Math 156, Fall 2023
Numerical Analysis for Data Science and Statistics

Instructor: Prof. Jon Wilkening
Office: 1051 Evans Hall
Office Hours: Tues 3-4 PM, Fri 3:30-4:30 PM
e-mail: wilkening@berkeley.edu (emergencies & administration only. No questions about HW, please)
online discussion forum for our class: edstem.org , GSI/grader: TBA
Course Announcements, Homework Solutions, etc.: https://bcourses.berkeley.edu/

Lectures: Tues/Thurs 12:40-2:00 PM, 3109 Etcheverry

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
"Numerical Optimization," Jorge Nocedal and Stephen J. Wright

Prerequisites: Multivariable Calculus (Math 53) and Linear Algebra (Math 54 or 56)

Syllabus: Introduction to applied linear algebra and optimization with applications in data science. We will cover Parts I, II, V, VI (and a few subtopics of III, VII) of Strang's book, as well as Higham (chap. 2-3), Demmel (chap. 2), Nocedal/Wright (chap. 2-6). In more detail:

- Floating-point arithmetic, relative and absolute error, running error analysis
- Vector spaces (real/complex), bases, dimension, subspaces, linear operators, rank-nullity theorem
- Inner products, adjoint operator, orthogonal projections, closest point property
- Orthogonal and unitary matrices, QR via Gram-Schmidt or Householder, LAPACK implementation
- Eigenvalues and eigenvectors, singular value decomposition (SVD)
- Gaussian elimination, LU factorization, positive definite matrices, Cholesky factorization
- Least squares via QR or the SVD, pseudo inverse, polar decomposition
- Operator norms (1,2,infinity,nuclear norms), linear functionals, Hahn-Banach theorem
- Perturbation theory, condition number, backward stability analysis
- Low rank approximation, Eckart-Young theorem (2-norm, Frobenius norm)
- Discrete probability, sample mean and covariance, principal component analysis
- Advanced calculus review (Jacobian, gradient, Hessian), multivariable Newton's method
- Nonlinear least squares, trust-region framework, Levenberg-Marquardt method
- Line-search methods, Wolfe conditions, conjugate gradients, GMRES, BFGS
- Rank-revealing QR, sparse solutions, deep neural networks, adjoint methods, backpropagation

Grades: Programming assignments: 20% (due Sep 13, Oct 14, Nov 11, Dec 5)
Homework: 20%. 10 assignments, 2 lowest scores dropped
Midterm: 25% (Thursday, Oct 12)
Final: 35% or 60% (Friday, Dec 15, 8-11 AM. Can replace 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: 10 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. Python will be the official programming language for the course (submitted to gradescope via a Jupyter notebook.)