Optimization Theory

Lecture 03

Fudan University

luoluo@fudan.edu.cn

Outline

1 Second-Order Characterization

Outline

1 Second-Order Characterization

Second-Order Characterization

Theorem (Smoothness and Convexity)

Let $f(\cdot)$ be a twice differentiable function defined on $\mathbb{R}^d$

1. *It is L-smooth if and only if $-L\mathbf{I} \preceq \nabla^2 f(\mathbf{x}) \preceq L\mathbf{I}$ for all $\mathbf{x} \in \mathbb{R}^d$.*
2. *It is convex if and only if $\nabla^2 f(\mathbf{x}) \succeq \mathbf{0}$ for all $\mathbf{x} \in \mathbb{R}^d$.*
3. *It is μ-strongly-convex if and only if $\nabla^2 f(\mathbf{x}) \succeq \mu\mathbf{I}$ for all $\mathbf{x} \in \mathbb{R}^d$.*

We also say $f(\cdot)$ is ℓ-weakly convex if

$$\nabla^2 f(\mathbf{x}) \succeq -\ell\mathbf{I}.$$

for some $\ell > 0$.

Second-Order Characterization

Some examples:

1. For unconstrained quadratic problem

$$\min_{\mathbf{x}\in\mathbb{R}^d} f(\mathbf{x}) \triangleq \frac{1}{2}\mathbf{x}^\top \mathbf{A}\mathbf{x} - \mathbf{b}^\top \mathbf{x},$$

   where $\mathbf{A} \in \mathbb{R}^{d\times d}$. We have

$$\nabla^2 f(\mathbf{x}) = \mathbf{A}.$$

2. For regularized generalized linear model

$$\min_{\mathbf{x}\in\mathbb{R}^d} f(\mathbf{x}) \triangleq \frac{1}{n}\sum_{i=1}^{n} \phi_i(\mathbf{a}^\top \mathbf{x}) + \frac{\lambda}{2}\|\mathbf{x}\|_2^2.$$

   where $\phi_i : \mathbb{R}^d \to \mathbb{R}$ is twice differentiable. We have

$$\nabla f(\mathbf{x}) = \frac{1}{n}\sum_{i=1}^{n} \phi_i'(\mathbf{a}_i^\top \mathbf{x})\mathbf{a} + \lambda\mathbf{x} \quad \text{and} \quad \nabla^2 f(\mathbf{x}) = \frac{1}{n}\sum_{i=1}^{n} \phi_i''(\mathbf{a}_i^\top \mathbf{x})\mathbf{a}_i\mathbf{a}_i^\top + \lambda\mathbf{I}.$$