2.2 Lipschitz Continuity

In this section, we focus on the Lipschitz continuity of convex functions.
First, we start with some lemmas.

Lemma: Let $\{e_1, ..., e_n\}$ denote the standard basis of $\mathbb{R}^n$. Let $A := \{x \pm \epsilon e_i\}$ Then the following holds:

1. $x + \delta e_i \in \text{conv}(A)$ for $|\delta| \leq \epsilon$
2. $B(x; \epsilon/n) \subset \text{conv}(A)$

Proof. 1. Since $|\delta| \leq \epsilon$, there exists λ such that $\delta = \lambda(-\epsilon) + (1-\lambda)\epsilon$. Then,

$$x + \delta e_i = \lambda(x - \epsilon e_i) + (1-\lambda)(x + \epsilon e_i) \in \text{conv}(A)$$

2. Let $y \in B(x; \epsilon/n)$. Then $y = x + \frac{\epsilon}{n}u$, where $||u|| \leq 1$. Write $u = \sum_{i=1}^n \lambda_i e_i$, then

$$|\lambda_i| \leq \sqrt{\sum_{i=1}^n \lambda_i^2} \leq 1$$

So

$$y = x + \frac{\epsilon}{n}u = x + \frac{\epsilon}{n}\sum_{i=1}^n \lambda_i e_i = \sum_{i=1}^n \frac{1}{n}(x + \epsilon\lambda_i e_i)$$

Since $x + \epsilon\lambda_i e_i \in \text{conv}(A)$, $y \in \text{conv}(A)$. Hence $B(x; \frac{\epsilon}{n}) \subseteq \text{conv}(A)$.

□

Lemma: If a convex function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is bounded above on $B(\overline{x}; \delta)$ for some $\overline{x} \in \text{dom}f$ and $\delta > 0$, then f is bounded on $B(\overline{x}, \delta)$.

Proof. Suppose $f(x) \leq M$ for all $x \in B(\overline{x}, \delta)$. Let $f(\overline{x}) = m$.
Suppose $x \in B(\overline{x}; \delta)$ Let $u := \overline{x} + (\overline{x} - x) = 2\overline{x} - x$. Then $u \in B(\overline{x}, \delta)$. We have

$$m = f(\overline{x}) = f(\frac{x+u}{2}) \leq \frac{1}{2}f(x) + \frac{1}{2}f(u)$$

Therefore, $f(x) \geq 2f(\overline{x}) - f(u) \geq 2m - M$. Hence f is bounded on $B(\overline{x}, \delta)$. □

Theorem: Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be convex with $\overline{x} \in \text{dom}f$. Suppose f is bounded on $B(\overline{x}, \delta)$ for some $\delta > 0$, then f is Lipschitz continuous on $B(\overline{x}; \frac{\delta}{2})$.

Proof. Let $x, y \in B(\overline{x}; \frac{\delta}{2})$ with $x \neq y$. Suppose $f \leq M$ on $B(\overline{x}; \delta)$. Let

$$u := x + \frac{\delta}{2||x-y||}(x-y)$$

then $u \in x + \frac{\delta}{2}B \subset x + \delta B$. Also

$$x = \frac{1}{\alpha+1}u + \frac{\alpha}{\alpha+1}y$$

where $\alpha = \frac{\delta}{2||x-y||}$. Then

$$\begin{aligned} f(x) - f(y) &\leq \frac{1}{\alpha+1} f(u) + \frac{\alpha}{\alpha+1} f(y) - f(y) \\ &= \frac{1}{\alpha+1}(f(u) - f(y)) \leq \frac{2M}{\alpha+1} \\ &= \frac{4M||x-y||}{\delta + 2||x-y||} \leq \frac{4M||x-y||}{\delta} \end{aligned}$$

□

Proposition: A convex function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is locally Lipschitz continuous on $\text{int}(\text{dom}f)$.

Proof. Let $\overline{x} \in \text{int}(\text{dom}f)$ and let $\epsilon > 0$ be such that $\overline{x} \pm \epsilon e_i \in \text{dom}f$ for all i. Let $A := \{\overline{x} \pm \epsilon e_i\}$. Then $B(\overline{x}; \frac{\epsilon}{n}) \subseteq \text{conv}(A)$. Let $M := \max\{f(a)|\ a \in A\}$. Pick $x \in B(\overline{x}; \frac{\epsilon}{n})$, then

$$x = \sum \lambda_i (\overline{x} + \epsilon e_i), \text{ with } \sum \lambda_i = 1$$

Hence

$$f(x) \leq \sum \lambda_i f(\overline{x} + \epsilon e_i) \leq M$$

Then f is bounded above on $B(\overline{x}; \frac{\epsilon}{n})$. Hence, by the previous theorem, f is Lipschitz continuous on $B(\overline{x}; \frac{\epsilon}{2n})$ □

2.3 Conjugate Functions

In the next chapter, we will consider the concept of duality. One notion that is crucial in the theory of duality is the conjugate function.

Definition:(Conjugate function) Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be a function. The *conjugate function* of f is the function $f^* : \mathbb{R}^n \to [-\infty, \infty]$ defined by

$$f^*(y) = \sup_{x \in \mathbb{R}^n} \{\langle x, y \rangle - f(x)\}$$

Note that f^* is convex even if f is not convex.

Examples of conjugate functions

1. $f(x) = ||x||_1$

$$\begin{aligned} f^*(a) &= \sup_{x \in \mathbb{R}^n} \langle x, a \rangle - ||x||_1 \\ &= \sup \sum (a_n x_n - |x_n|) \\ &= \begin{cases} 0 & ||a||_\infty \leq 1 \\ \infty & \text{otherwise} \end{cases} \end{aligned}$$

2. $f(x) = ||x||_\infty$

$$\begin{aligned} f^*(a) &= \sup_{x\in\mathbb{R}^n} \sum a_n x_n - \max_n |x_n| \\ &\le \sup \sum |a_n||x_n| - \max_n |x_n| \\ &\le \max_n |x_n| ||a||_1 - \max_n |x_n| \\ &\le \sup ||x||_\infty (||a||_1 - 1) \\ &= \begin{cases} 0 & ||a||_1 \le 1 \\ \infty & \text{otherwise} \end{cases} \end{aligned}$$

If $||a||_1 \le 1$, $\langle 0, a\rangle - ||0||_\infty = 0$, $f^*(a) \ge 0$ in this case.
If $||a||_1 > 1$, then $\langle x, a\rangle - ||x||_\infty$ is unbounded. Hence

$$f^*(a) = \begin{cases} 0 & ||a||_1 < 1 \\ \infty & \text{otherwise} \end{cases}$$

We can also consider the conjugate of f^* (double conjugate of f). It is given by

$$f^{**}(x) = \sup_{y\in\mathbb{R}^n} \{\langle y, x\rangle - f^*(y)\}$$

It is natural to ask whether $f = f^{**}$. Indeed, this is true under some conditions.

Theorem: Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be a function. Then:

1. $f(x) \ge f^{**}(x)$ for all $x \in \mathbb{R}^n$.
2. If f is closed, proper and convex, then $f(x) = f^{**}(x)$.

Proof. $\underline{1}$ For all x and y, we have

$$f^*(y) \ge \langle x, y\rangle - f(x)$$

So $f(x) \ge \langle x, y\rangle - f^*(y)$ for all x, y. (*)
Therefore, $f(x) \ge \sup\{\langle x, y\rangle - f^*(y)\} = f^{**}(x)$.
$\underline{2}$ By (1), we have epi$f \subseteq$ epif^{**}. We need to show epi$f^{**} \subseteq$ epi$f$.
It suffices to show that $(x, f^{**}(x)) \in$epi$f$. So suppose not.
Since epi$f$ is a closed convex set, $(x, f^{**}(x))$ can be strictly separated from epi$f$.
Hence

$$\langle y, z\rangle + bs < c < \langle y, x\rangle + bf^{**}(x)$$

for some y, b, c, and for all $(z, s) \in$ epi$f$.
We may assume $b \ne 0$ (If not, add $\epsilon(\overline{y}, -1)$ to (y, b) for some $\overline{y} \in$domf^*).
We must have $b < 0$. Since if $b > 0$, we have a contradiction by choosing s large.
Therefore, we further assume $b = -1$. Hence, in particular, we have

$$\langle y, z\rangle - f(z) < c < \langle y, x\rangle - f^{**}(x)$$

Then taking supremum over z, we have

$$f^*(y) + f^{**}(x) < \langle x, y \rangle$$

This is a contradiction to (*). Hence epif^{**} = epi$f$.
Therefore, $f = f^{**}$. □

2.4 Subgradient of Convex Function

In this section, we introduce the crucial concept of subgradient for convex functions. It acts as generalized derivative for nonsmooth functions and has many applications in optimization theory.

Definition:(Subgradient) Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be a convex function and let $\overline{x} \in \text{dom}f$. An element $g \in \mathbb{R}^n$ is called a *subgradient* of f at $\overline{x}$ if

$$f(x) - f(\overline{x}) \geq \langle g, x - \overline{x} \rangle \text{ for all } x \in \mathbb{R}^n$$

The collection of all subgradients of f is denoted by $\partial f(\overline{x})$.

Proposition: Let f be a convex function and let $\overline{x} \in \text{int}(\text{dom}f)$, then $\partial f(\overline{x})$ is nonempty and compact.