Stats 102B Lec 1: Syllabus

Introduction to Computation and Optimization for Statistics
UCLA, Summer 2024

Course Description

Introduction to computational methods and optimization useful for statisticians. Use of computer programming to solve statistical problems. Topics include vector/matrix computation, multivariate normal distribution, principal component analysis, clustering analysis, gradient-based optimization, EM algorithm for missing data, and dynamic programming.

Course Staff

Instructor: Michael Tsiang
Pronouns: He/Him
Email: mike@stat.ucla.edu
Office Hours: Mondays and Wednesdays 4:15–5:30pm
Physical Office: Math Sciences 8105J
Virtual (Zoom) Office: Meeting ID: 579 439 842, Passcode: 982973
`https://ucla.zoom.us/j/579439842?pwd=TE1oTjJ0NFJTSzdKcUtJNHRjbjJ3UT09`

Teaching Assistants: Abhishek Devarajan and Yingqi Gao
Email: abhishekkd@g.ucla.edu, yqg36@g.ucla.edu
Office Hours: Abhishek: Thursdays 12–1pm, on Zoom
`https://ucla.zoom.us/j/6446488096`
Yingqi: Tuesdays 5–6pm, on Zoom
`https://ucla.zoom.us/j/95635772557`

Course Materials

Course Website: `https://bruinlearn.ucla.edu/courses/188773`

Discussion Forum: `https://campuswire.com/p/G1FCBC400` (Class Code 7507)

We will be using Campuswire for all class Q&A and discussions. Please use it to ask any class-related questions. You are encouraged to discuss and collaborate with each other and answer each other's questions. The instructor and TAs will check Campuswire to answer unresolved questions.

Prerequisites: Stats 100B or Math 170S, Stats 102A, and Math 33A

Required Textbook: None. Course notes and any other required readings will be provided on Bruin Learn.

Reference (Optional) Textbooks:

- Chong, E., Lu, W. S., and Żak, S., An Introduction to Optimization: With Applications to Machine Learning, 5th edition, Wiley, 2023.
- Deisenroth, M. P., Faisal, A. A., and Ong, C. S., Mathematics for Machine Learning, Cambridge University Press, 2020.
- Rogers, S. and Girolami, M., A First Course in Machine Learning, 2nd edition, Chapman & Hall, 2017.

- Watt, J., Borhani, R., and Katsaggelos, A. K., Machine Learning Refined, 2nd edition, Cambridge University Press, 2020.

Whine and Cheese Club: Every other Friday (Weeks 1, 3, and 5), 7:30–9pm Pacific
Physical Location: Math Sciences 8105
Virtual (Zoom) Location: Meeting ID: 967 5963 5259, Passcode: 908362
`https://ucla.zoom.us/j/96759635259?pwd=UHVhY042VWF0VzhvTStsMkxuZ2hWQT09`

The Whine and Cheese Club is intended to be an informal meeting place for community building and an open discussion about the topic du jour. Stop by to ask questions (related or unrelated to course content), vent about the week, get to know your classmates/instructor, or just come listen. Turning video on is encouraged but not required.

Assignments and Evaluation

Grading Policy: Your grade will be based on the following components:

- 33% Homework
- 30% Midterm (Friday of Week 3, August 23)
- 35% Final exam (Friday of Week 6, September 13)
- 2% Campuswire Participation

If everyone does well, everyone will get a good grade. Grades *start from* an absolute scale (90% is an A, 80% is a B, etc.), but the cutoffs can be flexible to accommodate distributions that do not fit that very well. Both absolute scale (e.g., raw percentage) and relative standing (e.g., z-scores and percentiles) are considered when determining final grades. Plus and minus cutoffs will be determined at the end of the quarter.

Homework:

There will be weekly homework assignments posted on the course website. Please include your first and last name and student ID on every homework assignment.

Exams:

There will be one midterm and a final exam. Further details of the exams will be announced closer to the exam dates.

The exams will be open for a 24 hour period on the specified day, but you will have a limited amount of time once you start the exam. You will have two (2) hours to complete the midterm and three (3) hours to complete the final exam.

Communicate with the instructor immediately if you foresee any difficulty in taking the exams at the scheduled time. Not taking the final exam without approval by the instructor will result in an F in the course.

Campuswire Participation:

Campuswire has a reputation system that rewards students for asking thoughtful questions, having engaging discussions, and helping out other students by answering their questions. Participation on Campuswire accounts for 2 percentage points of your final grade. Campuswire's reputation report will be downloaded on an arbitrary day of Week 6. Credit will be assigned as follows:

- Students who remain at N00b level (baby chick) will get 0 points.
- Students who achieve Starter level (yellow bird) will get 0.5 points.
- Students who achieve Starter level AND earn 50 reputation points will get 1 points.
- Students who achieve Intermediate level (red bird) AND 100 reputation points will get 2 points (full credit).
- Students who achieve Advanced level (eagle) will get 3 points (full credit plus 1 extra credit point).

Low quality posts may be removed at the instructor's discretion and will not count towards a student's reputation score. A blatant abuse of the point system may result in the report being downloaded earlier and/or credit being cancelled.

When creating your Campuswire account, it is important to enter your first and last name as they appear on Bruin Learn. If you have a preferred name, email the instructor so that credit can be assigned properly.

Course Policies

Mask Policy:

While not currently mandated, properly wearing a face mask (covering the nose and mouth) is **strongly recommended** during all class related sessions (lecture, discussion, or office hours). Should any COVID-19 related policies change during the quarter, the course policy will always follow and comply with the latest University and LA County requirements. Non-compliance of any and all COVID-19 health and safety mitigation measures set by the University is a violation of the UCLA Student Conduct Code and may be reported to the Dean of Students.

Copyright Policy:

All course materials (lecture slides, assignments, exam questions, etc.) provided by the instructor or TAs are intended for personal use only for students who are enrolled in Stats 102B Lec 1 in Summer 2024. **Do not post, share, or distribute any portion of any course materials to anyone or anywhere, either electronically or physically, without explicit written consent by the instructor, *even after the quarter is over*.** Failure to comply is a violation of academic integrity and copyright infringement against the author(s) of the course materials.

Recording Policy:

To respect and protect the right to privacy of the instructor, the TAs, and each student, students are strictly prohibited from any and all photographic, audio, and/or video recording by any means (e.g., phone, camera, screenshot, etc.) in all class related sessions (lecture, discussion, or office hours), whether in-person, pre-recorded, or livestreamed. The lectures will be recorded (either by the instructor or through Bruincast) solely for use by currently enrolled students and may not be downloaded or shared.

Email Policy:

Any email to the instructor or TAs must include your enrolled lecture and discussion section number (e.g., "Stats 102B Disc 1A") in the subject heading (in addition to the subject of the email). Any emails without this information may be ignored without warning. Do not send messages through Canvas.

Please allow at least 24 hours (more on weekends and holidays) to expect a reply before sending a follow up email.

All homework assignments must be turned in through the respective submission portals on Bruin Learn. Email submissions will not be accepted: **Do not attach assignments to email**.

Assignment Policies:

There will be a 24-hour grace period for late submissions of homework. Exceptions to the late penalty are considered on a case-by-case basis to accommodate extenuating circumstances.

Submissions must be reasonably presentable, properly organized, and easily readable to the grader. All R code is expected to follow the Tidyverse Style Guide. Any submissions that are difficult for the grader to read will receive a penalty for style/readability.

Collaboration Policy:

Limited collaboration will be allowed in solving homework problems, but each student must submit independently written solutions that adhere to the university policy on academic integrity. **Searching for, finding, and/or using solutions online or from previous students is considered academically dishonest.** The full collaboration policy and an academic integrity agreement will be posted on Bruin Learn. **No work will be accepted or graded until the academic integrity agreement is signed and uploaded to Bruin Learn.**

Regrade Policy:

Requests for regrades on homework should be sent to your TA, who will liaise with the grader in addressing the grading concern. Homework submissions will be regraded in their entirety, and any grade adjustments (whether an increase or decrease) will be considered final.

Course Grade Changes:

It is the student's responsibility to check grades on Bruin Learn in a timely manner so that any grade issues (e.g., missing or incorrect grades) are resolved well before the quarter is over. All grades are considered final 24 hours after the final exam and cannot be appealed.

After course grades have been submitted to the Registrar, grades are final. Grade changes will only be considered if there has been a clerical or procedural mistake. Students have one quarter to make requests for a grade change. Graded exams and other materials will be kept for one quarter. After one quarter, course grade changes will not be made.

University and Departmental Policies

Academic Integrity: As a student and member of the University community, you are here to get an education and are, therefore, expected to demonstrate integrity in your academic endeavors. All students must uphold University of California Standards of Student Conduct as administered by the Office of the Dean of Students. Students are subject to disciplinary action for several types of misconduct, including but not limited to: cheating, multiple submissions, plagiarism, prohibited collaboration, facilitating academic dishonesty, or knowingly furnishing false information. For more information about academic integrity, please go to `http://www.deanofstudents.ucla.edu/`.

In addition, each student is the sole owner of their own work and/or code and must NOT:

- Submit work that is not original.
- Publish solutions or code online.
- Post the course questions on forums other than the designated course discussion forum. This means students cannot post questions on places like Stack Overflow, Chegg, ChatGPT, or other similar places.

- Submit someone else's work, or a modification of that work, with or without that person's knowledge.
- Allow someone else to submit their work, or a modification of their work.
- Contract course work out to others.
- Plan or execute with another student some form of cheating during an exam.
- Make use of unauthorized material during an exam.

Zero Tolerance Policy: Any and all issues of potential academic dishonesty will be reported to the Dean of Students without warning after the end of the quarter.

Accessible Education: If you have a disability that will require academic accommodation, please contact the UCLA Center for Accessible Education (CAE). Please contact the CAE as soon as possible to allow for sufficient time to coordinate accommodations.

Title IX: Title IX prohibits gender discrimination, including sexual harassment, domestic and dating violence, sexual assault, and stalking. If you have experienced sexual harassment or sexual violence, you can receive confidential support and advocacy at the CARE Advocacy Office for Sexual and Gender-Based Violence, 1st Floor Wooden Center West, CAREadvocate@caps.ucla.edu, (310) 206-2465. In addition, Counseling and Psychological Services (CAPS) provides confidential counseling to all students and can be reached 24/7 at (310) 825-0768. You can also report sexual violence or sexual harassment directly to the University's Title IX Coordinator, 2241 Murphy Hall, titleix@conet.ucla.edu, (310) 206-3417. Reports to law enforcement can be made to UCPD at (310) 825-1491. Faculty and TAs are required under the UC Policy on Sexual Violence and Sexual Harassment to inform the Title IX Coordinator should they become aware that you or any other student has experienced sexual violence or sexual harassment.

Undergraduate Mentoring: I am an undergraduate mentor for the Department of Statistics and Data Science. This means that you may visit me during office hours (or by appointment) for an informal meeting where you can ask questions unrelated to course content, and about research opportunities, graduate studies, career paths, or any other topic pertinent to your education. You are welcome to visit and meet even after the course ends.

Miscellaneous Advice

1. **Focus on the learning, not on the grade.** You are not a student at UCLA simply to get letters on a transcript. The point of this course is to introduce fundamental tools and concepts in statistical optimization and machine learning so that you are not floundering on the basics in later courses or applications when there are higher stakes and larger consequences. My main hope for you is eventually to become a strong and self-sufficient statistician/data scientist. I know that that takes more than one short quarter to happen. What is ultimately more important is to have a willingness to learn, a growth mindset, and integrity in everything you do.

   No one is born knowing statistical programming, data science, statistics, or any other thing that you may feel "comes naturally" to some people and not others. It takes hard work. It takes practice (sometimes a LOT of it). It takes failing and trying again. It takes asking others for advice. Being comfortable with disappointment and knowing how to deal and respond to it is part of the growth process. It is hard, it often sucks, but it is ultimately necessary to become the best you can be.

2. **Ask for help.** Post to the discussion forum and attend office hours if you have questions or concerns. The teaching team holds many office hours spread out over the week to be available and accessible to you. Please let the instructor know if there are time conflicts with these.

3. **No grade in any class, including this one, is more important than your physical well-being, your mental well-being, and your integrity.** Take time to rest, eat, exercise, go for a walk, hang out with friends, speak to a counselor, or do whatever it takes to take care of yourself. It might not feel like it, but you are not alone. Be mindful of others' struggles as well.

   `https://stand.ucla.edu/tips`

   "We need to remember what's important in life: friends, waffles, work. Or waffles, friends, work. Doesn't matter, but work is third." – Leslie Knope (from Parks and Recreation)

(Tentative) Possible Topics Covered

1. Basic Concepts in Mathematical Optimization

   - Global versus local optimization
   - Zero-order, first-order, second-order optimization methods
   - Convexity

2. Algorithms in Optimization

   - Gradient Descent
   - The BFGS Algorithm
   - Conjugate Gradient Method

3. Optimization in Statistics and Machine Learning

   - Regression Methods (least squares, ridge regression)
   - Classification Methods (KNN)
   - Clustering Methods (k-means, Gaussian mixture models)
   - Principal Component Analysis
   - Expectation-Maximization (EM) Algorithm