Optimization Theory

Lecture 05

Fudan University

luoluo@fudan.edu.cn

Outline

Outline

Optimal Condition

Theorem

Consider proper convex function f and convex set $\mathcal{C}$. The point $\mathbf{x}^* \in \mathcal{C}$ is a solution of convex optimization problem

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

if and only if

$$\mathbf{0} \in \partial(f(\mathbf{x}^*) + \mathbb{1}_{\mathcal{C}}(\mathbf{x}^*)).$$

The point $\mathbf{x}^*$ is an optimal solution of the problem if there exists a subgradient $\mathbf{g}^* \in \partial f(\mathbf{x}^*)$ such that for all $\mathbf{y} \in \mathcal{C}$ satisfies

$$\langle \mathbf{g}^*, \mathbf{y} - \mathbf{x}^* \rangle \geq 0.$$

In particular, the point $\mathbf{x}^*$ is the solution of the problem in unconstrained case if

$$\mathbf{0} \in \partial f(\mathbf{x}^*).$$

Outline

Regularity Conditions

The following regularity conditions are useful in the convergence analysis of convex optimization problems.

1. We say that a function $f : \mathcal{C} \to \mathbb{R}$ is G-Lipschitz continuous if for all $\mathbf{x}, \mathbf{y} \in \mathcal{C}$, we have

$$|f(\mathbf{x}) - f(\mathbf{y})| \leq G \|\mathbf{x} - \mathbf{y}\|_2 .$$

2. We say a differentiable function $f : \mathbb{R}^d \to \mathbb{R}$ is L-smooth if it has L-Lipschitz continuous gradient. That is, for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$, we have

$$\|\nabla f(\mathbf{x}) - \nabla f(\mathbf{y})\|_2 \leq L \|\mathbf{x} - \mathbf{y}\|_2 .$$

3. If the function

$$g(\mathbf{x}) = f(\mathbf{x}) - \frac{\mu}{2} \|\mathbf{x}\|_2^2$$

is convex for some $\mu > 0$, we say f is μ-strongly convex.

Strong Convexity

Theorem

The function $f : \mathcal{C} \to \mathbb{R}$ defined on convex set $\mathcal{C}$ is μ-strongly-convex if and only if

$$f(\alpha \mathbf{x} + (1-\alpha)\mathbf{y}) \leq \alpha f(\mathbf{x}) + (1-\alpha) f(\mathbf{y}) - \frac{\mu\alpha(1-\alpha)}{2} \|\mathbf{x} - \mathbf{y}\|_2^2$$

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

Theorem

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

$$f(\mathbf{y}) \geq f(\mathbf{x}) + \langle \nabla f(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle + \frac{\mu}{2} \|\mathbf{y} - \mathbf{x}\|_2^2$$

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

Strong Convexity

If there exists some

$$\mathbf{x}^* = \arg\min_{\mathbf{x}\in\mathcal{C}} f(\mathbf{x}),$$

then it is the unique minimizer.

Moreover, the solution is stable such that any approximate solution $\hat{\mathbf{x}}$ satisfying

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

leads to

$$\|\mathbf{x}^* - \hat{\mathbf{x}}\|_2^2 \leq \frac{2\epsilon}{\mu}.$$

Lipschitz Continuity and Smoothness

Theorem

A convex function f is G-Lipschitz continuous on $(\mathrm{dom}\, f)^\circ$ if and only if

$$\|\mathbf{g}\|_2 \leq G$$

for all $\mathbf{g} \in \partial f(\mathbf{x})$ and $\mathbf{x} \in (\mathrm{dom}\, f)^\circ$.

Theorem

A function $f : \mathbb{R}^d \to \mathbb{R}$ is L-smooth (possibly nonconvex), then it holds

$$|f(\mathbf{y}) - f(\mathbf{x}) - \langle \nabla f(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle| \leq \frac{L}{2} \|\mathbf{y} - \mathbf{x}\|_2^2$$

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

Smoothness and Convexity

Theorem

A function $f : \mathbb{R}^d \to \mathbb{R}$ is convex and L-smooth, then we have

1. $0 \leq f(\mathbf{y}) - f(\mathbf{x}) - \langle \nabla f(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle \leq \frac{L}{2} \|\mathbf{y} - \mathbf{x}\|_2^2$
2. $f(\mathbf{x}) + \langle \nabla f(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle + \frac{1}{2L} \|\nabla f(\mathbf{y}) - \nabla(\mathbf{x})\|_2^2 \leq f(\mathbf{y})$
3. $\frac{1}{L} \|\nabla f(\mathbf{y}) - \nabla(\mathbf{x})\|_2^2 \leq \langle \nabla f(\mathbf{x}) - \nabla f(\mathbf{y}), \mathbf{x} - \mathbf{y} \rangle$

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