CS 269: Special Topic: Optimization Methods for Deep Learning
Chih-Jen Lin

Alert: this seminar course is project-heavy and highly research-oriented

Course information:

• Instructor:
  – Chih-Jen Lin, Room 293, Building Engineering VI.
  – Office hour: Wed 2-3pm (or by appointment)

• Time and place:
  – Monday, Wednesday 12pm-1:50pm, ROYCE 154.
  – We do a 10-minute break at 12:50pm.

• Course web page
  – It’s constantly updated to include slides, project materials, and other information

• FAQ

• Course outline
  Deep learning involves a difficult non-convex optimization problem. The goal of this course is to study the efficiency and the robustness of optimization methods for deep learning. We will run this course in the following formats:
    – lectures (by the instructor)
    – project presentations (by students)

This course is highly research oriented. It’s suitable for students interested in the research of optimization on deep learning.

• Prerequisites:
  – Machine learning
- Optimization
- Programming skill: Python and Matlab/Octave

**Course policy**

- Grading policy: 100% projects
- Accommodations for students with disabilities: To insure that disability-related concerns are properly addressed from the beginning, students with disabilities who require assistance to participate in this class are asked to see me as soon as possible. We aim to be as accommodating and fair as possible.
- Use of outside resources: When using outside resources, proper citation is necessary. This includes papers, text books, software libraries, websites, and helps from others. For the details of Honor Code, please refer to UCLA Student Conduct. If you have any doubt, please check with me in advance. You may get F in the final letter grade if we detect any cheating.
- Late policy: No late submission will be accepted.
- Collaboration: By default, students in the same team will get the same score unless special circumstances. We encourage students to use a version control system (e.g., github, gitlab, etc...). It is important to keep your hard work in a safe place and log the contributions of individuals. If your team members complain about you and you cannot provide evidence of your contribution, we may lower your score.

**Tentative Schedule:**

- Week 1-2: neural network optimization problems
- Week 3-4: gradient descent and stochastic gradient
- Week 5-6: gradient calculation and detailed implementation
- Week 7-8: Newton method and its implementation
- Week 9-10: Performance evaluation and practical considerations