Matrix Computations & Scientific Computing Seminar

Organizer: James Demmel & Ming Gu

Wednesday, 11:00AM–12:00Noon, 380 Soda

Sept. 25   **Michael Mahoney**, Stanford University
Implementing Randomized Matrix Algorithms in Parallel and Distributed Environments

Motivated by problems in large-scale data analysis, randomized algorithms for matrix problems such as regression and low-rank matrix approximation have been the focus of a great deal of attention in recent years. These algorithms exploit novel random sampling and random projection methods; and implementations of these algorithms have already proven superior to traditional state-of-the-art algorithms, as implemented in Lapack and high-quality scientific computing software, for moderately-large problems stored in RAM on a single machine. Here, we describe the extension of these methods to computing high-precision solutions in parallel and distributed environments that are more common in very large-scale data analysis applications. In particular, we consider both the Least Squares Approximation problem and the Least Absolute Deviation problem, and we develop and implement randomized algorithms that take advantage of modern computer architectures in order to achieve improved communication profiles on, e.g., MapReduce and on clusters that have high communication costs such as on an Amazon Elastic Compute Cloud cluster.