Math 191, Spring 2023
Numerical Analysis for Data Science and Statistics

Instructor: Prof. Jon Wilkening
Office: 1051 Evans Hall
Office Hours: Tues 3:45-4:45 PM (both classes), Fri 3:15-4 PM (Math 191), Fri 4-4:45 PM (Math 228B)
e-mail: wilkening@berkeley.edu (emergencies & administration only. No questions about HW, please)
online discussion forum for our class: edstem.org
GSI: Jiahao Yao, (jiahao@math.berkeley.edu), 935 Evans, Fri 9-11 (228B), Fri 10-noon (Math 191)

Course Announcements, Homework Solutions, etc.: https://bcourses.berkeley.edu/

Lectures: MWF 2:10-3:00 PM, 60 Evans

Required textbook: "Linear Algebra and Learning from Data," by Gilbert Strang.
Recommended reading: "Accuracy and Stability of Numerical Algorithms," by Nicholas J. Higham
"Applied Numerical Linear Algebra," James W. Demmel

Prerequisites: Math 53 and 54 or equivalent (e.g., Math 91 from Fall 2022 can replace Math 54)

Syllabus: Introduction to applied linear algebra and optimization with applications in data science. We will cover Parts I, II, III, VI, VII (and a few subtopics of IV, V) of Strang's book, as well as chapters 2 and 3 of Higham's book and Chapter 2 of Demmel's book. In more detail:

- Floating-point arithmetic, condition number, perturbation theory, backward stability analysis
- Matrix decompositions (LU/QR/Cholesky/SVD), least squares problems, orthogonal matrices
- Eigenvalues, eigenvectors, Rayleigh quotients, generalized eigenvalues
- Principal components, low rank approximation, compressed sensing, matrix completion
- Convexity, Newton's method, Levenberg-Marquardt method, quasi-Newton methods
- Randomized linear algebra, stochastic gradient descent
- Machine learning, neural networks (deep/convolution), adjoint methods, backpropagation

Possible additional topics (time permitting):

- Discrete Fourier transform, circulant and Toeplitz matrices, clustering by spectral methods
- Multivariate Gaussian and weighted least squares, Kalman filter

Grades: Programming assignments: 20% (due Feb 15, Mar 15, Apr 12, May 3)
Homework: 20%. 11 assignments, 2 lowest scores dropped
Midterms 1 and 2: 15% each. (Friday, Feb 24; Friday, April 7)
Final: 30% or 45% (Tues, May 9, 11:30-2:30. Can replace a midterm score with the score on the final if helpful.) *No make-up exams for any reason... don't miss the final exam!*

Grade cutoffs: 98 A+, 90 A, 86 A–, 82 B+, 78 B, 74 B–, 70 C+, 66 C, 62 C–, 58 D+, 50 D (no D- given)
(raw scores on exams will be mapped to scaled scores, keeping these cutoffs in mind. The scaled score will never be lower than your raw score expressed as a percentage)

Incomplete grades: (University policy) A grade of I will only be given if "your work in a course has been of passing quality but is incomplete for reasons beyond your control"

More details: 11 homework assignments, 4 programming assignments. You may discuss the homework and programming assignments with your classmates, but **you must write up your own solutions.** The two lowest homework scores will be dropped. **You must complete all 4 programming assignments to pass the course.** (The grade cutoffs only apply if you complete all four.) Matlab or Python are suggested languages for the programming assignments and many of the homework problems, but if you prefer another programming language, just let me know.