MATH118 MWF 9:00–9:50 AM, Room: https://ucla.zoom.us/j/94572326850

Instructor: Daniel Mckenzie, mckenzie@math.ucla.edu.

Teaching Assistant: Denali Molitor, dmolitor@math.ucla.edu.

Office Hours:

- Daniel Mckenzie: TBD.
- Denali Molitor: TBD.

Textbooks:

1. Elden, Lars. *Matrix Methods in Data Mining and Pattern Recognition.* The Society for Industrial and Applied Mathematics, 2019.

2. Chong, E. and Zak, S. *An Introduction to Optimization, 4th Edition.* Wiley, 2013.

3. Bandeira, A. and Singer, A. and Strohmer, T. *Mathematics of Data Science, draft.* Available at https://people.math.ethz.ch/~abandeira/BandeiraSingerStrohmer-MDS-draft.pdf

The first two books will be used extensively. The third book is a work-in-progress, but has very good exposition of certain topics and will take you deeper into the mathematical theory of data science. Some background in probability and statistics will be very useful. If you have not yet taken a course in it, I recommend glancing over the first few chapters of *Practical Statistics for Data Scientists* (https://www.oreilly.com/library/view/practical-statistics-for/9781491952955/)

Prerequisite: Math 115A and some experience with scientific computing in MATLAB or Python. Note that on CCLE Math 42 is listed as a prerequisite, but this is not essential.

Course Objectives: This course will cover the key techniques from linear algebra and optimization that underlie much of modern data science. While there will be an emphasis on theory, we will always develop it with an eye towards applications. Theoretical analysis of algorithms will be complemented with programming exercises in Python.

Communication: This class will use Piazza for classroom discussion. If you have questions about the course content (*e.g.* "Where do I start with question 14 on the homework?" or "Can someone explain that thing we did at the end of class yesterday?") or of an administrative nature (*e.g.* "When is the second midterm?"), post them here. The sign up link is https://piazza.com/ucla/winter2021/math118/home. We (*i.e.* your instructor and T.A.) will be checking Piazza regularly, and will try our absolute best to ensure that all questions are answered within 24 hours. If you

have questions about your grades, or other questions/comments of a personal nature, you should contact me or your T.A. via email. Office hours or immediately after class are also good times to ask questions.

Tentative Schedule		
Week	**Date**	**Sections**
1	Mon 01/04	Introduction, relation of this course to optimization, algorithms etc. Finite precision arithmetic, big O notation and computational complexity. Ethical implications of Data Science.
	Wed 01/06	Vectors and matrices. Notation. Inner products, norms etc. Outer products. *(Elden Chpt 1–2)*
	Fri 01/08	Linear Systems. Least Squares. LU decomposition. Condition number. Matrix Least squares.*(Elden Chpt. 3)*
2	Mon 01/11	Orthogonal matrices, QR factorization, application to least squares, pre-recorded. *(Elden Chpt. 4–5)*
	Wed 01/13	Singular Value Decomposition. Interpretation. Condition number for least squares. *(Elden Chpt. 6 and Blum Chpt. 3)*
	Fri 01/15	PCA and applications. *(Elden Chpt. 6 and Bandeira Chpt. 3)*
3	**Mon 01/18**	**MLK day, no class**
	Wed 01/20	PCA and applications. *(Elden Chpt. 6 and Bandeira Chpt. 3)*
	Fri 01/22	Non-neg. matrix factorization. Text preprocessing using tf-idf. Topic modeling. *(Elden Chpt 9 & 12)*
4	Mon 01/25	Tensors and CP decomposition. *Based on* `https://arxiv.org/pdf/1711.10781.pdf` *and* `https://www.kolda.net/talk/2020-07-23-minds/`
	Wed 01/27	Graphs and Spectral Clustering.*Elden Chpt. 10*
	Fri 01/29	Spectral Clustering for Euclidean data. *Elden Chpt. 9*
5	Mon 02/01	PageRank. *Elden Chpt. 13*
	Wed 02/03	Review
	Fri 02/05	**Midterm**

6	Mon 02/08	Introduction to optimization. Convex Functions. Gradient Descent (GD) and Newton's method. *C. & Z. Chpt 6,8 & 9*
	Wed 02/10	Convergence analysis for GD.
	Fri 02/12	Constrained Optimization. Projected GD. Penalty methods. Regularization. *C. & Z. Chpt. 23*
7	**Mon 02/15**	**Presidents day. No class**
	Wed 02/17	Introduction to linear programming *C. & Z. Chpt. 15*
	Fri 02/19	The simplex method for linear programs *C. & Z. Chpt. 16*
8	Mon 02/22	Introduction to Dynamic Programming. Sequence alignment.
	Wed 02/24	Dynamic Programming. Markov Processes, Hidden Markov Models. Viterbi's algorithm
	Fri 02/26	Dynamic Programming. Reinforcement learning and the Bellman equation.
9	Mon 03/01	Introduction to supervised learning. Logistic Regression and SVM.
	Wed 03/03	Two layer neural networks. Backpropagation.
	Fri 03/05	Two layer neural networks.
10	Mon 03/08	Further topics on neural networks.
	Wed 03/10	Ethics, privacy and fairness in Data Science.
	Fri 03/12	Review/ Catch-up day.
11	**Mon 03/15**	**Final Exam**

Grading: Your grade will be determined using the following formula: 45% Homework + 25% Midterm + 30% Final exam. You may earn up to 2% extra credit by asking and answering questions on Piazza (see: "Communication").

Homework: There will be four homework assignments for this course, with due dates:

- Homework 1: Friday, January 15th.
- Homework 2: Friday, January 29th.

- Homework 3: Friday, February 19th.
- Homework 4: Friday, March 5th. *(earlier due to Thanksgiving break)*

You are to submit your homework using GradeScope. We strongly encourage you to use word processing software such as LaTeX to type your solutions. The penalty for late homework is 10% of your score per day after the due date. Homework problems will be a mix of pen-and-paper calculations and programming exercises. You are encouraged to discuss homework questions on Piazza or in person. However copying other's solutions or programs is considered a serious violation, and you should never share your written solutions with anyone.

Software: I strongly recommend you use Python for all programming exercises.

Exams: There will be one in-class midterm exam and a cumulative final. For both, you will have a 24 hour window to complete and submit the exam. The midterm will contain the same amount of content as a 1 hour in-person exam. The final will contain the same amount of content as a 3 hour in-person exam. Make-ups for the final and midterm are permitted only under exceptional circumstances, as outlined in the UCLA student handbook. **Note that you must take the final exam in order to pass this class.** The exams are scheduled for the following dates:

- Midterm: Friday, February 5th.
- Final Exam: Monday, March 15th.

Note that both exams will be of the pen-and-paper variety, but you may be asked to write pseudocode or implement, by hand, a few steps of an algorithm on a toy example.

Classroom Expectations

- Slides and Jupyter notebooks for all lectures will be posted on Piazza, **at least 24 hours before the lecture.** I encourage you to download them before class, so that you can skim over them, and so that you can add your own annotations to your copy during class. All lectures will be recorded and available on CCLE. You are expected to attend lectures and discussion sections if you are able. If you are in a different time zones, the expectation is that you watch the recordings of all lectures and discussion sections.
- On exams I expect you to give legible, well-justified solutions. A numerical answer without any supporting calculations is unlikely to receive credit, nor is a solution that I can't read.
- Standard Zoom etiquette applies. For example, mute your mic unless you have a question. You are encouraged, but not required to turn your cameras on.
- In this classroom you can expect to be treated with respect, regardless of your age, background, beliefs, ethnicity, gender, gender identity, gender expression, national origin, religious affiliation, sexual orientation and other visible and non-visible differences. All members of this class are expected to contribute to a respectful, welcoming and inclusive environment for every other member of the class.

How to succeed in this course

- Prepare for class by by looking over the slides or reading ahead in the textbook. Spending even 15 minutes skimming a section before it is taught in class can be useful!
- Office hours are time that we (your instructor and T.A.'s) set aside in order to discuss math with you. You can drop in at any time during office hours. They are a great time to discuss: homework problems you are stuck on, things you didn't quite understand in class, strategies for preparing for or taking tests or going over tests after they have been returned. We may also be able to help you debug your code
- Forming a study group can be very helpful. Remember that often the best way to fully understand a concept is to try to explain it to someone else.
- Make use of Piazza! (See "Communication" section). Again, answering other students questions is a great way to make sure you really understand a concept.

Academic Integrity: From the office of the Dean of Students:

> *With its status as a world-class research institution, it is critical that the University uphold the highest standards of integrity both inside and outside the classroom. As a student and member of the UCLA community, you are expected to demonstrate integrity in all of your academic endeavors. Accordingly, when accusations of academic dishonesty occur, The Office of the Dean of Students is charged with investigating and adjudicating suspected violations. Academic dishonesty, includes, but is not limited to, cheating, fabrication, plagiarism, multiple submissions or facilitating academic misconduct.*

Students are expected to be aware of the University policy on academic integrity in the UCLA Student Conduct Code*. In particular, please note the sections on (1) cheating, (2) plagiarism, and (3) unauthorized study aids.

Accessibility: Students needing academic accommodations based on a disability should contact the Center for Accessible Education (CAE) at (310)825-1501 or in person at Murphy Hall A255. In order to ensure accommodations, students need to contact the CAE within the first two weeks of term.

Notice about sexual harassment, discrimination and assault: Title IX prohibits gender discrimination, including sexual harassment, domestic and dating violence, sexual assault and stalking. Students who have experienced sexual harassment or sexual violence can receive **confidential** support and advocacy from a CARE advocate:

> The CARE Advocacy Office for Sexual and Gender-Based Violence
> 1st Floor, Wooden Center West
> `CAREadvocate@caps.ucla.edu`
> (310) 206-2465

You can also report sexual violence or sexual harassment directly to the University's Title IX Coordinator:

*`http://www.deanofstudents.ucla.edu/Portals/16/Documents/UCLACodeOfConduct_Rev030416.pdf`

Kathleen Salvaty
2241 Murphy Hall
titleix@conet.ucla.edu
(310) 206-3417