Optimization Theory

Lecture 03

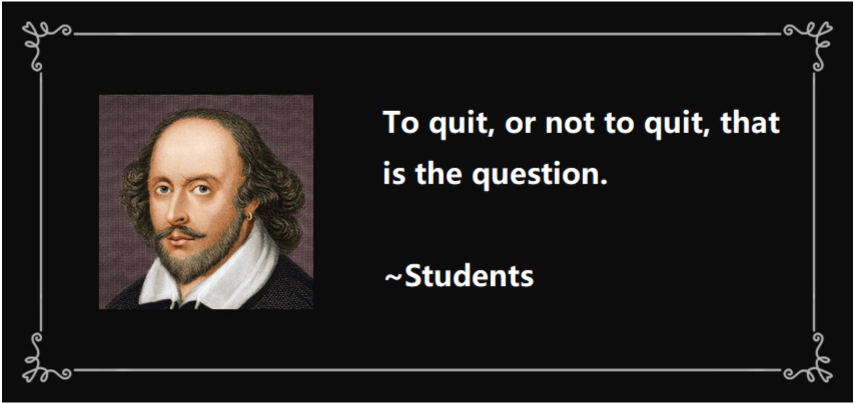
Fudan University

luoluo@fudan.edu.cn

Outline

Outline

Convex Analysis



You can make the decision after the sections of convex analysis.

Convex Set

We say a set $\mathcal{C} \subseteq \mathbb{R}^n$ is convex if for all $\mathbf{x}, \mathbf{y} \in \mathcal{C}$ and $\alpha \in [0, 1]$, it holds that

$$\alpha\mathbf{x} + (1 - \alpha)\mathbf{y} \in \mathcal{C}.$$

Geometrically, a set $\mathcal{C}$ is convex means that the line-segment connecting any two points in $\mathcal{C}$ also belongs to $\mathcal{C}$.

Given any collection of convex sets (finite, countable or uncountable), their intersection is itself a convex set.

Projection

Given a closed and convex set $\mathcal{C} \subseteq \mathbb{R}^n$ and any point $\mathbf{y} \in \mathbb{R}^d$, we define the projection of $\mathbf{y}$ onto $\mathcal{C}$ in Euclidean norm as the point in $\mathcal{C}$ that is closest to $\mathbf{y}$ as

$$\mathrm{proj}_{\mathcal{C}}(\mathbf{y}) = \operatorname*{arg\,min}_{\mathbf{x} \in \mathcal{C}} \|\mathbf{y} - \mathbf{x}\|_2^2 .$$

Projection

Some properties of the projection:

1. The projection $\operatorname{proj}_{\mathcal{C}}(\mathbf{y})$ exists and is uniquely defined.
2. If $\mathbf{y} \notin \mathcal{C}$, then $\mathbf{z} = \operatorname{proj}_{\mathcal{C}}(\mathbf{y})$ lies on the boundary of $\mathcal{C}$ and the hyperplane

$$\{\mathbf{x} : \langle \mathbf{y} - \mathbf{z}, \mathbf{x} - \mathbf{z} \rangle = 0\}$$

separates $\mathbf{y}$ and $\mathcal{C}$ in that they lie on different sides, that is

$$\langle \mathbf{y} - \mathbf{z}, \mathbf{y} - \mathbf{z} \rangle > 0 \quad \text{and} \quad \langle \mathbf{y} - \mathbf{z}, \mathbf{x} - \mathbf{z} \rangle \leq 0$$

for any $\mathbf{x} \in \mathcal{C}$. It implies

$$\|\mathbf{x} - \mathbf{z}\|_2^2 < \|\mathbf{x} - \mathbf{y}\|_2^2$$

for any $\mathbf{x} \in \mathcal{C}$.

Outline

Convex Function

A function $f : \mathcal{C} \to \mathbb{R}$, defined on a convex set $\mathcal{C}$, is convex if it holds

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

for all $\mathbf{x}, \mathbf{y} \in \mathcal{C}$ and $\alpha \in [0, 1]$.

Epigraph

The epigraph of a function $f : \mathcal{C} \to \mathbb{R}$ is defined as the set

$$\operatorname{epi} f \triangleq \{(\mathbf{x}, u) \in \mathcal{C} \times \mathbb{R} : f(\mathbf{x}) \leq u\}.$$

We say a function $f(\mathbf{x})$ is closed if its epigraph is closed.

Theorem

A function $f(\mathbf{x})$ is convex if and only if its epigraph is a convex set.

Extended Arithmetic Operations

We shall define convex function with possibly infinite values, which leads to arithmetic calculations involving $+\infty$ and $-\infty$:

- $-(-\infty) = +\infty$
- $\alpha \pm (+\infty) = (+\infty) \pm \alpha = +\infty$ for $\alpha \in \mathbb{R}$,
- $\alpha \pm (-\infty) = (-\infty) \pm \alpha = -\infty$ for $\alpha \in \mathbb{R}$,
- $\alpha \cdot (\pm\infty) = (\pm\infty) \cdot \alpha = \pm\infty$ for $\alpha \in (0, +\infty)$
- $\alpha \cdot (\pm\infty) = (\pm\infty) \cdot \alpha = \mp\infty$ for $\alpha \in (-\infty, 0)$
- $\alpha/(\pm\infty) = 0$ for $\alpha \in (-\infty, +\infty)$
- $(\pm\infty)/\alpha = \pm\infty$ for $\alpha \in (0, +\infty)$
- $(\pm\infty)/\alpha = \mp\infty$ for $\alpha \in (-\infty, 0)$
- $\inf \emptyset = +\infty, \ \ \sup \emptyset = -\infty$

The extended real number system $\overline{\mathbb{R}}$, defined as

$$[-\infty, +\infty] \qquad \text{or} \qquad \mathbb{R} \cup \{-\infty, +\infty\}.$$

Extended Arithmetic Operations

The expressions

$$(+\infty)-(+\infty), \quad (-\infty)+(+\infty), \quad \frac{+\infty}{-\infty} \text{ and } \frac{-\infty}{+\infty}.$$

are undefined and are avoided.

In the context of convex analysis, we also define

$$0\cdot\infty=\infty\cdot 0=0 \qquad \text{and} \qquad 0\cdot(-\infty)=(-\infty)\cdot 0=0.$$

Proper Convex Function

One may extend a convex function with domain $\mathcal{C} \subset \mathbb{R}^d$ to a proper convex function

$$f_{\mathcal{C}}(\mathbf{x}) = \begin{cases} f(\mathbf{x}), & \text{if } \mathbf{x} \in \mathcal{C}, \\ +\infty, & \text{if } \mathbf{x} \notin \mathcal{C}. \end{cases}$$

We define

$$\operatorname{dom} f \triangleq \{\mathbf{x} : f(\mathbf{x}) < +\infty\}.$$

We say a convex function is proper if its domain is non-empty and its values are all larger than $-\infty$.

We say a function $f(\mathbf{x})$ on $\mathbb{R}^d$ is concave if $-f(\mathbf{x})$ is convex. Linear functions are both convex and concave.

Convex Function

Some properties of convex function:

1. Given any $\mathbf{g} : \mathbb{R}^d \to \mathbb{R}^k$ such that each component $g_j(\mathbf{x})$ is convex, then the set $\mathcal{C} = \{\mathbf{x} : \mathbf{g}(\mathbf{x}) \leq \mathbf{0}\}$ is convex.
2. The supremum over a family of convex functions is convex.
3. The positively weighted sum of convex functions is convex.
4. The partial infimum of a convex function is convex.
5. The composition of convex functions may not preserve convexity.

Indicator Function

Given a closed convex set $\mathcal{C} \in \mathbb{R}^d$, we can define a convex function $\mathbb{1}_{\mathcal{C}}(\mathbf{x})$ on $\mathbb{R}^d$, called the indicator function of $\mathcal{C}$ on $\mathbb{R}^d$, as

$$\mathbb{1}_{\mathcal{C}}(\mathbf{x}) \triangleq \begin{cases} 0, & \text{if } \mathbf{x} \in \mathcal{C}, \\ +\infty, & \text{if } \mathbf{x} \notin \mathcal{C}. \end{cases}$$

We may write $f_{\mathcal{C}}(\mathbf{x}) = f(\mathbf{x}) + \mathbb{1}_{\mathcal{C}}(\mathbf{x})$ and the problem

$$\min_{\mathbf{x} \in \mathcal{C}} f(\mathbf{x})$$

is equivalent to

$$\min_{\mathbf{x} \in \mathbb{R}^d} f_{\mathcal{C}}(\mathbf{x}) = f(\mathbf{x}) + \mathbb{1}_{\mathcal{C}}(\mathbf{x}).$$

Closed Convex Function

We shall focus on closed functions in convex optimization.

1. All convex functions can be made closed by taking the closure of its epigraph.
2. In some pessimistic case, a closed convex function may not be continuous at the boundary of its domain. Consider the function

$$f(x, y) = \begin{cases} \dfrac{x^2}{y}, & (x, y) \neq (0, 0) \\ 0, & (x, y) = (0, 0) \end{cases}$$

   with domain $\{(x, y) : y > 0\} \cup \{(0, 0)\}$.
3. We focus on only consider problems where the optimal solution can be achieved at a point that is continuous.

Outline

Convex Optimization

Why do we love convex optimization?

Theorem

Let $f(\mathbf{x})$ be a convex function defined on a convex set $\mathcal{C}$ and $\mathbf{x}^*$ be a local solution of

$$\min_{\mathbf{x}\in\mathcal{C}} f(\mathbf{x}). \tag{1}$$

That is, there exist some $\delta > 0$ such that any $\hat{\mathbf{x}} \in \mathcal{B}_\delta(\mathbf{x}^*) \cap \mathcal{C}$ holds

$$f(\mathbf{x}^*) \leq f(\hat{\mathbf{x}}).$$

Then the local solution $\mathbf{x}^*$ is a global solution of problem (1).

First-Order Condition

Theorem

If a function f is differentiable on open set $\mathcal{C}$, then it is convex on $\mathcal{C}$ if and only if

$$f(\mathbf{y}) \geq f(\mathbf{x}) + \langle \nabla f(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle$$

hols for any $\mathbf{x}, \mathbf{y} \in \mathcal{C}$.

However, the gradient may not exist in general case.

Outline

Subgradient and Subdifferential

We say a vector $\mathbf{g} \in \mathbb{R}^d$ is a subgradient of a proper convex function $f : \mathbb{R}^d \to \mathbb{R}$ at $\mathbf{x} \in \operatorname{dom} f$ if

$$f(\mathbf{y}) \geq f(\mathbf{x}) + \langle \mathbf{g}, \mathbf{y} - \mathbf{x} \rangle$$

holds for any $\mathbf{y} \in \mathbb{R}^d$.

The set of subgradients at $\mathbf{x} \in \operatorname{dom} f$ is called the subdifferential of f at $\mathbf{x}$, defined as

$$\partial f(\mathbf{x}) \triangleq \left\{ \mathbf{g} \in \mathbb{R}^d : f(\mathbf{y}) \geq f(\mathbf{x}) + \langle \mathbf{g}, \mathbf{y} - \mathbf{x} \rangle \text{ holds for any } \mathbf{y} \in \mathbb{R}^d \right\}.$$

Examples of Subdifferential

1. The subdifferential of $f(x) = |x|$ at 0 is the set

$$\partial f(x) = [-1, 1].$$

   What about the general norm?

2. The subdifferential of an indicator function $\mathbb{1}_{\mathcal{C}}(\mathbf{x})$ is

$$\partial \mathbb{1}_{\mathcal{C}}(\mathbf{x}) = \mathcal{N}_{\mathcal{C}}(\mathbf{x}),$$

   where

$$\mathcal{N}_{\mathcal{C}}(\mathbf{x}) = \left\{\mathbf{g} \in \mathbb{R}^d : \langle \mathbf{g}, \mathbf{y} - \mathbf{x} \rangle \leq 0 \text{ for all } \mathbf{y} \in \mathcal{C}\right\}$$

   is called the normal cone of $\mathcal{C}$ at $\mathbf{x}$.

3. If a convex function f is differentiable at $\mathbf{x}$, then

$$\partial f(\mathbf{x}) = \{\nabla f(\mathbf{x})\}.$$