# ELEG 667
# Convex Optimization

Xiugang Wu

Dept of Electrical & Computer Engineering

https://www.eecis.udel.edu/~xwu

Aug 27, 2019

# Agenda

- What is (convex) optimization and why is it important?
- What can we expect to learn from this course?
- Logistics

# What is convex optimization and why is it important?

# Mathematical Optimization

Optimization problem:

$$\begin{aligned} \text{minimize} \ & f_0(x) \\ \text{subject to} \ & f_i(x) \le b_i, \ i = 1, 2, \ldots, m \end{aligned}$$

- $x = (x_1, x_2, \ldots, x_n)$: optimization variables
- $f_0 : \mathbf{R}^n \to \mathbf{R}$: objective function
- $f_i : \mathbf{R}^n \to \mathbf{R}, i = 1, \ldots, m$: constraint functions

(optimal) solution $x^*$ has smallest value of $f_0$ among all vectors that satisfy the constraints

# Optimization in Machine Learning

The common theme of ML is a prediction problem:

- Assume that $(X, Y) \sim P$, where $X \in \mathcal{X}$ is called the feature and $Y \in \mathcal{Y}$ is called the label (or target)

- Design a predictor $h$ that takes $X$ as input and outputs $\hat{Y} = h(X)$


- Performance measure: The risk associated with predictor $h$ under distribution $P$ is defined as $L(h, P) = \mathrm{E}_P[\ell(Y, h(X))]$, where $\ell : \mathcal{Y} \times \hat{\mathcal{Y}} \to \mathbf{R}_+$ is referred to as the loss function

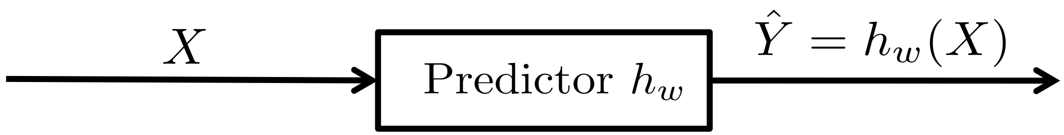
- Objective: minimize $L(h, P)$ over all possible $h$

# Optimization in Machine Learning

Machine Learning—a data-driven approach

Learning:


Prediction:




Empirical Risk Minimization:

$$\text{minimize} \quad \frac{1}{k}\sum_{i=1}^{k} \ell(y^{(i)}, h_w(x^{(i)}))$$

$$\text{subject to} \quad w \in \mathcal{H}$$

- the objective function is called the empirical risk and can be written as $L(h_w, P_k)$, where $P_k$ denotes the empirical distribution
- $w \in \mathcal{H}$ is the constraint for $w$; e.g., $\|w\| \leq B$

# Solving Optimization Problems

General optimization problem

- can be very difficult to solve
- methods involve some compromise, e.g., very long computation time, or not always finding the solution

Exceptions: certain problem classes can be solved efficiently and reliably

- least-squares problems
- convex optimization problems

# Least-Squares

$$\text{minimize } \|Ax - b\|_2^2$$

solving least-squares problems

- analytical solution: $x^* = (A^T A)^{-1} A^T b$
- reliable and efficient algorithms and software
- computation time proportional to $n^2 k$ ($A \in \mathbf{R}^{k \times n}$); less if structured
- a mature technology

using least-squares

- least-squares problems are easy to recognize
- a few standard techniques increase flexibility (e.g., including weights, adding regularization terms)

# Convex Optimization Problems

$$\begin{aligned} &\text{minimize } f_0(x) \\ &\text{subject to } f_i(x) \le b_i, \ \ i = 1, 2, \ldots, m \end{aligned}$$

- objective and constraint functions are convex:

$$f_i(\theta x + (1-\theta)y) \le \theta f_i(x) + (1-\theta) f_i(y)$$

for all $x, y$ and $\theta \in [0, 1]$

- includes least-squares problem as special case

# Convex Optimization Problems

solving convex optimization problems

- no analytical solution
- reliable and efficient algorithms
- polynomial computation time
- almost a technology

using convex optimization

- often difficult to recognize
- many tricks for transforming problems into convex form
- surprisingly many problems can be solved via convex optimization

# (More) Reasons to Learn Convex Optimization

- Different algorithms can perform better or worse on different problems

- Analyzing via the optimization approach can add insight into the original underlying statistical problem (e.g., recall the waterfilling algorithm for power allocation in information theory)

- Knowledge of optimization can help you formulate problems that are more interesting/useful/ easier to solve

- Knowledge of convex optimization may be extended to dealing with non-convex problems

## What can we expect to learn from this course?

# Course Goal

The goal of this course is to help you to develop the skills and background needed to recognize, formulate and solve convex optimization problems.

Course structure:

- Part I: Foundation
- Part II: Applications
- Part III: Algorithms

# Part I: Foundation

Involve working knowledge of

- Real analysis, calculus, linear algebra

Topics include:

- Convex sets
- Convex functions
- Convex optimization problems
- Duality

# Part II: Applications

Involve working knowledge of

- Problems in machine learning and statistics

Topics include:

- Approximation and fitting
- Statistical estimation

# Part III: Algorithms

Involve working knowledge of

- Data structure, computational complexity, programming (R, Python, Matlab)

Topics include:

- Unconstrained minimization
- Equality constrained minimization
- Interior-point methods

# Logistics

# Lecture and Office Hour

- Lecture
  - TR 9:30−10:45 AM
  - Colburn Lab 109
- Office hour
  - TR 11:00-12:00 AM
  - Evans 314
- Course website

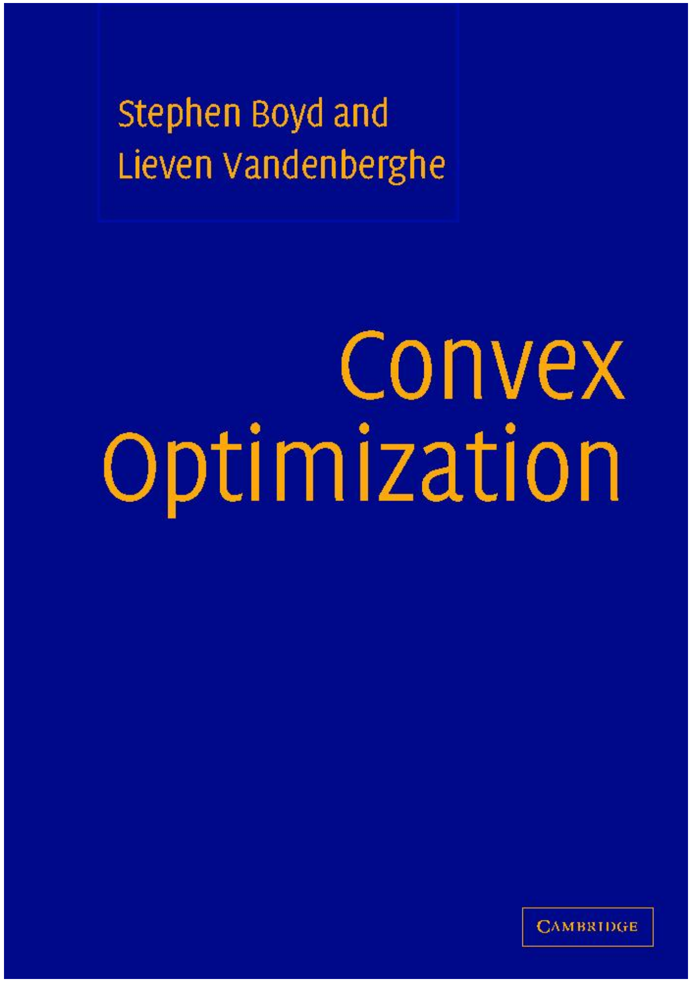
  https://www.eecis.udel.edu/~xwu/class/ELEG667/

# Prerequisite

- Undergraduate-level linear algebra, calculus, and probability theory; mathematical maturity in general

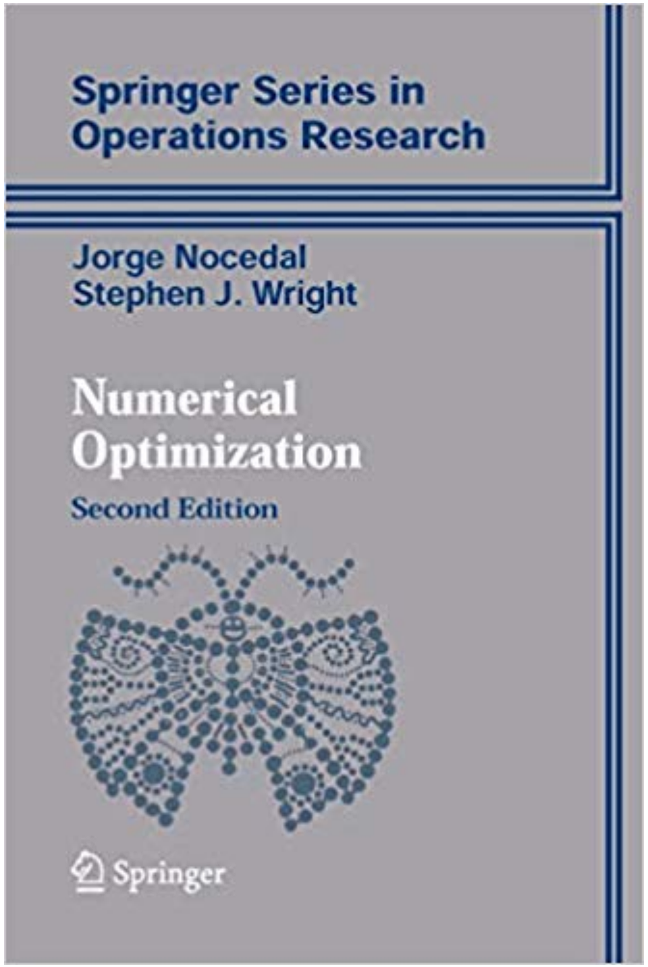
- Previous exposure to optimization is preferred but not mandatory

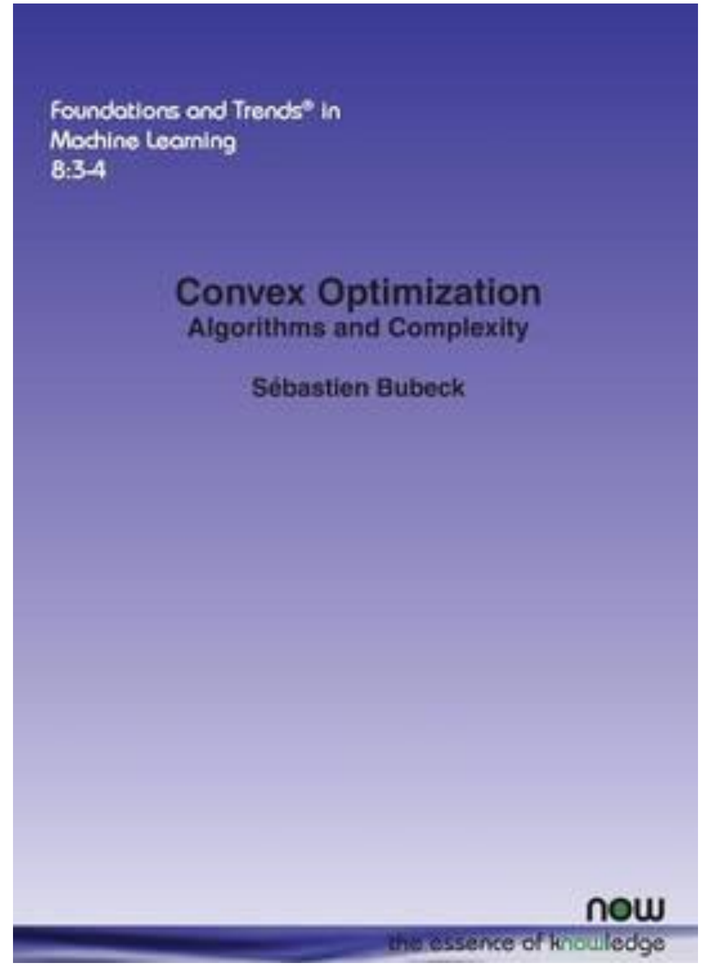
# Textbook




- Free pdf version online
- Lecture slides on course website

# References




Both free to download

# Grading

- Attendance: 10 points

  - Six random sign-in's (after the class is stabilized)

  - Two points for each attendance (5 out of 6 gives you max 10 pts)

- Homework: 40 points

  - Three in total, one for each part

  - Due in 1-2 weeks after being posted

- Final: 50 points + 10 bonus points

  - Closed book with one letter-size aid-sheet allowed

Questions?