Optimization Methods (CS1.404), Spring 2024
Lecture 22

Naresh Manwani

Machine Learning Lab, IIIT-H

April 15th, 2024

Strong Duality Theorem

Theorem

Consider the optimization problem

$$\min f(\mathbf{x})$$
$$s.t.\ h_j(\mathbf{x}) \leq 0,\ j = 1 \ldots l$$

where f, h_j, $j = 1 \ldots l$ are convex functions over $\mathbb{R}^n$. Suppose that there exists $\hat{\mathbf{x}}$ such that $h(\hat{\mathbf{x}}) < 0$, $j = 1 \ldots l$ and the problem has finite optimal value. Then the optimal value d^* of the dual problem $g(\boldsymbol{\lambda}) = \min_{\mathbf{x}} \mathcal{L}(\mathbf{x}, \boldsymbol{\lambda})$ is attained and is same as optimal value p^* of the primal problem.

Example: Convex Optimization Problem But Slater's Condition not Satisfied

- Consider the problem: $\min_{x,y>0}\ e^{-x}, \quad s.t.\ \frac{x^2}{y} \leq 0$ with variable (x, y) and domain $D = \{(x, y) \mid y > 0\}$.
- We have $p^* = 1$.
- The Lagrangian is $\mathcal{L}(x, y, \lambda) = e^{-x} + \lambda\frac{x^2}{y}$ and the dual function is

$$g(\lambda) = \min_{x,y>0}\ e^{-x} + \lambda\frac{x^2}{y} = \begin{cases} 0, & \lambda \geq 0 \\ -\infty, & \lambda < 0 \end{cases}$$

- So, we can write the dual problem as $d^* = \max_{\lambda}\ 0\ :\ \lambda \geq 0$ with optimal value $d^* = 0$.
- The optimal duality gap is $p^* - d^* = 1$. In this problem, Slater's condition is not satisfied, since $x = 0$ for any feasible pair (x, y).

Complimentary Slackness Conditions

Theorem

Consider the optimization problem

$$\begin{aligned} &\min\ f(\mathbf{x}) \\ &s.t.\ h_j(\mathbf{x}) \leq 0,\ j = 1 \dots l \end{aligned}$$

where $f,\ h_j,\ j = 1 \dots l$ are convex functions over $\mathbb{R}^n$. Assume that strong duality holds. If $\mathbf{x}^*$ and $\boldsymbol{\lambda}^*$ are optimal solutions of the primal and dual problems respectively, then

$$\begin{aligned} \mathbf{x}^* &\in \arg\min\ \mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}^*) \\ \lambda_j^* h_j(\mathbf{x}^*) &= 0,\ j = 1 \dots l \end{aligned}$$

Optimization Over a Convex Set

- Consider optimization problem $(P): \ \min f(\mathbf{x}) \ \ s.t. \ \mathbf{x} \in C$, where f is continuously differentiable function and C is closed-convex set.

Definition: Stationary Points of Constrained Problem

Let f be a continuously differentiable function over a closed convex set C. Then $\mathbf{x}^*$ is called a **stationary point** of (P) if $\nabla f(\mathbf{x}^*)^\top (\mathbf{x} - \mathbf{x}^*) \geq 0, \ \forall \mathbf{x} \in C$.

Theorem: Stationarity as a Necessary Optimality Condition

Let f be a continuously differentiable function over a closed convex set C, and let $\mathbf{x}^*$ be a local minimum of (P). Then, $\mathbf{x}^*$ is a stationary point of (P).

Examples

Feasible Set C	Explicit Stationary Condition
$\mathbb{R}^n$	$\nabla f(\mathbf{x}^*) = \mathbf{0}$
$\mathbb{R}^n_+$	$\frac{\partial f}{\partial x_i}(\mathbf{x}^*) \begin{cases} = 0, & x_i^* > 0 \\ \geq 0, & x_i^* = 0 \end{cases}$
$\{\mathbf{x} \in \mathbb{R}^n \ : \ \sum_{i=1}^n x_i = 1\}$	$\frac{\partial f}{\partial x_1}(\mathbf{x}^*) = \ldots = \frac{\partial f}{\partial x_n}(\mathbf{x}^*)$
$B[0,1]$	$\nabla f(\mathbf{x}^*) = \mathbf{0}$ or $\|\mathbf{x}^*\| = 1$ and $\exists \lambda \leq 0 \ : \ \nabla f(\mathbf{x}^*) = \lambda \mathbf{x}^*$

Stationarity in Convex Problems

Theorem

Let f be a continuously differentiable convex function over a closed and convex set $C \subseteq \mathbb{R}^n$. Then $\mathbf{x}^*$ is a stationary point of

$$(P) \quad \min \; f(\mathbf{x})$$
$$s.t. \; \mathbf{x} \in C$$

if and only if $\mathbf{x}^*$ is an optimal solution of (P).

Orthogonal Projection

Projection Theorem

Let C be a closed convex set and $\mathbf{x} \in \mathbb{R}^n$. Let $P_C(\mathbf{x})$ denote the orthogonal projection of $\mathbf{x}$ on the set C. Then $\mathbf{z} = P_C(\mathbf{x})$ if an only if $\mathbf{z} \in C$ and

$$(\mathbf{x} - \mathbf{z})^\top (\mathbf{y} - \mathbf{z}) \leq 0$$

for any $\mathbf{y} \in C$.

- Geometrically, it states that for a given closed and convex set C, $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{y} \in C$, the angle between $\mathbf{x} - P_C(\mathbf{x})$ and $\mathbf{y} - P_C(\mathbf{x})$ is greater than or equal to 90 degrees.