{d} \geq 0$ where $\mathbf{d}$ is a feasible direction. For any $\mathbf{x} \in \mathcal{D}$, $\mathbf{d} = \mathbf{x} - \mathbf{x}^*$ is a feasible direction. Then we obtain:

$$\nabla f(\mathbf{x}^*)^\top (\mathbf{x} - \mathbf{x}^*) \geq 0$$

Thus, this completes the proof in the forward direction.

"$\Leftarrow$"
By definition, we have that:

$$f(\mathbf{x}) \geq f(\mathbf{x}^*) + \nabla f(\mathbf{x}^*)^\top (\mathbf{x} - \mathbf{x}^*) \text{ for any } \mathbf{x} \in \mathcal{D}.$$

Thus, if $\nabla f(\mathbf{x}^*)^\top (\mathbf{x} - \mathbf{x}^*) \geq 0$, then $f(\mathbf{x}) - f(\mathbf{x}^*) \geq \nabla f(\mathbf{x}^*)^\top (\mathbf{x} - \mathbf{x}^*) \geq 0$, which means $\mathbf{x}^*$ is a global minimum of f over $\mathcal{D}$. ■

In the following, we will show another way to prove that a function is convex. First of all, let's introduce the restriction of a function to a line.

Let $f : \mathbb{R}^d \to \mathbb{R}$ be a function. The restriction of f to a line $\mathbf{x} + t\mathbf{v}$ is defined as $g : \mathbb{R} \to \mathbb{R} : g(t) = f(\mathbf{x} + t\mathbf{v})$, where $\mathbf{dom}(g) = \{t : \mathbf{x} + t\mathbf{v} \in \mathbf{dom}(f)\}$.

Theorem 3 (Restriction of a convex function to a line) *$f : \mathbb{R}^d \to \mathbb{R}$ is a convex function if and only if the function $g : \mathbb{R} \to \mathbb{R} : g(t) = f(\mathbf{x} + t\mathbf{v}), \mathbf{dom}(g) = \{t : \mathbf{x} + t\mathbf{v} \in \mathbf{dom}(f)\}$ is convex for any $\mathbf{x} \in \mathbf{dom}(f), \mathbf{v} \in \mathbb{R}^d$*

Proof: "$\Rightarrow$": f is convex $\to$ g is convex.
For any $t_1, t_2 \in \mathbf{dom}(g)$ and any $\alpha \in [0, 1]$, we have

$$\begin{aligned} g(\alpha t_1 + (1-\alpha)t_2) &= f(\mathbf{x} + (\alpha t_1 + (1-\alpha)t_2)\mathbf{v}) \\ &= f(\alpha \mathbf{x} + \alpha t_1 \mathbf{v} + (1-\alpha)\mathbf{x} + (1-\alpha)t_2 \mathbf{v}) \\ &= f(\alpha(\mathbf{x} + t_1\mathbf{v}) + (1-\alpha)(\mathbf{x} + t_2\mathbf{v})) \end{aligned}$$

Since $f(\mathbf{x})$ is convex, it then follows that

$$\begin{aligned} g(\alpha t_1 + (1-\alpha)t_2) &\leq \alpha f(\mathbf{x} + t_1\mathbf{v}) + (1-\alpha) f(\mathbf{x} + t_2\mathbf{v}) \\ &= \alpha g(t_1) + (1-\alpha) g(t_2), \end{aligned}$$

where the last equality follows by the definition of $g(t)$. Thus, by definition, $g(t)$ is convex.

"$\Leftarrow$" g is convex $\to$ f is convex.
For any $\mathbf{x}, \mathbf{y} \in \mathbf{dom}(f)$ and any $\alpha \in [0, 1]$, we want to show

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

Let $\mathbf{v} = \mathbf{y} - \mathbf{x}$, and consider $g(t) = f(\mathbf{x} + t(\mathbf{y} - \mathbf{x}))$. It is easy to verify that $g(0) = f(\mathbf{x})$, $g(1) = f(\mathbf{y})$, and $g(1-\alpha) = f(\alpha\mathbf{x} + (1-\alpha)\mathbf{y})$. We then have

$$\begin{aligned} f(\alpha\mathbf{x} + (1-\alpha)\mathbf{y}) &= g(1-\alpha) \\ &= g(\alpha 0 + (1-\alpha)\cdot 1) \\ &\leq \alpha g(0) + (1-\alpha)g(1) \\ &= \alpha f(\mathbf{x}) + (1-\alpha)f(\mathbf{y}). \end{aligned} \tag{6}$$

Therefore, by definition, $f(\mathbf{x})$ is a convex function. ■

Theorem 3 basically suggests that a function is convex if and only if the restriction of this function to any lines is convex. It enables us to check convexity of f by checking convexity of functions of one variable.