CSE 392 – Topics in Computer Science: Matrix and Tensor Algorithms for Data

Instructor: Shashanka Ubaru

Course Description

Advances in modern technologies have resulted in huge volumes of data being generated in several scientific, industrial, and social domains. With ever increasing size of data comes the necessity to develop fast and scalable machine learning and data algorithms to process and analyze them. In this course, we study the mathematical foundations of large-scale data processing, with focus on designing algorithms and learning to (theoretically) analyze them. We explore randomized numerical linear algebra (sketching and sampling) and tensor methods for processing and analyzing large-scale multidimensional data, graphs, and data-streams. We will also have presentations on the linear algebra concepts of quantum computing.

Prerequisites

The minimum requirements for the course are basics concepts of probability, algorithms, and linear algebra. Knowledge and experience with machine learning algorithms will be helpful. For the course, we will rely most heavily on probability, linear and tensor algebra, but we will also learn few concepts related to approximation theory, high dimensional geometry, and quantum computing. The course will involve rigorous theoretical analyses and some programming (practical implementation and applications).

Programming language: The programming languages for the course will be *Matlab* and *Python*.

Grading

Grading is based on problem sets, project/presentation, and class participation. There will be no exams. The breakdown is as follows:

- (50%) 4 to 5 assignments each contributing an equal amount to the grade. Assignments will include problem sets and programming exercises.
- (40%) Class project and presentation.
- (10%) Participation in the class.

Resources

There is no official textbook for the class. Course material will mainly consist of lecture notes/slides, along with online resources such as papers, notes from other courses, and publicly available surveys.

Course Schedule

The following schedule is tentative and subject to change.

- Week 1 - **Introduction and basics**
  - Lecture 1: Vector spaces, matrices, norms.
  - Lecture 2: Probability review, concentration of measure.
- Week 2 - **Regression and low rank approximation**
  - Lecture 3: Least squares regression, kernel methods.
  - Lecture 4: Matrix factorizations I - Singular Value Decomposition (SVD), QR factorization.
- Week 3 - **Matrix factorization and Randomized projection**
  - Lecture 5: Matrix factorizations II - eigenvalue decomposition, Principal Component Analysis.
  - Lecture 6: Approximate matrix product, sampling.
- Week 4 - **Randomized Sketching**
  - Lecture 7: Johnson–Lindenstrauss(JL) lemma, subspace embedding.
  - Lecture 8: Sketching, types of sketching matrices.
- Week 5 - **Randomized linear algebra I**
  - Lecture 9: Sketch and solve - least squares regression.
  - Lecture 10: Sampling for least squares, preconditioned LS.
- Week 6 - **Randomized linear algebra II**
  - Lecture 11: Randomized SVD.
  - Lecture 12: Stochastic trace estimation.
- Week 7 - **Iterative methods**
  - Lecture 13 : Subspace iteration (power) method.
  - Lecture 14 : Krylov subspace method.
- Week 8 - **Tensor methods - CP foundations**
  - Lecture 15: Introduction to tensors, tensor-matrix product.
  - Lecture 16: Canonical Polyadic (CP) decomposition.
- Week 9 - **Randomized CP decomposition**
  - Lecture 17: Kronecker Fast JL, randomized CP-ALS.
  - Lecture 18: CP-ALS with leverage scores, Generalized CP.
- Week 10 - **Tucker Decomposition**
  - Lecture 19: Tucker decomposition, HOSVD.
  - Lecture 20: Randomized Tucker.

- Week 11 - **Matrix mimetic tensor algebra I**
  - Lecture 21: Tube-fiber product, t-product.
  - Lecture 22: Tensor-tensor-SVD (t-SVD).
- Week 12 - **Matrix mimetic tensor algebra II**
  - Lecture 23: Randomized t-SVD, t-product applications.
  - Lecture 24: Tensor networks.
- Week 13 - **Quantum computing** (Optional)
  - Lecture 25: Introduction to quantum computing, vector states, Pauli matrices, tensor product,
  - Lecture 26: Quantum circuits and quantum measurements.
- Week 14 - **Presentations**
  - Lecture 27: Project presentations I
  - Lecture 28: Project presentations II

The quantum computing topics (Week 13) above will be discussed as presentations, provided we are able to cover other topics as per schedule. There will not be any problem sets related to these topics.