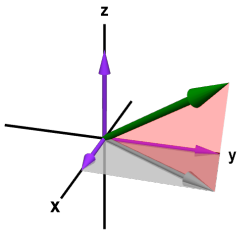
COLUMBIA | ENGINEERING
The Fu Foundation School of Engineering and Applied Science

Computing at Scale with Randomized Numerical Linear Algebra



When: Thursdays 1:10–3:40 , hybrid format
Instructor: James Anderson (`ja3451@columbia.edu`)
Course website: `https://tinyurl.com/e6899`
Who: Graduates with an interest in numerical methods for big data

Course Description

This graduate level course will introduce students to key ideas in randomized numerical linear algebra. We will explore how randomization can be leveraged as a powerful computational tool for developing algorithms with faster run-times, lighter storage overhead, and improved numerical stability, than its well studied deterministic counterpart. Such speedups are essential for both analyzing massive-scale data-sets and as black-box subroutines embedded in classical algorithms. Topics we will cover include:

- Matrix approximation using sampling (approximate multiplication, norm approximation)
- Randomized embeddings (Johnson-Lindenstrauss lemma)
- Low rank approximations (randomized range finder and SVD)
- Optimization at scale: linear regression, semidefinite programming, descent algorithms
- Scalable kernel methods for large-scale machine learning

Organization

The first 3 weeks will be led by the instructor, where we will cover motivating examples, numerical linear algebra, and probability fundamentals. Subsequently, classes will be student led with the instructor providing background material and key concepts as and where needed. A tentative schedule is available on the course website.

Grading

Your final grade will be determined by performance in **homework assignments** (20%), **presentation(s)** (20%), and a **course project** (60%). Homework will comprise of 3–4 problem sets that mix theory with algorithm implementation. You will choose a topic for which you will be responsible for teaching to the class. You will assign background reading, and deliver a lecture on the topic. For the course project, you can focus on theory or applications – almost any application is possible, provided there is a coherent objective and narrative. You will be required to submit a short project report, deliver a conference-style short presentation, and a project proposal. You may work alone or with a partner.

Prerequisites

A strong linear algebra background is essential. Probability and random signals (ELEN 4815 or equivalent). Familiarity with optimization techniques helpful but not essential (e.g. EEOR 4650 or equivalent). Basic coding skills using a scripting language such as MATLAB or Python.