Optimization for Machine Learning

DSCC 435 / CSC435 / ECE 412 Fall 2023

Meeting Information

Tuesday/Thursday 9:40-10:55 am
Hylan Building Room 203

Instructor

Jiaming Liang
Email: jiaming.liang@rochester.edu
Office: Wegmans Hall 2403

Teaching Assistant

Lin Zang
Email: lzang@simon.rochester.edu

Office Hours

4:00-5:00 pm, Wednesday, Wegmans Hall 2403 (Jiaming Liang)
4:00-5:00 pm, Friday, Wegmans Hall 1219 (Lin Zang)

Prerequisites

Students should be familiar with multivariate calculus, linear algebra, basic probability, and have good MATLAB or Python programming skills. Prior knowledge of optimization is helpful but not required.

Course Description

This course primarily focuses on algorithms for large-scale optimization problems arising in machine learning and data science applications. The first part will cover first-order methods including gradient and subgradient methods, mirror descent, proximal gradient method, accelerated gradient method, Frank-Wolfe method, inexact proximal point methods, and more advanced topics. The second part will introduce algorithms for nonconvex optimization, stochastic optimization, distributed optimization, manifold optimization, reinforcement learning, and those beyond first-order.

Topics (subject to change)

1. Introduction
    - Applications in ML
    - Convex analysis and complexity analysis
2. First-order methods I: basic concepts
    - Gradient method
    - Subgradient method

- Mirror descent
- Proximal gradient method
- Accelerated gradient method
- Frank-Wolfe method

3. First-order methods II: advanced topics
    - Inexact proximal point methods
    - Operator splitting
    - Randomized block coordinate descent
    - Optimization in relative scale
    - Smoothing techniques
4. Selected topics in machine learning
    - Nonconvex optimization
    - Stochastic optimization
    - Distributed optimization
    - Manifold optimization
    - Reinforcement learning
5. Beyond first-order methods
    - Second-order methods
    - Higher-order methods

Textbooks

1. Amir Beck. *First-order methods in optimization*. SIAM, 2017.
2. Yurii Nesterov. *Lectures on convex optimization*. Springer, 2018.

Recommended Readings

1. Guanghui Lan. *First-order and Stochastic Optimization Methods for Machine Learning*. Springer, 2020.
2. Benjamin Recht and Stephen Wright. *Optimization for Data Analysis*. Cambridge University Press, 2022.
3. Suvrit Sra, Sebastian Nowozin, and Stephen Wright, eds. *Optimization for Machine Learning*. MIT Press, 2011.

This course will cover classical and modern algorithmic results in nonlinear optimization. For those results, we mainly use the two textbooks. The course also introduces several important and emerging applications of optimization in machine learning. Some topics by themselves could make up an entire course. This course is designed as an introductory gateway to these areas, for which we will draw from multiple references. The instructor will provide some additional resources (e.g., papers and excerpts from other books) on Blackboard and/or the course website.

Readings & Assignments

This is a tentative schedule of the readings and assignments.

Lecture	Date	Topic	Reading	Problem sets
1	08/31	Introduction		
2	09/05	Convex analysis, complexity analysis	AB 2, YN 1.1	
3	09/07	Gradient method	AB 10.1, YN 1.2.3	

4	09/12	Subgradients	AB 3, YN 3.1.5, 3.1.6	
5	09/14	Subgradient method	AB 8.2, YN 3.2.3	PS1 out
6	09/19	Mirror descent	AB 9	
7	09/21	Proximal gradient method	AB 10.2, 10.4, 10.6	
8	09/26	Accelerated gradient method	AB 10.7	
9	09/28	Frank-Wolfe method	AB 13.1	PS1 in, PS2 out
10	10/03	Inexact proximal point methods I	AB 10.5	
11	10/05	Inexact proximal point methods II	AB 14	
12	10/10	Operator splitting		PS2 in
13	10/12	Midterm		
14	10/19	Randomized block coordinate descent	AB 11.5	PS3 out
15	10/24	Optimization in relative scale	YN 7	
16	10/26	Smoothing techniques	AB 10.8, YN 6.1, 6.3	
17	10/31	Nonconvex optimization I	AB 10.3	PS3 in, PS4 out
18	11/02	Nonconvex optimization II		
19	11/07	Stochastic optimization I		Proposal due
20	11/09	Stochastic optimization II		
21	11/14	Distributed optimization I		
22	11/16	Distributed optimization II		PS4 in, PS5 out
23	11/21	Manifold optimization I		
24	11/28	Manifold optimization II		
25	11/30	Reinforcement learning I		
26	12/05	Reinforcement learning II		
27	12/07	Second and higher-order methods	YN 4.1, 4.2, 4.3	PS5 in
28	12/12	Project presentation		

Assessments & Grading

- 50% from homework assignments (5 problem sets x 10% each)
- 20% take-home midterm exam (Oct. 10-12)
- 10% from project proposal (Nov. 7)
- 10% from project presentation (Dec. 12)
- 10% from project write-up (Dec. 15)

The final grade will be assigned as a letter grade according to the following scale:

A	90-100%
B	80-89%
C	70-79%
D	60-69%
F	0-59%

Project

The final course project is intended to give students the opportunity for in-depth exploration of a topic in modern optimization methods for machine learning and data science. Course projects could look like one of the following:

- An in-depth survey of one of the topics covered in the class. A survey consists of a rigorous academic review of the literature related to the topic interpreted in the student's own words, and a possible discussion of future areas of research.
- An application of the algorithms discussed in class on a data set or an engineering application. The application can be non-standard; in fact, proposing new applications for the material developed in class is encouraged. However, the methodology in the implementation needs to involve techniques that we discuss during the class.
- A conceptual (i.e., either theoretical or simulation-based) topic of novel research related to the class. Preliminary results and directions for future work are common outcomes.

Project proposals that do not neatly fall into one of these categories are also welcome. Students will be asked to submit a short abstract around mid-October (ungraded), submit a proposal on November 7 (10%), make a presentation on December 12 (10%), and submit a final report on December 15 (10%).

Academic Integrity

Academic integrity is a core value of the University of Rochester. Students are strongly encouraged to discuss homework problems with one another. However, each student must write up and turn in their own solutions, written in their own words/consisting of their own code. All assignments and activities associated with this course must be performed in accordance with the University of Rochester's Academic Honesty Policy. More information is available at: http://www.rochester.edu/college/honesty.

Absences/Late Submissions

Out of fairness to the entire class, late submission of homework will not be accepted in the absence of a prior agreement between the student and instructor. In particular, excused absences include illnesses, religious observations, career fairs and job interviews. In the event than an excused absence such as above prevents a student from submitting an assignment, their homework grade will be calculated on a prorated basis.

Diversity, Equity, Inclusion, & Belonging

Instructors, teaching assistants, and students should work together to ensure that our class is a welcoming, inclusive, respectful, and vibrant place for all of its members to share, learn, and grow. Our class will not tolerate discrimination, prejudice, or harassment of any kind. More resources can be found at: https://www.rochester.edu/diversity/.

Accessibility

Students needing academic adjustments or accommodations because of a documented disability must contact the Disability Resource Coordinator for the school in which they are enrolled. I am happy to accommodate any and all accommodations, so long as they are documented with the Office of Disability Resources. I am glad to meet to discuss your specific situation or to help ensure you have the support you need. For additional information, please see: https://www.rochester.edu/college/disability/.