COMM 616: Modern Optimization with Applications in ML and OR 2024-25 Fall

Information Sheet

Instructor: Jiajin Li August 26, 2024

1 General Information

- INSTRUCTOR: Professor Jiajin Li
  - OFFICE: Henry Angus - HA 472
  - OFFICE HOURS: By appointment
  - EMAIL: `jiajin.li@sauder.ubc.ca`
- CLASS TIME & LOCATION:
  - Tuesday 10:00 am — 13:30 pm, in ANGU 432
- CLASS WEBSITE: https://gerrili1996.github.io/comm616/
- Online Grading System: Canvas

2 Course Objective

This course offers a comprehensive exploration of modern optimization theory and algorithms, with applications in machine learning and operations research. It begins with an examination of fundamental concepts and problem properties in optimization, such as convexity, duality, smoothness, and subdifferentials. Students will then learn to design first-order optimization algorithms tailored to various problem characteristics, with a strong focus on how these structures impact convergence analysis.

The course also covers lower complexity bounds for different function classes, providing students with a critical understanding of the computational limits inherent in optimization algorithms. In the final part of this class, students will apply these advanced optimization techniques to tackle real-world challenges in machine learning, data science, and operations research, with particular attention to topics like optimal transport and distributionally robust optimization.

No prior optimization background is required for this class. However, students should have workable knowledge in (real) analysis and linear algebra.

3 Syllabus

Part I Optimization Theory

- **Lecture 1**: Introduction and classes of optimization problems; Convexity analysis and duality theory
- **Lecture 2**: Strong convexity, weak convexity and smoothness; Gradient descent (GD)
- **Lecture 3 (a)** : Nonsmoothness and various subdifferential concepts

Part II First-Order Optimization Algorithms and Their Convergence Analysis

- **Lecture 3 (b):** Algorithm design principles and optimality residuals
- **Lecture 4**: Subgradient method and Proximal Point Algorithms (PPA) for (non)convex problems
- **Lecture 5**: Proximal Gradient Descent (PGD) for structured (non)convex optimization problems
- **Lecture 6**: Convergence analysis using error bound-type conditions; Review of subgradient, PGD, and PPA under weaker conditions
- **Lecture 7**: Augmented Lagrangian Methods (ALM) and Alternating Direction Method of Multipliers (ADMM) for constrained convex/nonconvex optimization problems
- **Lecture 8**: Minimax optimization problems and their algorithmic development

Part III Lower Complexity Bounds of Different Function Classes

- **Lecture 9**: Lower bounds for strongly convex and smooth, convex and smooth, and smooth functions

Part IV Applications in ML and OR

- **Lecture 10:** Iterative methods for Optimal Transport and Gromov-Wasserstein distance; Bregman geometry and algorithm extensions
- **Lecture 11:** Tractable formulations for various distributionally robust optimization problems
- **Lecture 12:** Tailored algorithm design for distributionally robust optimization problems
- **Lecture 13**: Student oral presentations

4 Grading

- HOMEWORK 40%: There will be about 2 problem sets during the term. Typically, they are due two weeks after being assigned. Eeach student is granted a total of two late days, which can be used to submit homework assignments after their due date without penalty. However, once you have used up the late days, **no more late homeworks will be accepted unless prior arrangement has been made with the instructor.**

  You are encouraged to discuss the homework with your classmates; however, you must write up your solutions independently. Please indicate any individuals with whom you have discussed or collaborated. Plagiarism and other anti-scholarly behavior will be dealt with severely.

- PROJECT (final in-class presentation and report) 50%: As part of this course, you are required to complete a research or pedagogical project. A list of problems will be provided, from which you may select one to address using the optimization techniques covered in class. Alternatively, you may propose a problem of your own interest, subject to approval by the instructor. Additional details will be provided as the course progresses.

- SCRIBE NOTE 10%: Each student will be assigned a session to take detailed notes on the lecture. These notes will be shared with the class as a resource. Your scribe notes should be thorough, accurate, and well-organized. The quality of your scribe notes will be evaluated based on clarity, completeness, and adherence to the assigned format. Further guidelines will be provided at the start of the course.

References

[1] Amir Beck. *First-order methods in optimization.* SIAM, 2017.

[2] Dimitri Bertsekas. *Convex optimization algorithms.* Athena Scientific, 2015.

[3] Stephen Boyd and Lieven Vandenberghe. *Convex optimization.* Cambridge university press, 2004.

[4] Ying Cui and Jong-Shi Pang. *Modern nonconvex nondifferentiable optimization.* SIAM, 2021.

[5] Jean-Baptiste Hiriart-Urruty and Claude Lemaréchal. *Fundamentals of convex analysis.* Springer Science & Business Media, 2004.

[6] Yurii Nesterov. *Introductory lectures on convex optimization: A basic course*, volume 87. Springer Science & Business Media, 2013.

[7] Jorge Nocedal and Stephen J Wright. *Numerical optimization.* Springer, 1999.

[8] R Tyrrell Rockafellar and Roger J-B Wets. *Variational analysis*, volume 317. Springer Science & Business Media, 2009.

Other supplementary material will be posted on the course website as the course progresses.