Numerical Optimization with Deep Learning

Course Title
Advanced Topics in Machine Learning: Numerical Optimization with Deep Learning

Short Title
Numerical Optimization with Deep Learning

Introduction:
Optimization lies at the heart of machine learning. By quantitatively formulating the objective of modeling, it allows machine learning methods to flexibly incorporate domain knowledge in applications such as computer vision and natural language processing. To effectively and efficiently solve the optimizations problem, a large body of techniques have been developed recently. The course will investigate them in three steps: 1) introduce the mathematical foundations and general ideas; 2) customize general techniques to specific machine learning problems by exploiting additional structures; 3) study the practical performance on convex and nonconvex problems (e.g. deep learning).

General reason for this request:
To create a course on the machine learning topic of convex and nonconvex optimization that will prepare graduate students to conduct research in this area.

Purpose of this Course:
Exposes students to recently developed methods for solving large scale convex and nonconvex optimization problems that arise in machine learning and data analytics.

Relationship to similar courses:
Introduction to Machine Learning (CS 412) provides a broad overview of machine learning techniques without an advanced focus on the task of convex and nonconvex optimization.

Course Goals:
The course will consist of lectures, reading and discussion of research papers, and individual projects. The goals of the course are:

- To introduce students to the primary machine learning tasks and the techniques that have been developed for those tasks.
- To enable students to apply cutting edge optimization techniques to solve machine learning problems, and effectively communicate their findings. This includes both mathematical formulations and computer implementations, including parallel programming.
- To enable students to understand the pros and cons of various optimization techniques, and to *customize* solutions for new problems that arise in practice.

Prerequisites:
CS 251; and MATH 310 or MATH 320 (or equivalent linear algebra); or consent of the instructor.

Restrictions:

Restricted to students in the following colleges/schools: Engineering or Graduate College.

Credit Hours:
4 graduate hours

Type of Instruction: Lecture-Discussion
Contact Hours/Week Over 15-Week Term: 3

Course Learning Outcome Assessment Methods:
Course participation, presentations, and course projects

Major Topics:

1. Convex optimization formulations (3 weeks)
    - Convex sets and convex functions
    - Problems in Convex Optimization (linear/quadratic/Semi-definite programming)
    - Strong and weak duality

2. Applications in machine learning (2 weeks)
    - Maximum margin models with structured output
    - Sparse coding and dictionary learning
    - Convex relaxations

3. Algorithms for solving convex optimization (5 weeks)
    - Simplex and ellipsoid algorithm
    - First order optimization techniques: (stochastic) gradient descent, conjugate/conditional gradient, ADMM
    - Second order optimization techniques: quasi-Newton and Hessian free methods
    - Solving machine learning problems and off-the-shelf solvers
    - Rates of convergence (1.5 hours at most, not for examination)
    - Parallel distributed solvers (Map-Reduce)

4. Nonconvex optimization and deep learning (5 weeks)
    - Introduction to deep learning
    - Optimization for training deep models
    - Convex relaxations of deep learning

Instructor:
The course will be taught by Prof. Xinhua Zhang.

Expected Registration:
Graduate College

Students will be MS and PhD-level students primarily in Computer Science and Electrical and Computer Engineering.

Readings:

Stephen Boyd and Lieven Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004. ISBN: 9780521833783. (PDF free at http://stanford.edu/~boyd/cvxbook/)

Ian Goodfellow, Yoshua Bengio, and Aaron Courville, *Deep Learning*. Book in preparation for MIT Press, 2016. http://goodfeli.github.io/dlbook. (Chapters 6 and 8)

Recent research papers and excerpts of relevant background material from available textbooks
Supplemental notes for specific topics