Introduction to Optimization

Jong-June Jeon

Department of Statistics,
University of Seoul

Sep, 2023

과목 소개

- 과목명: 최적화, 빅데이터최적화 방법론
- 강의실강의시간 : 33-607 / (목) 18시-20시 20분
- 평가: 중간고사(40%), 기말고사(40%), 과제(20%), 출석 (공지사항 참조).
    - 중간 혹은 기말고사 미응시인 경우 F.
- 교재: Convex optimization, Stenphen Boyd

교수 소개

- 이름: 전종준
- 연구실: 33-712 (미래관)
- 관련연구목록: 최적화 방법론 개발
  - Learning Multiple Quantiles with Neural Networks (2021)
  - Primal path algorithm for compositional data analysis (2020)
  - The sparse Luce model (2018)
  - Adaptive stochastic gradient method under two mixing heterogeneous models(2017)
  - Abrupt change point detection of annual maximum precipitation using fused lasso (2016)
- 현재 진행 연구
  - Optimization for sparse Ranking models

Keywords

- Optimization problem
- Objective function, variable, constraints, feasible set, solution
- Unconstrained optimization problem, constrained optimization problem
- ...

What is the optimization problem? Decision-making under constraints.

Let $f : \mathbb{R}^p \mapsto \mathbb{R}$ and $X \subset \mathbb{R}^p$. The optimization problem is typically given by

$$\begin{aligned} \min_{x} \quad & f(x) \\ \text{st.} \quad & x \in X. \end{aligned}$$

Here we denote f as $f : x \mapsto f(x)$. We call x a variable, f an objective function, X a feasible set, and $x \in X$ the constraint.

Example 1 (Regression problem)

Let $(y_i, x_i) \in \mathbb{R} \times \mathbb{R}^p$ for $i = 1, \cdots, n$ be response-predictor pairs (constants) and $\beta \in \mathbb{R}^p$ be variable.

$$L : \beta \mapsto \frac{1}{n} \sum_{i=1}^{n} (y_i - x_i^\top \beta)^2.$$

The solution of the following optimization problem is obtained by solving

$$\begin{aligned} \min_{\beta} \quad & L(\beta) \\ \text{st.} \quad & \beta \in \mathbb{R}^p. \end{aligned}$$

The solution is called the OLS (ordinary least squares) estimator. Since the feasible set is the whole set of $\mathbb{R}^p$, the optimization problem is called an unconstrained optimization problem.

Example 2 (Logistic regression problem)

Let $(y_i, x_i) \in \{0,1\} \times \mathbb{R}^p$ be given response-predictor pairs and $\beta \in \mathbb{R}^p$.

$$L : \beta \mapsto -\frac{1}{n} \sum_{i=1}^{n} \left(y_i x_i^\top \beta - \log(1 + \exp(x_i^\top \beta)) \right)$$

The MLE of the logistic regression model is the solution of the following optimization problem:

$$\begin{aligned} \min_{\beta} \quad & L(\beta) \\ \text{st.} \quad & \beta \in \mathbb{R}^p. \end{aligned}$$

Example 3 (MLE for a general case)

Let $y|x \sim f(y|x;\theta)$ be the conditional density function of y for a given x, where $\theta \in \Theta \subset \mathbb{R}^p$ is the parameter of the density function. Let (y_i, x_i) be a random sample from $f(y|x;\theta)$. In our optimization problem, (y_i, x_i)s are constants, and θ is a variable.

$$L : \theta \mapsto -\sum_{i=1}^{n} \log f(y_i|x_i;\theta)$$

The MLE of the unknown θ is given by

$$\begin{array}{ll} \min\limits_{\theta} & L(\theta) \\ \text{st.} & \theta \in \Theta. \end{array}$$

Example 4 (L_1-Compressed Sensing)
Let $Y \in \mathbb{R}^m$ be an original signal vector and $X \in \mathbb{R}^{m \times n}$ be a filter matrix and $\beta \in \mathbb{R}^n$ the sparse signal.

$$L : \beta \mapsto \|\beta\|_1$$

The L_1-Compressed Sensing is the solution to the following problem:

$$\begin{aligned} \min_{\beta} \quad & L(\beta) \\ \text{st.} \quad & \beta \in \{\beta : Y = X\beta\} \end{aligned}$$

Example 5 (Other machine learning models: Empirical Risk Minimization (ERM) Framework)

- Spatial Regression problem
- Nonnegative matrix factorization
- Learning a Deep Neural Network model
- A lot of problems we can solve!

Why did I study optimization methods?

- Naive approach: Let f be our objective function and $x \in X$ be constraint. Then investigate $f(x)$ for each $x \in X$ and choose the minimum value of $f(x)$.
- Unfortunately, this method is intractable in practice. Consider $x \in \mathbb{R}^{20}$.

Is it tough to learn and practice?

- If $X = \mathbb{R}^p$ (no constraint), it is surprisingly easy.
- If $X \subsetneq \mathbb{R}^p$, it depends on the problem.

We will deal with

(1) Unconstrained convex optimization problem.

(2) Linear constrained convex optimization problem.

(3) To solve (2) easier, we will learn general optimization methods.

Contents

Let's start our journey to OPTIMIZATION!