Introduction to this course

Dabeen Lee

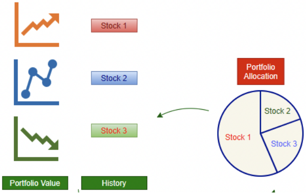
Industrial and Systems Engineering, KAIST

IE 539: Convex Optimization

August 30, 2022

What is "Optimization"?

Portfolio optimization



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

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

What is “Optimization”?

Facility location

▶ Solve

Solution summary

Score:	0hard/-1408654soft
Total cost:	$359,417.0 (23.3%)
Total distance:	137 km

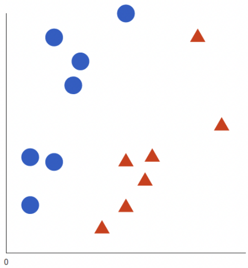
Facility list

Name	Usage	Setup cost
Facility 4	75/15	$43.6K
Facility 11	105/15	$46.0K
Facility 15	150/150	$43.4K
Facility 19	105/15	$48.6K
Facility 22	135/150	$33.1K
Facility 24	150/150	$43.2K
Facility 29	135/150	$61.8K

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

What is "Optimization"?

Support vector machine



Given 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

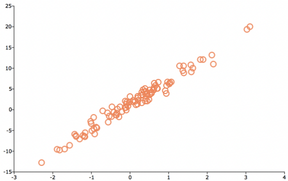
$$w^\top x = b$$

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



Based on 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: we want to find β minimizing the "mean squared error", given by

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

Who is Dabeen?

The door is open to anyone!

At KAIST since July 2022...

Office: E2-2 #2109

Email: dabeenl@kaist.ac.kr

Office hours: Wed 2:00 - 4:00 pm

Research interests:

- Optimization (mainly discrete and stochastic, but some works on continuous),
- Algorithms (for combinatorial and continuous problems),
- + Machine learning, Quantum computing.

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, etc.)
Optimality Conditions	Proximal and Projected GD	Statistics (Uncertainty Quantification, Inverse Covariance Selection)
Semidefinite Programming	Mirror Descent	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 Online GD, Stochastic GD, Frank-Wolfe, and Interior Point Methods.

Logistics

Class times: Tuesday and Thursday 4:00 - 5:30 pm.

Assessment (typesetting in Latex is required for all submissions):

- 6-7 biweekly assignments (50%)
- Course project (20%)
- Take-home final (30%)

Assignments: Being comfortable with making mathematical arguments, writing proofs and programming is required throughout this course.

Project: (1) Choose a problem (possibly from your own reseaarch) that admits a (non)convex optimization formulation, (2) Propose solution methods, and (3) Test candidate algorithms. (Details will be announced soon)

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

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.