SYS 6003: Optimization	**Fall 2016**
Lecture 5	
Instructor: Quanquan Gu	*Date: Sep 7^{th}*

In last lecture, I have introduced two convexity preserving operations: intersection of sets, and affine functions. I continue to introduce two more convexity preserving operations.

Definition 1 *A perspective function is a function $f : \mathbb{R}^d \to \mathbb{R}^{d-1}$ such that*

$$f(\mathbf{x}) = \begin{bmatrix} x_1/x_d \\ x_2/x_d \\ \vdots \\ x_{d-1}/x_d \end{bmatrix} \in \mathbb{R}^{d-1}, \text{ where } \mathbf{x} \in \mathbb{R}^d \text{ and } x_d > 0.$$

In general, a perspective function can also be written as $f(\mathbf{x}, t) = \mathbf{x}/t$, where $\mathbf{x} \in \mathbb{R}^{d-1}$, and $t > 0$.

Claim 1 *If f is a perspective function defined on a convex set $\mathcal{S}$, then $f(\mathcal{S})$ is also convex.*

Proof: First note that $f(\mathcal{S}) = \{\mathbf{y} \in \mathbb{R}^{d-1} : \exists(\mathbf{x}, t) \in \mathcal{S}, \text{ such that } \mathbf{y} = \mathbf{x}/t\}$. Consider any $\mathbf{y}_1, \mathbf{y}_2 \in f(S)$ where $\mathbf{y}_1 = \mathbf{x}_1/t_1$, and $\mathbf{y}_2 = \mathbf{x}_2/t_2$. We need to prove that for any $\alpha \in [0, 1]$, $\mathbf{y} = \alpha\mathbf{y}_1 + (1-\alpha)\mathbf{y}_2 \in f(S)$. In other words, there exist $(\mathbf{x}', t') \in \mathcal{S}$ such that $\mathbf{y} = \mathbf{x}'/t'$.

Since $(\mathbf{x}_1, t_1), (\mathbf{x}_2, t_2)$ are two points in the convex set $\mathcal{S}$, so are their convex combinations. We assume that $\mathbf{x}' = \theta\mathbf{x}_1 + (1-\theta)\mathbf{x}_2$ and $t' = \theta t_1 + (1-\theta)t_2$. The goal is then to find a $\theta \in [0, 1]$ such that the following equality holds:

$$\alpha\frac{\mathbf{x}_1}{t_1} + (1-\alpha)\frac{\mathbf{x}_2}{t_2} = \frac{\theta\mathbf{x}_1 + (1-\theta)\mathbf{x}_2}{\theta t_1 + (1-\theta)t_2}. \tag{1}$$

It can be verified that (1) holds when

$$\theta = \frac{\alpha t_2}{(1-\alpha)t_1 + \alpha t_2}.$$

Also note that since $\alpha \in [0, 1]$, the θ given by the equation above lies in $[0, 1]$.

Thus, we have proved that given any two points $\mathbf{y}_1, \mathbf{y}_2 \in f(\mathcal{S})$, and any $\alpha \in [0, 1]$, their convex combination $\mathbf{y}$ can be represented as $\mathbf{x}'/t$ where $(\mathbf{x}', t) \in \mathcal{S}$. This means $\mathbf{y} \in f(\mathcal{S})$. By definition, $f(\mathcal{S})$ is convex. ■

Claim 2 *If $\mathcal{S}$ is convex, then $f^{-1}(\mathcal{S})$ is convex where f is a perspective function.*

Proof: First note that $f^{-1}(\mathcal{S})$ can be explicitly written as:

$$f^{-1}(\mathcal{S}) = \left\{(\mathbf{x}, t) : \frac{\mathbf{x}}{t} \in \mathcal{S}, t > 0\right\}.$$

For any $(\mathbf{x}_1, t_1), (\mathbf{x}_2, t_2) \in f^{-1}(\mathcal{S})$, and $\alpha \in [0,1]$ we want to show that

$$\alpha \begin{pmatrix} \mathbf{x}_1 \\ t_1 \end{pmatrix} + (1-\alpha) \begin{pmatrix} \mathbf{x}_2 \\ t_2 \end{pmatrix} \in f^{-1}(S)$$

That is, we want to show that

$$\mathbf{y} = \frac{\alpha \mathbf{x}_1 + (1-\alpha)\mathbf{x}_2}{\alpha t_1 + (1-\alpha) t_2} \in \mathcal{S}.$$

We assume that this $\mathbf{y}$ can be represented as $\mathbf{y} = \theta \mathbf{x}_1/t_1 + (1-\theta)\mathbf{x}_2/t_2$, where $\theta \in [0,1]$. Then the goal is to find $\theta \in [0,1]$ such that the the following equality holds:

$$\frac{\alpha}{\alpha t_1 + (1-\alpha) t_2} \mathbf{x}_1 + \frac{1-\alpha}{\alpha t_1 + (1-\alpha) t_2} \mathbf{x}_2 = \frac{\theta}{t_1} \mathbf{x}_1 + \frac{1-\theta}{t_2} \mathbf{x}_2.$$

Matching the coefficients of $\mathbf{x}_1$ and $\mathbf{x}_2$ in the equation above gives

$$\theta = \frac{\alpha t_1}{(1-\alpha) t_1 + \alpha t_1}.$$

Since $\alpha \in [0,1]$, $\theta \in [0,1]$. Hence $\mathbf{y}$ is a convex combination of $\mathbf{x}_1/t_1$ and $\mathbf{x}_2/t_2$, which are in $\mathcal{S}$. Due to the assumption that $\mathcal{S}$ is convex, it follows that $\mathbf{y} \in \mathcal{S}$, and $f^{-1}(\mathcal{S})$ is a convex set. ■

Definition 2 *The following function is called a linear fractional function:*

$$f(\mathbf{x}) = \frac{\mathbf{A}\mathbf{x} + \mathbf{b}}{\mathbf{c}^\top \mathbf{x} + u},$$

where $\mathbf{x} \in \mathbb{R}^d$, $\mathbf{A} \in \mathbb{R}^{m \times d}$, $\mathbf{b} \in \mathbb{R}^m$, $\mathbf{c} \in \mathbb{R}^d$, *and* $u \in \mathbb{R}$. *The domain of* f *is* $\{\mathbf{x} : \mathbf{c}^\top \mathbf{x} + u > 0\}$.

Claim 3 *If a set* $\mathcal{S}$ *is convex, then* $f(\mathcal{S})$ *is convex, where* f *is a linear fraction function.*

Proof: The intuition behind the proof is that $f(x)$ is a perspective transform of an affine function. Define

$$g(\mathbf{x}) = \begin{pmatrix} \mathbf{A} \\ \mathbf{c}^\top \end{pmatrix} \mathbf{x} + \begin{pmatrix} \mathbf{b} \\ u \end{pmatrix} \in \mathbb{R}^{m+1}.$$

$$f(\mathbf{x}) = \text{perspective}(g(\mathbf{x})) \implies f(\mathcal{S}) = \text{perspective}(\text{g}(\mathcal{S})) \text{ is convex. } ■$$

So far, I have covered the basic concepts and proof techniques for convex sets. I'm now going to introduce another important component in convex optimization: convex functions.

Definition 3 (Convex function) *A function* $f : \mathbb{R}^d \to \mathbb{R}$ *is called a convex function if its domain* $\mathbf{dom} f$ *is a convex set, and for any two points* $\mathbf{x}, \mathbf{y} \in \mathbf{dom} f$ *and any* $\alpha \in [0,1]$, *the following inequality holds:*

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

Definition 4 (Strictly Convex function) *A function $f : \mathbb{R}^d \to \mathbb{R}$ is called a strictly convex function if its domain $\mathbf{dom} f$ is a convex set, and for any two points $\mathbf{x}, \mathbf{y} \in \mathbf{dom} f$ and any $\alpha \in (0, 1)$, the following inequality holds:*

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

While the convex function is typically considered in minimization problem, its counterpart, the concave function, is often considered in maximization problem. A concave function is the negative of a convex function. Its formal definition is as follows.

Definition 5 (Concave function) *A function $f : \mathbb{R}^d \to \mathbb{R}$ is called a concave function if $-f$ is a convex function.*

Definition 6 (Strictly Concave function) *A function $f : \mathbb{R}^d \to \mathbb{R}$ is called a strictly concave function if $-f$ is a strictly convex function.*

Example 1 *The following functions are all convex:*

(1) Affine function: $f(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b}, \mathbf{A} \in \mathbb{R}^{m \times d}, \mathbf{b} \in \mathbb{R}^m$;

(2) Vector norms: $\|\mathbf{x}\|_p = (\sum_{i=1}^d |x_i|^p)^{1/p}$, *where* $p \geq 1$;

(3) Exponential function: $f(x) = e^{\alpha x}$, *where* $\alpha \in \mathbb{R}$;

(4) Powers: $f(\mathbf{x}) = x^a$, *where* $\alpha \geq 1$ *or* $\alpha < 0$, *for* $x > 0$;

(5) Negative entropy: $f(\mathbf{x}) = x \ln(x)$, *for* $x > 0$.