STA 243 (Spring 2022)
Computational Statistics

Classes:	M (10-11:40AM) & W (10-10:50AM)	**Room:**	TLC 2218
Instructor:	Krishna Balasubramanian	**E-mail:**	kbala@ucdavis.edu
Office Hours:	M 4:30-6:00PM	**Office:**	Zoom
Discussion:	Asynchronous	**Room:**	Zoom
Teaching Assistant:	Tesi Xiao	**E-mail:**	texiao@ucdavis.edu
Office Hours:	W 4:30-6:00PM	**Office:**	Zoom

Prerequisite: **Linear Algebra** and **mathematical maturity** are absolutely essential for this course. There will be a review on Linear Algebra during the first week, but you are expected to have taken a full course on Linear Algebra before. Also brush up you probability and statistics background. Please be aware that this is graduate-level course and hence, **mathematical maturity** at the graduate-level is expected. **If you are uncomfortable with the prerequisites, you are encouraged to drop this course and take suggested replacement courses in your curriculum.**

Course Objectives: The goal of this course will be to introduce computational and statistical techniques for analysis of big-data. The main emphasize of this course is **not** on providing a list of computational algorithms for statistical problem, and their implementations in a programming language. Rather, the core emphasize is on **understanding the main principles** behind the computational algorithms used in statistics. The primary purpose will be to introduce the interplay between statistics and computation. Students will be expected to understand and apply the techniques to real-life problems.

Text: Notes (and/or related online materials) will be provided. Apart from that, the following books are *highly* recommended although not necessary:

- *Convex Optimization Algorithms* by Dimitri P. Bertsekas, 2015.
- *Monte Carlo Statistical Methods* by Christian Robert and George Casella, 2004.

Privacy: **The course lectures and course materials, including notes, lecture capture videos, tests, outlines, and similar materials, are protected by U.S. copyright law and by University policy. The instructor is the exclusive owner of the copyright of those materials. You may take notes and make copies of course materials for your own use. You may also share those materials with another student who is enrolled in or auditing this course. You may not reproduce, distribute or display (post/upload) lecture notes or recordings or course materials in any other way–whether or not a fee is charged–without the instructor's prior written consent. You also may not allow others to do so. If you do so, you may be subject to student conduct proceedings under the UC Davis Code of Academic Conduct.**

Content: The course will cover: (1) Randomized Linear Algebra, (2) Convex and Nonconvex Optimization including stochastic settings, and (3) Sampling. Detailed schedule (tentative) in provided in page 3.

Homeworks: There will be **4** homework assignments. You can team-up with a different partner for each homework. Homeworks involve both derivation based problems and implementations of algorithms. For the derivations, you must use LaTeX. Online compilers like Overleaf are extremely easy to use and require no installation. Code should be turned in along with the homework. Late homework submissions will **not be accepted** unless you present a well-documented case of a medical or family emergency. Copying of homework, is absolutely forbidden and constitutes a violation of the Code of Academic Conduct. You are encouraged to ask the instructor or the TA for help on homework problems after you have tried to solve the problems on your own. Please respect your TA's schedule and do not hand in assignments late. TAs will make **no special allowances** to accept late assignments.

Project: There will be a **final group project**. The group size will be set at the end of the second week. You will be required to summarize a paper or papers. There are two options: (1) you can use papers from the list provided around middle of the quarter or (2) you can select papers of your choice and get it approved with the instructor. (Note: For option (2), the paper has to be sufficiently technical and relevant. Otherwise, it will not be approved.) You will be required to submit a **8-page summary** of the paper (references excluded and with unlimited supplementary pages to include any details that you wish to include). This will mainly involve **understanding and writing the methodological/theoretical contribution of the paper(s), in your own words**. You are encouraged to implement the algorithm proposed in the paper for supporting your understanding. However, **only simulations are allowed. Real-world datasets are not allowed**. The purpose of this restriction is so that you carefully design your simulation setup to provide meaningful insights on the performance of the algorithm. **Please keep in mind that if you add simulations that are not carefully thought-out (and are provided to just fill up space), you will end up losing points.**

Language: We will use **Python** as the coding language for this course. The second discussion will be introduction to Python.

Grading: Homework will count for 60% of the total points. Each of the 4 homeworks will count for 20% of the total points, but the homework with the least points will be dropped. Final project score will used for the remaining 40%.

Disability Service: UC Davis is committed to educational equity in the academic setting, and in serving a diverse student body. I encourage all students who are interested in learning more about the Student Disability Center (SDC) to contact them directly at sdc.ucdavis.edu, sdc@ucdavis.edu or 530-752-3184.

If you are a student who currently receives academic accommodation(s), please submit your SDC Letter of Accommodation to me as soon as possible, ideally within the first two weeks of this course.

Code of Conduct: There will be **zero tolerance** regarding any form of academic dishonesty like copying homework from any online sources or from classmates. Such dishonesty will be immediately reported and appropriate action will be taken. Please take a look at `https://ossja.ucdavis.edu/code-academic-conduct` for more details.

Mental Health: It is important that you take care of yourself and do your best in maintaining a healthy lifestyle. This will help you achieve your goals and cope with stress. In the event that you need help, please visit `https://shcs.ucdavis.edu/`, or `https://shcs.ucdavis.edu/about/how-get-help-when-were-closed` or call (530) 752-2349 for further information.

Date	Topic	HW (tentative)
03/28	Linear Algebra: Basics	
03/30	Linear Algebra: Basics	
04/04	Randomized Linear Algebra: Matrix Multiplication	
04/06	Randomized Linear Algebra: Matrix Multiplication	HW1 Out
04/11*	Power Method	
04/13*	Least-squares and Sketching	
04/18	Convex Opt	
04/20	Convex Opt	
04/25	Convex Opt	HW2 Out
04/27	Convex Opt	
05/02	Individual project meeting	
05/04	Convex Opt	
05/09	Non-Convex Opt: EM	HW3 Out
05/11	Non-convex Opt	
05/16	Sampling	
05/18*	Sampling	
05/23*	Sampling	HW4 Out
05/25*	Sampling	
05/30	University Holiday	
06/01	Sampling	

* These lectures will be asynchronous (not in-person) as I will be traveling.