ISE 633 Large Scale Optimization for Machine Learning

Number of units: 03

Location and time: WPH 104, Tuesday/Thursday 12:30-1:50pm

Instructor: Meisam Razaviyayn
Email: razaviya@usc.edu
Office hours: Thursdays 2-3pm, OHE 310G

Teaching Assistant: Tianyu Hao
Email: tianyuha@usc.edu
Office hours: Mondays 12-2pm, GER 309

Goal: The objective of the course is to introduce large scale optimization algorithms that arise in modern data science and machine learning applications.

Course Description: Large scale optimization algorithms that arise in modern data science and machine learning applications. Stochastic Optimization, Accelerated Methods, Parallelization, Online Optimization, Randomized Linear Algebra

Textbook: There is no required textbook for the class. All course materials will be presented in class or will be available online as notes. The following textbooks cover parts of the course materials and you may find them useful:

- D. P. Bertsekas, Nonlinear Programming, Belmont: Athena scientific, 1999.
- S. Boyd and L. Vandenberghe, Convex optimization, Cambridge university press, 2004.
    - The book is available for free here: http://web.stanford.edu/~boyd/cvxbook
- S. Shalev-Shwartz and S. Ben-David, Understanding Machine Learning: From Theory to Algorithms. Cambridge University Press, 2014.
    - The book is available for free here: http://www.cs.huji.ac.il/~shais/UnderstandingMachineLearning/
- A. Shapiro, D. Darinka, and A. Ruszczynski, Lectures on Stochastic Programming: Modeling and Theory, SIAM, 2009.
    - The book is available for free here: http://www2.isye.gatech.edu/~ashapiro/publications.html

Course Plan:

- Week 1: (Aug 22-24)
  - Optimization overview, examples in machine learning, large scale optimization, memory/time/CPU requirement
  - Mathematical Basics

- Week 2: (Aug 29 – Aug 31)
  - Unconstrained optimization, necessary optimality conditions, smooth versus non-smooth optimization
  - Sufficient optimality conditions, convex versus non-convex optimization
  - Homework 1 assigned: Due Sep 12, 2017

- Week 3: (Sep 5 – Sep 7)
  - Gradient methods (unconstrained), choices of direction,
  - Asymptotic convergence, Newton method

- Week 4: (Sep 12 – Sep 14)
  - Rate of convergence of gradient descent, first order oracle model
  - Lower and upper bounds in the oracle model
  - Homework 2 assigned: Due Sep 26, 2017

- Week 5: (Sep 19 – Sep 21)
  - Accelerated Nesterov
  - Constrained optimization, optimality conditions

- Week 6: (Sep 26 – Sep 28)
  - KKT optimality conditions and Lagrange multipliers
  - Projection and algorithms, examples in machine learning
  - Homework 3 assigned: Due Oct 10, 2017

- Week 7: (Oct 3 – Oct 5)
  - Exploiting multi-block structure of the problem, examples in machine learning, block coordinate descent methods
  - Different block selection rules and convergence analysis

- Week 8: (Oct 10 – Oct 12)
  - Block successive upper-bound minimization and its convergence
  - **Midterm**

- Week 9: (Oct 17 – Oct 19)
  - Alternating direction method of multipliers
  - Non-smooth optimization and examples in machine learning
  - Homework 4 assigned: Due Oct 31, 2017

- Week 10: (Oct 24 – Oct 26)
  - Necessary and sufficient conditions in non-smooth optimization, successive upper-bound minimization, proximal operator
  - Multi-block methods in non-smooth optimization

- Week 11: (Oct 31 – Nov 2)
    - Stochastic/Online/Incremental optimization
    - Incremental gradient and its analysis
    - Homework 5 assigned: Due Nov 14, 2017

- Week 12: (Nov 7 – Nov 9)
    - Sample Average Approximation and Stochastic Approximation
    - Analysis

- Week 13: (Nov 14 – Nov 16)
    - Parallelization: synchronous vs asynchronous
    - Adversarial viewpoint and regret analysis
    - Homework 6 assigned: Due Nov 28, 2017

- Week 14: (Nov 21)
    - Non-convexity and examples in machine learning: principal component analysis, deep learning, non-negative matrix factorization
    - Local optimality results

- Week 15: (Nov 28 – Nov 30)
    - Randomized linear algebra: power method, faster than power method
    - Randomized linear algebra: analysis

Course Requirement and Grading:
- In-class midterm (30%)
- Final exam (35%)
- Homework assignments (Best 4 out of 6: 20%)
- Participation (5%)
- Scribing (10%)

Homework assignments:
- All homework assignments are due by 4:30pm on the date indicated.
- Homework assignments must be submitted via email to the TA with cc'ing the instructor. Only one pdf file should be submitted for each homework assignment. You can submit latex pdf files or scanned images which are converted to pdf format.
- Late homework submissions are not accepted under any circumstances. Start your homework assignments early.
- There will be six homework assignments. The two lowest scores will not be considered in your final grade.
- You are encouraged to discuss homework assignments with other students. However, each student is required to submit his/her own personal work.

Scribing: In order to gain experience with technical writing, each student is required to scribe notes for two lectures. These notes will be revised by the instructor and will be posted on the course website. The scribed notes should be written in a way that they are completely understandable to a student who may have missed the class.

University policies:

- *Statement for Students with Disabilities.* Any student requesting academic accommodations based on a disability is required to register with Disability Services and Programs (DSP) each semester. A letter of verification for approved accommodations can be obtained from DSP. Please be sure the letter is delivered to your course instructor (or TA) as early in the semester as possible. DSP is located in STU 301 and is open from 8:30am to 5:00pm, Monday through Friday. Website and contact information for DSP: http://sait.usc.edu/academicsupport/centerprograms/dsp/home_index.html, (213) 740 – 0776n (Phone), (213) 740-6948 (TDD only), (213) 740-8216 (FAX), ability@usc.edu.

- *Statement on Academic Integrity.* USC seeks to maintain an optimal learning environment. General Principles of academic honesty include the concept of respect for the intellectual property of others, the expectation that individual work will be submitted unless otherwise allowed by an instructor, and the obligations both to protect ones own academic work from misuse by others as well as to avoid using another's work as ones own. All students are expected to understand and abide by these principles. SCampus, The Student Guidebook, contains the Student Conduct Code in Section 11.00, while the recommended sanctions are located in Appendix A: http://usc.edu/dept/publications/SCAMPUS/gov/. Students will be referred to the Office of Student Judicial Affairs and Community Standards for further review should there be any suspicion of academic dishonesty. The Review process can be found at: http://usc.edu/student-affaris/SJACS/. Information on intellectual property at USC is available at: http://usc.edu/academe/acsen/issues/ipr/index.html.

- *Emergency Preparedness/Course Continuity in a Crisis*. In case of emergency, when travel to campus is difficult, if not impossible, USC executive leadership will announce a digital way for instructors to teach students in their residence halls or homes using a combination of the Blackboard LMS (Learning Management System), teleconferencing, and other technologies. Instructors should be prepared to assign students a ``Plan B" project that can be completed ``at a distance". For additional information about maintaining your classes in an emergency, please access: http://cst.usc.edu/services/emergencyprep.html.