COM S 578X: Optimization for Machine Learning

Fall 2018

Course Description

This course will introduce advances in optimization theory and algorithms with rapidly growing applications in machine learning, including first-order methods, stochastic optimization, sparsity, regularized optimization, compressed sensing, higher-order methods, interior-point methods, proximal methods, robust optimization, etc. The goal of this course is to prepare graduate students with a solid theoretical and mathematical foundation at the intersection of optimization and machine learning to conduct advanced research in the related fields. Compared to the optimization courses taught in operation research or related engineering fields, this course is focused more on topics that are of special interest in the machine learning community, such as convergence rate analysis, momentum-based acceleration, distributed and asynchronous algorithm design, sparsity regularized optimization problems, saddle point escaping, etc.

Prerequisites

Working knowledge of linear algebra and probability. Prior exposure to optimization is a plus but not necessary.

Course Information

Time: Tue/Thu 11:00am-12:15pm
Location: TBD
Instructor: Jia (Kevin) Liu (`jialiu@iastate.edu`)
Office: 209 Atansoff Hall
Office hours: Wed 5:00pm-6:00pm or by appointment
Course Web: `http://web.cs.iastate.edu/~jialiu/teaching/COMS578_F18/`

Course Overview

Since its inception as a discipline, machine learning has made extensive use of optimization formulations and algorithms. Likewise, machine learning has contributed substantially to optimization theory, driving the development of new optimization approaches that address the significant challenges presented by machine learning applications. This cross-inspiration continues to deepen, producing a vast literature at the integration of the two fields, while increasingly attracting leading researchers to this effort. This course gears toward such an intersection of the two fields.

Optimization techniques have enjoyed a critical role in machine learning because of their wide applications and theoretical appeals. While classical algorithms proposed decades ago continue to be refined, the ever-increasing complexity, size, and variety of today's machine learning applications demand a systematic reassessment of traditional assumptions and techniques. Besides describing the revival of classical algorithms in novel contexts, such as first-order methods, stochastic approximations, convex relaxations, interior-point methods, and proximal methods, the course devotes significant attention to newer themes such as regularized optimization, compressed sensing, robust optimization, a variety of gradient and subgradient methods with acceleration, and the use of splitting techniques and second-order information. We aim to provide a fresh account of optimization theory/algorithms relevant to machine learning – those that have played important roles in history as well as those that are rising in influence.

Textbooks and References

There is no required textbook. Most of the materials covered in the class will be based on classical books and recently published papers and monographs. A list of historically important and/or trending papers will be provided on the course website. Some useful reference books on optimization theory and mathematical backgrounds include:

- S. Boyd and L. Vandenberghe, "Convex Optimization," Cambridge University Press, 2004.
- Y. Nesterov, "Introductory Lectures on Convex Optimization: A Basic Course," Springer, 2004.
- M. Bazarra, H.D. Sherali, and C.M. Shetty, "Nonlinear Programming: Theory and Algorithms," John Wiley & Sons, 2006.
- D. Luenberger, "Optimization by Vector Space Methods," John Wiley & Sons, 1969.

Topics and Schedule

Below is an estimated class schedule, which is subject to change depending on lecture progress and/or class interests.

- **Week 1–2: Fundamentals of Convex Analysis**
  - Review of basic linear algebra and probability
  - Convex sets and functions
  - Strong and weak duality, constraint qualifications
  - Optimality conditions for machine learning problems (regressions, SVM, etc.)
- **Weeks 3–7: First-Order Methods**
  - Gradient descent convergence analysis
  - Convergence analysis for momentum-based acceleration methods: Heavy-ball, multi-step, Nesterov, FISTA, etc.
  - Convergence speedup with conjugacy
  - Convergence analysis for subgradient methods
  - Stochastic (sub) gradient descent (convergences in probability and distribution, almost sure convergence, parallelism, applications in deep learning, etc.)
- **Weeks 8–9: Higher-Order Methods**
  - Newton's method: convergence analysis (exact/inexact step-sizes, self-concordance), applications in regressions
  - Quasi-Newton Theory (Secant methods), convergence proofs for BFGS/DFP, L-BFGS in machine learning
- **Weeks 10–12: Sparse/Regularized Optimization**
  - ℓ^1-regularized sparse optimization for machine/statistical learning: compressed sensing, LASSO, logistic regression, etc.

  - Structured sparsity optimization for machine/statistical learning: low-rank matrix completion, nuclear norm regularization, inverse covariance inference, atomic norm regularization, etc.

- **Weeks 13–14: Proximal and Operator Splitting**
  - Dual decomposition and decentralization
  - Method of multipliers and ADMM methods: convergence analysis and proofs
  - Proximal operators and proximal methods
  - Design and analysis of distributed algorithms

- **Weeks 15-16: Nonconvex Optimization in Machine Learning**
  - Coordinate descent methods and convergence analysis
  - Special structured nonconvex optimization
  - Optimization landscape
  - Saddle point escape

Homework

There will be 4-5 homework assignments, assigned roughly biweekly. Homework *must* be typeset in LaTeX. The LaTeX homework template can be found on the course website.

Midterm

The in-class midterm exam will be closed-book and closed-notes, but you are allowed to bring a 1-page cheat sheet. The midterm will be comprehensive up to the finished lectures.

Final Project

Students could choose to finish the final project individually or by a team of no more than two persons. Project proposals will be due soon after midterm. Final reports will be due by the *beginning* of the final exam week. Final reports should follow IEEE or ACM format. Each project is required to have a 15-minute in-class presentation at the end of the semester. Attendance to your fellow students' presentations is *required* and will be accounted in your final grade. Potential project ideas include but are not limited to:

- Nontrivial extension of the results introduced in class.
- Novel applications in your own research area.
- New theoretical analysis of an existing algorithm.

Each project should contain something new and useful. It is important that you justify its novelty. **Note that an outstanding project could lead to a publication of yours!** ☺

Grading Policy

Homework: 30%; Midterm: 30%; Final project: 40%.

Late Policy

Without the consent of the instructor, late homework assignments or final report will not be accepted and will result in a grade of zero. In the case of a conference deadline or something of the like, a **5-day** written notice of extension is required. In the case of an emergency (sudden sickness, family problems, etc.), an after-the-fact notice is acceptable. But we emphasize that this is reserved for true emergencies.

Disabilities Statement

Any student who feels s/he may need an accommodation based on the impact of a disability should contact the instructor privately to discuss specific needs. Please contact the ISU Student Disability Resources Office (SDR) for assistance in verifying the need for accommodations and developing accommodation strategies.

Academic Misconduct Statement

Academic misconduct is any activity that tends to compromise the academic integrity of the university, or subvert the educational process, and is considered a serious offense. Any student found to have engaged in academic misconduct, as set forth in the ISU Code of Academic/Research Misconduct for Students, will be subject to disciplinary action by the university.

Student Conduct

Students are expected to abide by the provisions in the Code of Student Conduct. The University's Code of Student Conduct and Sexual Harassment Policy are available on the ISU Web page.