Introduction

Dabeen Lee

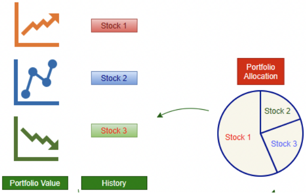
Industrial and Systems Engineering, KAIST

IE 539: Convex Optimization

August 28, 2023

What is “Optimization”?

Portfolio optimization

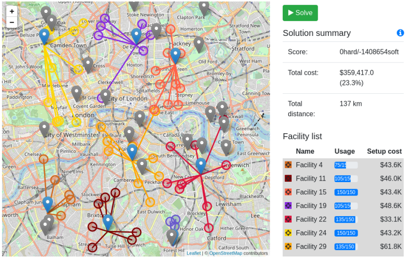


- d financial assets (stocks, bonds, etc).
- Asset i that has return μ_i.
- σ_{ij} is the covariance of assets i and j.
- We allocate x_i fraction of our budget to asset i.

Goal: find a portfolio (allocation) maximizing return while minimizing risk (measured as a function of the covariance).

What is “Optimization”?

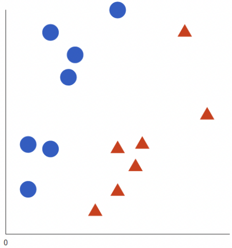
Facility location



Goal: build “fire stations” covering all households while minimizing the longest distance to a household.

What is "Optimization"?

Support vector machine



- n data $(x_1, y_1), \ldots, (x_n, y_n)$ where $y_i \in \{-1, 1\}$ are labels.
- We want to find a separating hyperplane

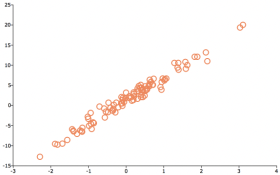
$$w^\top x = b$$

to classify data with $+1$ and data with -1.

Goal: find a separating hyperplane $w^\top x = b$ with the "gap" $(1/\|w\|)$ being maximized.

What is “Optimization”?

Linear regression



- n data points $(x_1, y_1), \ldots, (x_n, y_n)$.
- We want to find a linear rule

$$y = \beta^\top x$$

that best represents the relationship between x and y.

Goal: find β minimizing

$$\underbrace{\frac{1}{n}\sum_{i=1}^{n}(y_i - \beta^\top x_i)^2}_{\text{mean squared error}} + \underbrace{R(\beta)}_{\text{regularization}} \quad .$$

What is "Optimization"?

Policy optimization for Markov decision process

- Compute an optimal policy π maximizing the value function:

$$\max_{\pi} \quad \mathbb{E}_{s\sim\rho}\left[V_r^{\pi}(s)\right]$$

where

- ρ is the initial distribution of states,
- r is the reward function,
- $V_r^{\pi}(s) = \mathbb{E}\left[\sum_{t=0}^{\infty} \gamma^t r(s_t, a_t) \mid s_0 = s, \pi\right]$

What is "Optimization"?

Policy optimization for Markov decision process

- Compute an optimal policy π maximizing the value function:

$$\max_{\pi} \quad \mathbb{E}_{s \sim \rho} \left[V_r^{\pi}(s) \right]$$

where

- ρ is the initial distribution of states,
- r is the reward function,
- $V_r^{\pi}(s) = \mathbb{E} \left[\sum_{t=0}^{\infty} \gamma^t r(s_t, a_t) \mid s_0 = s, \pi \right]$

Who is Dabeen?

Office: E2-2 #2109

Email: dabeenl@kaist.ac.kr

Office hours: Tuesday 2:00 - 3:00 pm

Research interests:

- Optimization Theory (discrete, continuous, stochastic, online).
- Optimization for Machine Learning.
- Algorithm Design for Operations Research (resource allocation, scheduling, dynamic routing).

About this course

The following is a tentative list of topics covered in this course.

Theory	Algorithms	Applications
Convex Analysis (sets, functions, operations)	Gradient Descent (GD)	Machine Learning (SVM, LASSO, Ridge Regression, Policy Gradient, etc.)
Optimality Conditions	Proximal, Projected, Online, Stochastic GD	Statistics (Uncertainty Quantification, Inverse Covariance Selection)
Semidefinite Programming	Frank-Wolfe	Operations Research (Advertisement Allocation, Facility Location, Portfolio Optimization)
Quadratic Programming	Proximal Point Algorithm and Augmented Lagrangian Method	
	Operator Splitting and ADMM	
	Newton's method and Quasi Newton methods	

- Many more applications will be discussed on the way.
- We might also cover other algorithms such as Mirror descent, and Interior Point Methods.

Logistics

Class times: Monday and Wednesday 4:00 - 5:15 pm.

Assessment:

- 5 assignments (50%)
- Course project (20%)
- Take-home final (30%)

No attendance check (but be responsible!)

Assignment

- Being comfortable with making mathematical arguments, writing proofs and programming is required throughout this course.
- **Typesetting in LaTeX is required for any submission.**
- Easiest option: Overleaf (https://www.overleaf.com)

```latex
\documentclass{article}
\usepackage{graphicx} % Required for inserting images

\title{IE331 Assignment 1}
\author{Dabeen Lee}
\date{\today}

\usepackage{amsmath,amssymb}

\begin{document}

\maketitle

\begin{enumerate}
    \item My answer to question 1 is ...

    $$\text{epi}(f) = \left\{(x,t)\in\mathbb{R}^d\times\mathbb{R}:\ \exists y\in\mathbb{R}^p\text{ s.t. } Ax + Dy + h t \leq r \right\}$$

    \item \begin{enumerate}
        \item My answer to question 2(a) is ...

        \begin{align*}
    \min\quad & f(x)\\
    \text{s.t.}\quad & g_i(x)\leq b_i,\quad i\in[m],\\
    &x\in\mathbb{R}^d.
\end{align*}
        \item My answer to question 2(b) is ...

            $$d(x)=\sum_{i\in[n]}\left\|x-v^i\right\|_\infty.$$
        \item My answer to question 2(c) is ...

            $$d(x)=\sum_{i\in[n]}\left\|x-v^i\right\|_1.$$
        \item My answer to question 2(d) is ...
    \end{enumerate}
\end{enumerate}

\end{document}
```

IE331 Assignment 1

Dabeen Lee

March 14, 2023

1. My answer to question 1 is ...

$$\text{epi}(f) = \left\{(x,t)\in\mathbb{R}^d\times\mathbb{R}:\ \exists y\in\mathbb{R}^p\text{ s.t. } Ax + Dy + h t \leq r \right\}$$

2. (a) My answer to question 2(a) is ...

$$\begin{aligned} \min\quad & f(x)\\ \text{s.t.}\quad & g_i(x)\leq b_i,\quad i\in[m],\\ &x\in\mathbb{R}^d. \end{aligned}$$

   (b) My answer to question 2(b) is ...

$$d(x)=\sum_{i\in[n]}\left\|x-v^i\right\|_\infty.$$

   (c) My answer to question 2(c) is ...

$$d(x)=\sum_{i\in[n]}\left\|x-v^i\right\|_1.$$

   (d) My answer to question 2(d) is ...

Project

There are two options.

1. Review of a research paper related to convex optimization.

   - Choose a paper published in a journal or announced at a coference.

2. Formulation and implementation of algorithms or methods for certain optimization problem.

   - Numerical implementation is required.

Submit (1) a proposal and (2) a final report.

Objectives

We formulate a decision-making problem as an optimization model

$$P \; : \; \min_{x \in D} f(x).$$

Then

- We have to study the structure of the problem, f and D.
  - Is P convex? a linear program (LP)? a quadratic program (QP)? a semidefinite program (SDP)?
  - Is f smooth? strongly convex? both?
  - Is D convex? an affine subspace?
- We have to figure out and test candidate algorithms for solving P.
  - Gradient Descent, simply? Proximal Gradient Descent? Newton's method?

For this task, we need comprehensive knowledge in convex optimization.

Later, this knowledge will help you create a new optimization problem.

Example: dual methods

Consider

$$\begin{aligned} \min \quad & f(x) + g(y) \\ \text{s.t.} \quad & Ax + By = c \end{aligned}$$

where f, g are convex and A, B, c are matrices of appropriate dimension.

How do we solve the problem?

- If f and g are both strongly convex, then Gradient Ascent in the dual.
- If only f is strongly convex while g has an easy Prox, then Proximal Gradient in the dual.
- If neither f nor g is strongly convex, then Proximal Point Algorithm in the dual or ADMM.