Course outline: CO 673 / CS 794
Optimization for Data Science
Fall 2018

Stephen Vavasis

August 1, 2018

Course Goals

The course will cover optimization techniques used especially for machine learning and data science. Because these fields typically give rise to very large instances, first-order optimization (gradient-based) methods are typically preferred.

After finishing this course, students should be able to

- Recognize optimization problems arising in data science and know how to distinguish convex from nonconvex.
- Have a toolkit of algorithmic techniques that can be used on these problems.
- Know how to write code for common optimization algorithms.
- Know how to analyze certain classes of algorithms.

Prerequisites

Knowledge of linear algebra, multivariate calculus, basic analysis (convergence, limits), basic probability (common distributions, means, and so on). Knowledge of programming in one of Python, Matlab, Julia or R.

Preliminary list of topics

- Convex sets and functions
- Gradient descent
- Projection and alternating projection

- Convex relaxation
- Overfitting and regularization
- Subgradient methods
- Proximal gradient
- Nesterov's methods
- Conditional gradient (Frank-Wolfe) methods
- Coordinate descent
- Augmented Lagrangian methods
- ADMM
- Stochastic gradient methods
- The EM algorithm

Texts

- S. Bubeck, *Convex Optimization: Algorithms and Complexity*, Foundations and Trends in Machine Learning, 2015. Preliminary version available on arxiv.org.
- P. Jain and P. Kar, *Nonconvex optimization for Machine Learning*, 2017. Preliminary version available on arxiv.org.