CAAM 651 · TOPICS IN NUMERICAL LINEAR ALGEBRA

Randomized Matrix Algorithms

Fall 2013 · Rice University

Extremely large linear algebra problems are common to a wide variety of data driven applications. Sources include social networks, genome analysis, term-document data, machine learning, uncertainty quantification and many others. The sheer size, uncertainty, and diversity of sources of massive data make most existing matrix methods useless. Alternative approaches are essential.

This seminar will present an introduction to randomized algorithms for large-scale matrix problems arising in various scenarios involving massive data set analysis. Topics to be covered include: basic theory, decoupling of randomization and deterministic phases, the Johnson–Lindenstrauss lemma, random sampling and projection algorithms.

Specific problem areas include: least-squares regression, low-rank matrix approximation, Principal Component Analysis and approximate SVDs. We will address practical implementation issues and implications for machine learning and statistical applications in large-scale data applications.

In addition to CAAM students, this seminar is appropriate for graduate students and advanced undergraduates in computer science, statistics, mathematics, and electrical engineering, as well as computationally-inclined students from application domains.

Lectures: Thursdays, 4-5:30 pm, Duncan Hall 1075

Web Site: http://www.caam.rice.edu/∼caam651

Instructors: Mark Embree (embree@rice.edu), Duncan Hall 3019, (713) 348-6160
Dan Sorensen (sorensen@rice.edu), Duncan Hall 2035, (713) 348-5193

Text: We shall select topics from two review articles (Halko, Martinsson, Tropp, "Finding structure with randomness: Probabalistic algorithms for constructing approximate matrix decompositions," *SIAM Review*, 2011; Mahoney, "Randomized algorithms for matrices and Data," *FnTML*, 2011), and supplement this material with recent articles and other materials we discover as the course evolves.

Prerequisites: Students should have a solid understanding of numerical linear algebra, an acquaintance with probability theory at the advanced undergraduate level, and proficiency with MATLAB.

Credits: 1 credit: Read, digest, and summarize a recent paper (or part of a paper) in a 5 page report, and give a short presentation.
3 credit: Produce a more extensive review of a recent paper (or set of papers), including illustrative numerical experiments; present this work in a 30 minute presentation.

Grades: Students will be evaluated on their active participation in the course, the thoroughness of their review projects, and the literate presentation of those projects.

Objectives: CAAM 651 students learn special topics at the frontier of numerical linear algebra research.

Outcomes: Upon completing this course, students will have experience reading the recent research literature, analyzing algorithms in numerical linear algebra, and studying implementations in mathematical software.

Any student with a disability requiring accommodation in this course is encouraged to contact to contact the instructors, and also to contact Disability Support Services in the Ley Student Center.