ECE273: Convex Optimization and Applications — Spring 2024

Instructor: Prof. Jun-Kun Wang
Lectures: Tuesday and Thursday 3:30PM-4:50PM at PETER (Peterson Hall) 102
Email: jkw005@ucsd.edu
Office Hours: Friday 3PM-4PM and by appointment
Office Location: Jacobs Hall Room 6404

Teaching Assistant (TA): Marialena Sfyraki
Email: msfyraki@ucsd.edu
Office Hours: Wednesday 1PM-2PM
Office Location: Jacobs Hall Room 4506

Course Description: This course covers the theoretical and algorithmic foundations of optimization. We will cover some convex analysis and duality theory, and we will walk through classical optimization algorithms such as Gradient Descent, Coordinate Descent, the Frank-Wolfe Method, Accelerated Methods, Mirror Descent, Stochastic Gradient Descent, and Online Gradient Descent. We will put particular emphasis on understanding the behaviors and the convergence rate guarantees of these algorithms, as well as the tools and techniques to analyze them. Students will learn the basic foundations of deterministic convex optimization, stochastic optimization, online convex optimization, min-max optimization, and non-convex optimization.

Students are expected to sign up on **Piazza** and **GradeScope [Code: KK2Y57]**. Discussions will happen on Piazza. The homework should be turned in and will be graded on GradeScope. Important announcements will be made via Canvas.

Grading Policy: 40% homeworks (5 problem sets; the lowest one will be dropped), 25% midterm exam, 25% final project, 10% scribing and 5% extra credit for course attendance/participation.

List of Topics

- Mathematical background and Gradient Flow.

- Convexity, strong convexity, smoothness of a function.

- Gradient Descent and its convergence analysis

- Reduction

- Properties of convex sets

- Subgradients and optimality conditions

- Projected Gradient Descent and Frank-Wolfe Method

- Coordinate Descent

- Duality Theory, KKT conditions, and Stochastic Dual Coordinate Ascent

- Mirror Descent

- SGD

- SGD and variance reduction

- Online Convex Optimization

- Min-Max Optimization

- Nonconvex optimization (Gradient Dominance Condition, Finding a stationary point)

- Acceleration via the Chebyshev Polynomial

- Momentum methods

- Optimization as a game

Homework Policy: Students can work with up to two other classmates on the homeworks. However, the solutions submitted must be independently written in your own words. As a hard rule, you must write up your solution individually and cannot copy your classmates' solutions or let your classmates copy your solutions; otherwise, you may not receive any credit for the HW problem(s). You must write your collaborators names on the top of your assignment. If you do not work with collaborators, list Collaborators: None.

Students are encouraged to submit a typed solution. Written solutions are allowed but have to be clearly readable. Students might not receive credit for the problems if the written solutions are incomprehensible. We reserve the right to determine if a written solution is incomprehensible.

Each homework and project assignment will have a clearly stated due date. We expect you to turn in all completed problem sets on time. Late submissions and deadline extensions will not be possible for any reason.

Midterm: The midterm will take place during regularly scheduled class hours on **Thursday, May 23, 2024, from 3:30 pm to 4:50 pm**. It will be a closed-book exam; however, students are permitted to use a hand-written cheat sheet, limited to one page and two sides, on standard printer paper size (8.5 x 11 inches). The use of electronic devices, such as phones, tablets, laptops, and earbuds, is not allowed during the exam.

Students are expected to comply with UCSD's academic integrity policy throughout the duration of the midterm, https://senate.ucsd.edu/Operating-Procedures/Senate-Manual/Appendices/2. Instances of academic dishonesty will be referred to the Office of Student Conduct for adjudication.

Reasonable Accommodations: Students requesting accommodations for this course due to a disability must provide a current Authorization for Accommodation (AFA) letter (paper or electronic) issued by the Office for Students with Disabilities. Students are required to discuss accommodation arrangements with instructors and OSD liaisons in the department in advance of any exams or assignments. If you have submitted an AFA letter, please make arrangements to meet with the instructor no later than the end of Week 2 to ensure that reasonable accommodations can be arranged.

Final project: You can form a group of size up to 3 members for the final project.
Grading of the final project:
1% for a checkpoint report, 12% for a written report, 12 % for an oral presentation.
Written report of the final project:

- Describe an optimization paper in detail using your own words. The paper you focus on should be a peer-reviewed paper from top venues such as ICML, NeurIPS, ICLR, COLT, JMLR, Mathematical Programming, or SIAM Journal on Optimization from the past three years. Focusing on papers not from these venues requires discussion with the instructor and obtaining approval first.

  Students are required to **submit a typed write-up** for the final project consisting of 4-6 pages (the list of references does not count toward the page limit). Submitting a report that is less than 4 full pages (references excluded) will result in a score deduction. The write-up must contain a **significant theoretical component**. Please use the Latex template provided in this link `https://media.neurips.cc/Conferences/NeurIPS2024/Styles.zip`

  The report should include **the motivation, a description of the algorithm, the underlying algorithmic design principles, and a clear reproduction or summary of the theoretical analysis**. When reproducing the proof, ask yourself: What is the most essential part of the analysis? Can you significantly simplify the proofs without trivializing the results? Because of the page limit, you should avoid presenting every step in the analysis in your written report. Focus on conveying the logical flow and key steps of the theoretical analysis.

  Students are encouraged to perform simple simulations to test the proposed algorithm(s) in the paper and report their findings in the report, if the paper includes experiments. Code snippets and additional simulation results can be provided in the appendix (after the list of references if any) and do not contribute to the page limit, but the instructor does not have an obligation to read the appendix. The simulations of the method(s) in the paper can be conducted using small, toy, and synthetic data created by your own to showcase the convergent behavior of the algorithm(s) and their baselines. Students do not need to precisely follow the setting nor use the same data as described in the paper. Some papers are purely theoretical without any numerical experiments. In this case, students are still encouraged to create their own synthetic data or come up with a simulation scenario to verify the theoretical results. However, the evaluation of the written report will be focused on **whether you can present the motivation, the algorithm, the idea behind the algorithm design, and reproducing/summarizing the analysis clearly**. Simulations are optional and will only be considered for the assessment if the other aforementioned components of the written report are not strong.

  The due date of the written report is **Thursday, Jun 13, 2024, 11:59 pm**.

- **Additional Tasks (Optional)** include (1) identifying the limitations and proposing a way to overcome them (theoretically and/or empirically), (2) or improving the results in the paper, such as obtaining a tighter bound or modifying the algorithms to be faster (theoretically and/or empirically), (3) or proposing a new algorithm in the same setting. Final projects which include an additional task with a **solid** result may receive extra credits. If an additional task is conducted, students should highlight what they have accomplished in the main text of the written report and may provide further detail in the appendix after Page 6, but the instructor does not have an obligation to read the appendix.

Final presentation for the project: The session for in-person oral presentations is scheduled for **Monday, Jun 10, 2024, from 15:00 pm to 18:00 pm** (the final exam time). If you do not plan to attend the final presentation in person, please record your video presentation (recommended using Zoom) and upload/share the Zoom link under Canvas/Assignment no later than **Thursday, Jun 13, 2024, 11:59 PM**. Each team has 10 minutes (subject to change) for their presentation and 1 minute (subject to change) for the Q&A. A team can earn extra credits in their oral presentation by asking a good question or making a constructive remark on the presentations of other teams during the Q&A. Similarly, a team can earn extra credits by answering questions well during the Q&A. So, you are encouraged to attend the final presentaion.

Given the limited time available for each oral presentation, each team should avoid walking through the theoretical analysis in the slides presentation. Instead, the oral presentation should focus on giving a clear high-level picture and key messages of the paper, as if you were the authors aiming to deliver a great **lightning** talk to draw attention from your audience so that they would be interested in reading your paper offline.

The first couple of minutes/slides should be allocated to describe the setup as clear as possible and/or provide necessary background so that the audience can be on the same page. Subsequently, the main results of the paper should be clearly summarized. If you think you have completed an additional task listed in Additional Tasks above, make sure to allocate at least one-third of time to highlight what you have accomplished.

Checkpoint report of the Final Project: In the checkpoint report, students are required to include the following information:

- The name of each team member.

- The title of the paper that you will focus on.

It is emphasized that **every member** in the team should submit a checkpoint report. The checkpoint report is due on **May 17 Friday at 11:59 PM**.

Scribes: Starting from the second lecture, each lecture will have two to three students to scribe. This group of students will be responsible for taking detailed notes in class, and the quality of these notes will be graded. There will be a sign-up sheet for students to sign up to scribe for a preferred date.

Students should use the provided scribe template available on Canvas (also available on Overleaf `https://www.overleaf.com/4346235867pssnjnxszjtr#b6a6ff`) and email a typed PDF file **with** the LaTeX code within 6 days to the TA. The TA will review the scribe and may request students to make updates if any part of the notes is unclear or contains errors. It is expected that students revise the scribe accordingly within a couple of days. Students will receive no more than one request if there are issues in the scribe. If students do not address the issues or errors in the scribe timely, then TA may update the scribe; in this case, students may not receive full credit.

Similar courses taught at other universities:
1. Convex Optimization and Approximation at UC Berkeley by Moritz Hardt. `https://ee227c.github.io/`
2. Large-Scale Optimization for Data Science at University of Pennsylvania by Yuxin Chen. `https://yuxinchen2020.github.io/large_scale_optimization/index.html`

3. Optimization for Machine Learning at Boston University by Ashok Cutkosky. `https://optmlclass.github.io/`
4. Introduction to Optimization Theory at Standford by Aaron Sidford. `https://web.stanford.edu/~sidford/courses/19fa_opt_theory/fa19_opt_theory.html`
5. Convex Optimization at CMU by Ryan Tibshirani. `https://www.stat.cmu.edu/~ryantibs/convexopt/`

Tentative Schedule

Week 1	Apr 2	L1: Course logistics, Mathematical background	HW 1 out
	Apr 4	L2: Gradient Flow and Convex Analysis I	
Week 2	Apr 9	L3: Gradient Descent	
	Apr 11	L4: Reduction	
Week 3	Apr 16	L5: Convex Analysis II	HW 1 due; HW 2 out
	Apr 18	L6: Projected Gradient Descent and Frank-Wolfe Method	
Week 4	Apr 23	L7: SGD	
	Apr 25	L8: SGD and Variance Reduction	HW 2 due; HW 3 out
Week 5	Apr 30	L9: Duality Theory I: Lagrangian and Weak Duality	
	May 2	L10: Duality Theory II: KKT conditions and Strong Duality	
Week 6	May 7	L11: Duality Theory III: Conjugate functions and SDCA	HW 3 due; HW 4 out
	May 9	L12: Mirror Descent	
Week 7	May 14	L13: Online Convex Optimization	
	May 16	L14: Online Convex Optimization	HW 4 due
Week 8	May 21	L15: Min-Max Optimization	
	May 23	Midterm	
Week 9	May 28	L16: Min-Max Optimization	HW 5 out
	May 30	L17: Acceleration via Chebyshev Polynomial	
Week 10	Jun 4	L18: Optimization as a game	
	Jun 6	L19: Optimization as a game	HW 5 due

Textbooks: There is no required textbooks. The course will draw some of the materials from the following references. All the following materials are accessible online or accessible via UCSD library online portal. The pointer to the references for each lecture will be clearly given.
[1] "Convex Optimization: Algorithms and Complexity" by Sebastien Bubeck.
`https://arxiv.org/pdf/1405.4980.pdf`
[2] "Convex Analysis and Nonsmooth Optimization" by Dmitriy Drusvyatskiy `https://sites.math.washington.edu/~ddrusv/crs/Math_516_2021/bookwithindex.pdf`
[3] "Algorithms for Convex Optimization" by Nisheeth K. Vishnoi.
`https://convex-optimization.github.io/`
[4] "A Modern Introduction to Online Learning" by Francesco Orabona.
`https://arxiv.org/pdf/1912.13213.pdf`
[5] "Optimization Algorithms" by Aaron Sidford.
`https://drive.google.com/file/d/1BfMkt2glaZpJGwg7gwsJw9T_XxH3o8gx/view`
[6] "Convex Optimization" by Stephen Boyd and Lieven Vandenberghe.
`https://web.stanford.edu/~boyd/cvxbook/`
[7] "Techniques in Optimization and Sampling" by Yin Tat Lee and Santosh Vempala.
`https://github.com/YinTat/optimizationbook/blob/main/main.pdf`

[8] "Lectures on Modern Convex Optimization" by Arkadi Nemirovski.
`https://www2.isye.gatech.edu/~nemirovs/LMCOLN2022Fall.pdf`
[9] "Introduction to Online Convex Optimization" by Elad Hazan.
[10] "Introductory Lectures on Stochastic Convex Optimization" by John Duchi `https://stanford.edu/~jduchi/PCMIConvex/Duchi16.pdf`
[11] "Acceleration Methods" by Alexandre d'Aspremont, Damien Scieur, Adrien Taylor.
`https://arxiv.org/abs/2101.09545`
[12] "Optimization for Modern Data Analysis" by Benjamin Recht and Stephen J. Wright.
[13] "First-order and Stochastic Optimization Methods for Machine Learning" by Guanghui Lan.
[14] "Introductory Lectures on Convex Optimization" by Yurii Nesterov,
[15] "Introduction to Nonlinear Optimization - Theory, Algorithms and Applications" by Amir Beck.