Math 790 Fall 2025: Minimax problem, saddle point, and robust learning

Instructor: Xiuyuan Cheng Email: xiuyuan.cheng@duke.edu
Time: Oct. 30- Dec. 2, Tu Th 1:25-2:40pm
Classroom: Gross 304B

Overview: This minicourse will provide an introduction to some recent developments in the statistics/computer science literature on solving the minimax problem, that is, min_x max_y f(x,y), in search of a saddle point. The course mainly consists of lectures, and we will start from the basics. We will focus on first-order schemes like Gradient Descent Ascent (GDA) and Extra Gradient Method (EGM), and will go over some bounds on the iteration complexity. We will first look into the (strongly) convex-concave case, and then move to the nonconvex-nonconcave case under additional assumptions. Finally, we will introduce the application to distribution-robust machine learning, extending the problem formulation from Euclidean space to the space of probability distributions (Wasserstein-2 space).

The lectures are intended for beginning graduate students in applied math and related subjects, and the assumed background is gentle — a solid understanding of multivariate calculus will go a long way in following the lectures, and familiarity with algorithms/optimization/machine learning would be a plus. There are no planned homeworks or exams. Audition welcome and registration encouraged.

References:

[BTN2024] Ben-Tal, A. and Nemirovski, A., Lectures on modern convex optimization: analysis, algorithms, and engineering applications. (Latest version online)

We will mainly use Chapter 5. More related references will be provided in the lectures.

Tentative lectures:

L1: Overview, minimization vs minimax problem

L2: Gradient Descent Ascent (GDA), monotone field, Strongly convex-strongly concave case

L3: Proximal Point Method, Moreau envelope, saddle envelope

L4: Nonconvex-nonconcave case with interaction dominance, Extra gradient method

L5: Polyak-Łojasiewicz (PL) condition in minimization

L6: Two-timescale GDA, nonconvex-nonconcave case with PL condition

L7: Adversarial learning and distribution robust optimization