Large Scale Machine Learning and Optimization

Instructor: Dimitris Papailiopoulos, Assistant Professor, ECE Department
Email: papailiopoulos@gmail.com

Course Outline:

This course will explore the mathematical foundations of a rapidly evolving new field: large-scale optimization and machine learning. We will focus on recent texts in machine learning, optimization, and randomized algorithms, with the goal to understand the tradeoffs that are driving algorithmic design in this new discipline. These trade-offs will revolve around statistical accuracy, scalability, algorithmic complexity, and implementation.

Sample topics include:

1. Optimization and Learning
- Stochastic Methods for Convex and Nonconvex Settings
- Constrained and Projection-free Optimization
- Overfitting, Generalization, and Algorithmic Stability

2. Large Scale Learning and Systems
- Parallel Stochastic Methods
- Asynchronous Shared-memory Algorithms
- Distributed and Multi-core Learning Frameworks

3. Semidefinite and Linear Programs
- Semidefinite relaxations and Randomized Rounding
- The Multiplicative Updates Meta-algorithm
- Parallel Solvers

Key References

[1] *Convex Optimization: Algorithms and Complexity* (book)
http://research.microsoft.com/en-us/um/people/sebubeck/book.html
[2] *Understanding Machine Learning: From Theory to Algorithms (book)*
http://www.cs.huji.ac.il/~shais/UnderstandingMachineLearning/index.html
[3] *Robust stochastic approximation approach to stochastic programming*
http://www2.isye.gatech.edu/~nemirovs/SIOPT_RSA_2009.pdf

[4] *Stability and Generalization*
http://www.jmlr.org/papers/volume2/bousquet02a/bousquet02a.pdf
[5] *Optimization Methods for Large-Scale Machine Learning*
https://arxiv.org/pdf/1606.04838v1.pdf
[6] *Hogwild!: A Lock-Free Approach to Parallelizing Stochastic Gradient Descent*
https://people.eecs.berkeley.edu/~brecht/papers/hogwildTR.pdf
[7] *The Multiplicative Weights Update Method: A Meta-Algorithm and Applications*
http://isites.harvard.edu/fs/docs/icb.topic1465468.files/v008a006.pdf
[8] *Sparse Approximate Solutions to Semidefinite Programs*
http://ie.technion.ac.il/~ehazan/papers/SparseSDP.pdf

Student Evaluation

- Semester project: 50% (mid and end-semester report, and project presentation)
- Paper Presentation: 30%
- Scribe Notes: 20%

Scribing: All students are required to scribe notes for one lecture. Scribe notes will be due one week after their corresponding lecture. Depending on the size of the class, up to B students will be selected per lecture, so that #enrolled_students/B = #total_scribes.

Prerequisites

This course is ideal for advanced graduate students, who are interested in applying novel research concepts to their own research. This course is for students interested in the mathematical foundations of large-scale optimization for machine learning. Students are expected to be familiar with basic concepts in optimization and machine learning, and a solid background in linear algebra and probability.

Enrolled students are required to have attended at least 1 course, from at least 2 of the following categories, or equivalent, pending the instructor's approval.

ECE/CS/ME 532: Theory and Applications of Pattern Recognition
CS/ISyE 524 and ECE 601: Introduction to Optimization

ECE 729 Theory of Information Processing and Transmission
ECE 730: Probability and Stochastic Processes

ECE 901: Statistical Learning Theory
ECE 830: Statistical signal processing
CS 761: Advanced Machine Learning