Large-Scale Optimization (Fall 2023)
Instructor: Prof. Ulugbek S. Kamilov

Course Website: cigroup.wustl.edu/teaching/lso-f2023
Lecture: Mon and Wed, 1–2:20pm
Location: Hillman 70

Instructor Team:

Ulugbek Kamilov  Flora Sun  Weijie Gan
kamilov@wustl.edu zhixin.sun@wustl.edu weijie.gan@wustl.edu
Mon, 2:30-3:30pm  Fri, 4-5pm  Thu, 4-5pm

Prerequisites:
Knowledge of optimization is necessary (see ESE 415). Understanding of basic linear algebra and probability (measure-theoretic probability is not needed). Working knowledge of a scientific programming language such as Python.

Course Description:
Optimization over high-dimensions is essential for modern data science, artificial intelligence, and machine learning. This course introduces theoretical and algorithmic concepts on advanced optimization suitable for problems arising in these areas. We will learn several algorithms for both smooth and nonsmooth optimization, including gradient methods, subgradient methods, smoothing techniques, proximal methods, mirror descent, Nesterov acceleration, stochastic gradient methods, and alternating direction method of multipliers. We will also discuss the effectiveness of these methods from both theoretical and practical perspectives. The lectures by the instructor will be complemented by student lectures with topics drawn from recent literature on optimization.

Syllabus: The following list is subject to change, more likely by deletion than addition.
- Gradient methods
- Frank-Wolfe method
- Projected gradient methods
- Subgradient methods
- Mirror descent
- Nesterov accelerated techniques
- Proximal gradient methods
- Stochastic and block coordinate methods
- Alternating direction methods of multiplier

Last updated: 7 August 2023
Grading: The grade will be based on a weighted average of the following components:

- **Student lectures (20%)**: There will be regular student lectures. A team of 4-5 students will work together for each lecture. The list of topics with corresponding publications will be published on the course page. Topics will include the important publications in the field. The goal of each team is to understand the paper, present its key ideas, and provide additional commentary. Demos and implementations are highly encouraged. Each group as well as each student in the group will be individually evaluated by a set of peers. Each member of the group is expected to be able to answer questions from the whole paper, including from the parts presented by their team members.

- **Class participation (10%)**: Students are expected to actively participate in the class by attending the lectures, both by the instructor as well as by students. We will select a group of student evaluators for each student lecture, who are expected to read the paper, ask at least one question from each group, and submit the evaluation form.

- **Homework (20%)**: There will be a small number of assignments throughout the semester, with two weeks available to complete each one. Each assignment will have two types of questions: traditional “pencil-and-paper” questions, and programming exercises meant to give more insight into applying the methods we discuss. Assignments will not be formally corrected: If you make a reasonable attempt to answer a question, you will get the full credit. The solutions to each assignment will be provided online on the course website.

- **Class Project (50%)**: There will be a group project in the second half of the semester. Students will form groups of 5-6 people for the project. The goal of the project is to go deeper on a single topic related to large-scale optimization. You can explore topics by reading the publications that are part of student lectures or by selecting a topic that is aligned with your interests. The success of a project will be judged based on the methodological approach rather than the quantitative details of the final outcome. This is an opportunity for independently investigating theory, algorithms, and applications of large-scale optimization. Note that all the reports must be submitted by using the LaTeX template provided by the instructor. There are four components to the project:
  - **Project proposal (10%)**: The project proposal is due on Fri, 09/29. The proposal is a one page document describing the project. Feedback will be provided to each proposal.
  - **Project update (10%)**: The project update is due on Fri, 10/27. The project update is up to two pages with updates on the progress of the project.
  - **Project report (15%)**: The first full draft of the project report is due on Tue, 11/21. The report can have up to four pages explaining all the work and the results of the project. The team will be given a chance to revise the final report after the project presentation. The deadline for the final report is Fri, 12/08.
- **Project presentation (15%)**: Each group must prepare a 15-minute presentation describing the project. At least one member of the team must deliver the presentation and other team members must be available for answering the questions. The presentation should explain everything about the project to the instructor and other students in the class.

**Final course grade will be assigned using the following scale:**

<table>
<thead>
<tr>
<th>Grade</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>A+</td>
<td>≥ 97%</td>
</tr>
<tr>
<td>A</td>
<td>≥ 93%</td>
</tr>
<tr>
<td>A-</td>
<td>≥ 90%</td>
</tr>
<tr>
<td>B+</td>
<td>≥ 87%</td>
</tr>
<tr>
<td>B</td>
<td>≥ 83%</td>
</tr>
<tr>
<td>B-</td>
<td>≥ 80%</td>
</tr>
<tr>
<td>C+</td>
<td>≥ 77%</td>
</tr>
<tr>
<td>C</td>
<td>≥ 73%</td>
</tr>
<tr>
<td>C-</td>
<td>≥ 70%</td>
</tr>
<tr>
<td>D+</td>
<td>≥ 67%</td>
</tr>
<tr>
<td>D</td>
<td>≥ 63%</td>
</tr>
<tr>
<td>D-</td>
<td>≥ 60%</td>
</tr>
<tr>
<td>F</td>
<td>&lt; 60%</td>
</tr>
</tbody>
</table>

The **passing grade** is C- or better (≥ 70%).

**Optional Textbooks:**
The lecture notes for the class will be available on the class website. The following textbooks are optional:

- “Convex Optimization” by S. Boyd and L. Vandenberghe ([download here](#))
- “Introductory Lectures on Convex Optimization” by Y. Nesterov. ([video lectures](#))

**Disability:**
Washington University is committed to providing accommodations and/or services to students with documented disabilities. Students who are seeking support for a disability or a suspected disability should contact Disability Resources at 935-4153. Disability Resources is responsible for approving all disability-related accommodations for WashU students. If you have already been approved for accommodations, please inform the instructor within the first two weeks of the semester.