Numerical optimization and Large Scale Linear Algebra

Paris Vassalos, Assistant Professor, AUEB, pvassal@aueb.gr

Office: Pattision 76, Antoniadou Wing, 5rd floor, Tel: +30-210-8203187, Email: pvassal@aueb.gr

Overview

Numerical linear algebra is of great practical importance in scientific computation and is used in mathematics, natural sciences, computer science and social science. Even nonlinear problems usually involve linear algebra in their solution.□□This high level course provides in-depth knowledge of matrix computations. Learn the numerical solutions of systems of linear equations including direct methods, error analysis, structured matrices as well as iterative methods and least squares. Explore applications in industry including direct implications for Internet applications. In addition, one of the important objectives of this course is to use computers for solving scientific problems. Assignments can be done in PYTHON and specifically using SCI-PY.

Key Outcomes

By completing the course the students will learn:

1. The basic matrix factorization methods for solving systems of linear equations and linear least square problems.
2. Basic computer arithmetic and the concepts of conditioning and stability of a numerical method.
3. The basic numerical methods for computing eigenvalues.
4. The basic iterative methods for solving systems of linear equations.
5. Spectral methods for data analysis and compression.
6. How to implement the above numerical methods in Python (or another similar software package).

Requirements and Prerequisites

Elementary Linear Algebra, and matrix theory. Notions of analysis, numerical algorithms.

Books

There are many books on the subject, and a lot of free resources on the Internet; the following selection provides a good foundation:

- Trefethen, Lloyd N., and David Bau. *Numerical Linear Algebra*. Philadelphia, PA: Society for Industrial and Applied Mathematics, May 1, 1997. ISBN: 0898713617.
- Demmel, James W. *Applied Numerical Linear Algebra.* Philadelphia, PA: Society for Industrial and Applied Mathematics, September 1, 1997. ISBN: 0898713897.
- Strang, Gilbert. Linear algebra and its applications. Academic Press, New York, 2nd edition, 1980. ISBN 0-12-673660.

- Kutz, Nathan. Data-Driven Modeling & Scientific Computing. Oxford University Press. ISBN 9780199660346.

Grading

Homework (computational) assigments: 50%

Final examination: 50%

This is a practical course; students will be graded on their competency to work with data in realistic problems and to show practical results. There will be three course assignments at two unit intervals.

1. The first assignment will be announced after the first three lectures and will count 10% towards the final grade.
2. The second assignment will be announced after the sixth lecture and will count 10% towards the final grade.
3. The third assignment will announced after the eighth lecture and will count 10% towards the final grade.
4. The main course project will be announced at the sixth unit and will count 20% towards the final grade.

Course Syllabus

The course comprises ten units of three hours each.

Unit 1: Introduction to Scientific computing

Approximation in Scientific Computation, computer arithmetic and errors, sensitivity and conditioning. Matrices in large-scale scientific data analysis.

Unit 2: Curve fitting

Least-square fitting methods. Polynomial fits and splines. Least squares fitting with trigonometric functions. Beyond least-square fitting: The L1 norm. Case study: Scattered data interpolation in computer graphics

Unit 3: Linear Systems (Direct Methods)

Structured and unstructured linear systems. The basic matrix decompositions (LU, Cholesky, QR, QD, SVD) and their variations. Stability analysis. Some special matrices. Case study: Estimation of the price of a financial derivative via Black-Scholes equation.

Unit 4: Linear Systems (Iterative Methods)

The classical iterative methods (Jacobi, Gauss-Seidel, SOR). Convergence analysis. Krylov subspace methods. Preconditioning. Case study: Solving a problem arising from Markov chains and queuing theory.

Unit 5: The SVD analysis and its applications.

The Decomposition. Fundamental Subspaces. Matrix Approximation. Principal Component Analysis. Truncated SVD: Principal Component Regression. Independent component analysis. Case study: I Inferring Population Structure with PCA

Unit 6: Least squares Problem.

Linear and nonlinear models. Periodic data. Data linearization. Solving Least Squares Problems. Rank-Deficient and Underdetermined Systems. Damped least-squares and Tikhonov regularization. Reduced-Rank Least Squares Models. Nonlinear least squares. Gauss-Newton method. Models with nonlinear coefficients. Case study: Image registration.

Unit 7: Spectral methods

Signal reconstruction and circumventing Nyquist. Data (image) reconstruction from sparse sampling. Fourier Transform. The Discrete Cosine Transform and signal/image compression. Quantization. Huffman coding. The Modified Cosine Transform. Case study: spectral methods for detecting periodicity in big data sets.

Unit 8: Tensor Decomposition.

Basic Tensor Concepts. A Tensor SVD. Approximating a Tensor by HOSVD. Clustering and Nonnegative Matrix Factorization. The k-Means Algorithm. Nonnegative Matrix Factorization. Case study: TensorSVD: an approach to personalized Web search.

Unit 9: Randomized algorithms for matrices and data.

Randomization applied to matrix problems. Randomized algorithms for least-squares approximation. Randomized algorithms for low-rank matrix approximation. Empirical observations. Reality check: Study the population stratification.

Unit 10: Numerical Methods For Large Eigenvalue Problems.

Single Vector Iterations: The Power Method and its variations (Shifted Power Method, Inverse Iteration). Convergence analysis. Deflation with Several Vectors. General Projection methods. Krylov subspace methods: Arnoldi, Lanczos Case study: Page rank.

Participation

In-class contribution is a key instructional activity with many benefits, including:

- Active learning, resulting in better application of course material
- Knowledge gained from the shared perspectives and experiences of classmates
- Opportunities to integrate research and readings into discussions, resulting in deeper exploration and reflection of course topics
- Providing students a platform to express ideas and share prior learning experiences

In order for students to capitalize on these benefits, consistent and quality study of the past lectures is necessary.

Please, plan to arrive to class on time and to stay for the entire class period (or until dismissed). Random arrivals and exits are unprofessional, disrespectful and distracting. Repeated tardiness will have an impact on your grade.

Attendance Requirements

Class attendance is essential to succeed in this course and is part of your grade. An excused absence can only be granted in cases of serious illness or grave family emergencies and must be documented. Job interviews and incompatible travel plans are considered unexcused absences. Where possible, please notify the instructor in advance of an excused absence. Students are responsible for keeping up with the course material, including lectures, from the first day of this class, forward. It is the student's obligation to bring oneself up to date on any missed coursework.

Assignments

Late assignments will either not be accepted or will incur a grade penalty unless due to documented serious illness or family emergency. Exceptions to this policy for reasons of civic obligations will only be made available when the assignment cannot reasonably be completed prior to the due date, you make suitable arrangements, and give notice for late submission in advance

Code of Ethics

Students may not work together on individual graded assignments unless the instructor gives express permission. They must avoid

- Falsification and fabrication of data
- Plagiarism and other misappropriation of the work of another
- Cheating
- Other forms of academic misconduct that are commonly accepted within the scientific community

Exercise integrity in all aspects of one's academic work including, but not limited to, the preparation and completion of all other course requirements by not engaging in any method or means that provides an unfair advantage. In any case of doubt, students must be able to prove that they are the sole authors of their work by demonstrating their knowledge to the instructor.

Clearly acknowledge the work and efforts of others when submitting written work as one’s own. Ideas, data, direct quotations (which should be designated with quotation marks), paraphrasing, creative expression, or any other incorporation of the work of others should be fully referenced. No plagiarism of any sort will be tolerated. This includes any material found on the internet. Reuse of material found in question and answer forums, code repositories, other lecture sites, etc., is unacceptable. You may use online material to deepen your understanding of a concept, not for finding answers.