Optimization Theory

Lecture 04

Fudan University

luoluo@fudan.edu.cn

Outline

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 $\mathbf{x} \in \mathcal{C}$, then
$$\partial f(\mathbf{x}) = \{\nabla f(\mathbf{x})\}.$$

Outline

Subdifferential Calculus

Let f_1 and f_2 be proper convex functions on $\mathbb{R}^d$, then

$$\partial(f_1 + f_2)(\mathbf{x}) \supseteq \partial f_1(\mathbf{x}) + \partial f_2(\mathbf{x}).$$

If the sets $\mathrm{ri}(\mathrm{dom}\, f_1)$ and $\mathrm{ri}(\mathrm{dom}\, f_2)$ have a point in common (overlap sufficiently), we have

$$\partial(f_1 + f_2)(\mathbf{x}) = f_1(\mathbf{x}) + \partial f_2(\mathbf{x}).$$

We define the relative interior $\mathrm{ri}(\mathcal{C})$ for convex $\mathcal{C} \subseteq \mathbb{R}^d$ as

$$\begin{aligned}\mathrm{ri}(\mathcal{C}) = \{\mathbf{z} \in \mathcal{C} :\ & \text{for every } \mathbf{x} \in \mathcal{C} \text{ such that} \\ & \text{there exist a } \mu > 1 \text{ such that } (1-\mu)\mathbf{x} + \mu\mathbf{z} \in \mathcal{C}\}.\end{aligned}$$

It means every line segment in $\mathcal{C}$ having $\mathbf{z}$ as one endpoint can be prolonged beyond $\mathbf{z}$ without leaving $\mathcal{C}$.

Subdifferential Calculus

Nonempty subdifferential and convexity:

1. If any $\mathbf{x} \in \mathrm{dom}\, f$ satisfies $\partial f(\mathbf{x}) \neq \emptyset$, then f is convex.
2. If $f : \mathbb{R}^d \to \mathbb{R}$ is convex and $\mathbf{x}$ belongs to the interior of $\mathrm{dom}\, f$, then $\partial f(\mathbf{x}) \neq \emptyset$.

Theorem (Hyperplane Separation Theorem)

Let $\mathcal{X} \subseteq \mathbb{R}^d$ is a convex set and $\mathbf{x}_0$ belongs to its boundary. Then, there exists a nonzero vector $\mathbf{w} \in \mathbb{R}^d$ such that

$$\langle \mathbf{w}, \mathbf{x} \rangle \leq \langle \mathbf{w}, \mathbf{x}_0 \rangle.$$

Subdifferential Calculus

The subgradient of a convex function may not exist at a boundary point of the domain.

As an example, consider the function

$$f(x) = -\sqrt{x}$$

defined on $[0, +\infty)$, where we have $\partial f(0) = \emptyset$.

Given matrix $\mathbf{A} \in \mathbb{R}^{d \times m}$ and vector $\mathbf{b} \in \mathbb{R}^d$, define

$$h(\mathbf{x}) = f(\mathbf{A}\mathbf{x} + \mathbf{b}),$$

where f is a proper convex on $\mathbb{R}^d$.

Then the function $h(\mathbf{x})$ is convex and

$$\partial h(\mathbf{x}) \supseteq \mathbf{A}^\top \partial f(\mathbf{A}\mathbf{x} + \mathbf{b}).$$

If the range of $\mathbf{A}$ contains a point of $\mathrm{ri}(\mathrm{dom}\, h)$, then

$$\partial h(\mathbf{x}) = \mathbf{A}^\top \partial f(\mathbf{A}\mathbf{x} + \mathbf{b}).$$