IE 521: Convex Optimization	**Spring 2017, UIUC**

Lecture 5: Convex Functions, Part II – February 6

Instructor: Niao He *Scribe: Shuanglong Wang*

Courtesy warning: These notes do not necessarily cover everything discussed in the class. Please email TA (swang157@illinois.edu) if you find any typos or mistakes.

In this lecture, we cover the following topics

- Characterization of Convex Function
  - Epigraph
  - Level set
  - One dimensional property
  - First order condition
  - Second order condtion
- Continuity of Convex Functions

References: Boyd &Vandenberghe Chapter 2

5.1 Recall

A function $f : \mathbf{R}^n \to \mathbf{R}$ is convex if it satisfies

(i) convexity of domain: $dom(f) = \{x : |f(x)| < +\infty\}$ is convex

(ii) basic convex inequality: if $x, y \in dom(f), \lambda \in [0, 1]$ $f(\lambda x + (1 - \lambda)y) \le \lambda f(x) + (1 - \lambda)f(y)$

Remark (Extended-value function) f can be extended to a function from $\mathbf{R}^n \to \mathbf{R} \bigcup \{+\infty\}$ by setting $f(x) = +\infty$, if $x \notin dom(f)$. Now (ii) can rewritten as $\forall x, y \in \mathbf{R}^n, \lambda \in [0, 1], f(\lambda x + (1 - \lambda)y) \le \lambda f(x) + (1 - \lambda)f(y)$

Remark (General convex inequality) $\forall \lambda_i \ge 0, \sum_{i=1}^m = 1$, we have by induction that

$$f(\sum_{i=1}^m \lambda_i x_i) \le \sum_{i=1}^m \lambda_i f(x_i)$$

This is also known as Jensen's inequality. Let ψ be a finite discrete random variable. Then we always have

$$f(\mathbf{E}[\psi]) \le \mathbf{E}[f(\psi)]$$

5.2 Characterization of Convex Functions

5.2.1 Epigraph

The epigraph of a function is defined as

$$\text{epi}(f) = \left\{(x,t) \in \mathbf{R}^{n+1} : f(x) \leq t\right\}$$

Proposition 5.1 *f is convex on $\mathbf{R}^n$ if and only if its epigraph is a convex set in $\mathbf{R}^{n+1}$.*

Proof:

- ($\Leftarrow$ part) Firstly, $dom(f) = \{x : \exists t, \text{ s.t. } (x,t) \in \text{epi}(f)\}$ is convex. Let $(x_1,t_1),(x_2,t_2) \in epi(f)$, then $(\lambda x_1 + (1-\lambda)x_2, \lambda t_1 + (1-\lambda)t_2) \in \text{epi}(f), \forall \lambda \in [0,1]$. By definition of epigraph, this means $f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda t_1 + (1-\lambda)t_2$. Particularly, setting $t_1 = f(x_1), t_2 = f(x_2)$, we have the basic convex inequality.

- ($\Rightarrow$ part) On the other hand,

$$\begin{aligned} f \text{ convex} &\Rightarrow f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda)f(x_2), \forall \lambda \in [0,1] \\ &\Rightarrow f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda t_1 + (1-\lambda)t_2, \forall (x_1,t_1),(x_2,t_2) \in \text{epi}(f) \\ &\Rightarrow \lambda(x_1,t_1) + (1-\lambda)(x_2,t_2) \in \text{epi}(f) \\ &\Rightarrow \text{epi}(f) \text{is a convex set} \end{aligned}$$

■

5.2.2 Level set

For any $t \in \mathbf{R}$, the level set of f with level t is defined as

$$\text{lev}_t(f) = \{x \in dom(f) : f(x) \leq t\}.$$

Proposition 5.2 *If f is convex, then, every level set is convex.*

Proof: For all $x_1, x_2 \in \text{lev}_t(f)$ and $\lambda \in [0,1]$,

$$f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda)f(x_2) \leq \lambda t + (1-\lambda)t = t$$

i.e. $\lambda x_1 + (1-\lambda)x_2 \in \text{lev}_t(f)$ ■

Remark: The reverse is not true. A function is called *quasi-convex* if its domain and all level sets are convex, this holds if and only if

$$f(\lambda x + (1-\lambda)y) \leq \max\{f(x), f(y)\}, \forall \lambda \in [0,1]$$

5.2.3 One-dimensional Property

Recall that a set is convex if and only if its restriction on any line is convex. Similarly, we have

Proposition 5.3 *f is convex if and only is its restriction on any line is convex. i.e. $\forall x, h \in \mathbf{R}^n, \phi(t) = f(x + th)$ is convex on the axis.*

Proof: ($\Longrightarrow$) Firstly, $dom(\phi) = \{t \in \mathbf{R} : x + th \in dom(f)\}$ is convex. Also, $\forall t_1, t_2 \in \mathbf{R}, \forall \lambda \in [0, 1]$

$$\begin{aligned}\phi(\lambda t_1 + (1-\lambda)t_2) &= f(x + (\lambda t_1 + (1-\lambda)t_2)h) \\ &= f(\lambda(x + t_1 h) + (1-\lambda)(x + t_2 h)) \\ &\le \lambda f(x + t_1 h) + (1-\lambda) f(x_2 + t_2 h) \\ &= \lambda\phi(t_1) + (1-\lambda)\phi(t_2)\end{aligned}$$

Hence, $\phi(t)$ is convex. Proof for the reverse direction is omitted. ■

Remark: Checking convexity in $\mathbf{R}$ boils down to check convexity of one-dimensional function on the axis. (See example in HW1)
Recall from basic calculus that for a univariate function f on (a, b),

1. if f is differentiable, f convex $\Longleftrightarrow$ f' increasing
2. if f is twice-differentiable, f convex $\Longleftrightarrow$ $f" \ge 0$

Similar first and second order conditions hold for general convex functions.

5.2.4 First-order Condition

Proposition 5.4 *Assume f is differentiable, then f is convex if and only if $dom(f)$ is convex and $f(y) \ge f(x) + \bigtriangledown f(x)^T(y - x), \forall x, y$.*

Proof: ($\Leftarrow$ part) Let $z = \lambda x + (1-\lambda)y, \forall \lambda \in [0, 1]$

$$\begin{aligned}f(x) &\ge f(z) + \nabla f(z)^T(x - z) \\ f(y) &\ge f(z) + \nabla f(z)^T(y - z)\end{aligned}$$

Hence, we have

$$\lambda f(x) + (1-\lambda)f(y) \ge f(z) + \nabla f(z)^T(\lambda x + (1-\lambda)y - z) = f(z) = f(\lambda x + (1-\lambda)y).$$

($\Rightarrow$ part) By convexity, we have for all $\lambda \in [0, 1]$,

$$\begin{aligned}& f((1-\lambda)x + \lambda y) \le (1-\lambda)f(x) + \lambda f(y) \\ \Rightarrow\ & \frac{f(x + \lambda(y - x))}{\lambda} \le \frac{f(x)}{\lambda} + f(y) - f(x) \\ \Rightarrow\ & f(y) \ge f(x) + \frac{f(x + \lambda(y - x)) - f(x)}{\lambda}, \forall \lambda \in [0, 1] \\ \Rightarrow\ & f(y) \ge f(x) + \nabla f(x)^T(y - x), \text{ letting } \lambda \to 0\end{aligned}$$

■

Remark: The above proposition implies that we have obtain global underestimate of the entire function based on local information $(f(x), \nabla f(x))$. This is an important feature of convex functions. Moreover, one can see that a convex function can be viewed as the supremum of affine function.

5.2.5 Second-order condition

Proposition 5.5 *Assume f is twice-differentiable, then f is convex if and only if $dom(f)$ is convex and $\nabla^2 f(x) \succeq 0, \forall x \in dom(f)$*

Proof: ($\Leftarrow$ part) $\forall x, h, \phi(t) = f(x + th)$ is convex on the axis

$$\phi"(t) = h^T \nabla^2 f(x + th)h \geq 0$$

particularly, $\phi"(0) = h^T \nabla^2 f(x)h \geq 0, \forall h$ Hence $\nabla^2 f(x) \succeq 0$.

($\Rightarrow$ part) Any one dimensional restriction

$$\phi(t) = f(x + th) \text{ is convex since } \phi"(t) \geq 0$$

Hence f is convex. ■

Example: $f(x) = \frac{1}{2}x^T Qx + b^T x + c$ is convex if and only if $Q \succeq 0$.

5.3 Continuity of Convex Functions

Convex functions are almost everywhere continuous.

Theorem 5.6 *Let $f : \mathbf{R}^n \to \mathbf{R}$ be convex, then f is continuous on $rint(dom(f))$.*

Remark: Note that f needs not to be continuous on $\mathbf{R}^n$ or $dom(f)$. e.g.

$$f(x) = \begin{cases} 1, & x = 0 \\ 0, & x > 0 \\ +\infty, & o.w. \end{cases}$$

Proof: Without loss of generality, let's assume $dim(dom(f)) = n$, $0 \in int(dom(f))$ and $\{x :\| x \|_2 \leq 1\} \subseteq dom(f)$. Let us consider the continuity at point 0. Let $\{x_n\} \to 0$ with $\| x_n \|_2 \leq 1$.

(a) $\limsup_{n\to\infty} f(x_n) \leq f(0)$
This is because $x_k = (1- \| x_k \|_2) \cdot 0 + \| x_k \|_2 \cdot y_k$, where $y_k = \frac{x_k}{\|x_k\|_2} \in dom(f)$. By convexity of f, we have

$$f(x_k) \leq (1- \| x_k \|_2) \cdot f(0) + \| x_k \|_2 \cdot f(y_k)$$

Therefore, $\limsup_{k\to\infty} f(x_k) \leq f(0)$.

(b) $\liminf_{n\to\infty} f(x_n) \geq f(0)$
This is because $0 = \frac{1}{\|x_k\|_2+1} x_k + \frac{\|x_k\|_2}{\|x_k\|_2+1} z_k$, where $z_k = -\frac{x_k}{\|x_k\|_2} \in dom(f)$. By the convexity of f, we have

$$f(0) \leq \frac{1}{\| x_k \|_2 + 1} f(x_k) + \frac{\| x_k \|_2}{\| x_k \|_2 + 1} f(z_k)$$

Therefore, $\liminf_{k\to\infty} f(x_k) \geq f(0)$.

■