Matrix Computations and Scientific Computing Seminar

Organizer: Jim Demmel and Ming Gu

Wednesday, 12:10–1:00 pm, 380 Soda Hall

Nov 5 **Chris Melgaard**, UC Berkeley

Improved Column-Based Matrix Reconstruction

Column-based matrix reconstruction, or the CX decomposition, has become a popular tool in large-scale data analysis for constructing low-rank matrix approximations built using the features (or columns) of the data. In this paper, we present a modified deterministic algorithm for column selection, based on the seminal work of Boutsidis, Drineas and Magdon-Ismail. Our modified algorithm enjoys improved computational complexity, less memory usage and stronger error bounds. We also provide a randomized version for further computational efficiency. Our algorithms offer stronger controls on the spectral and Frobenius norm errors. Additionally, we provide novel rank revealing lower bounds on the individual singular values in the CX decomposition. For matrices with decaying singular values (the typical case with real-world data), our bounds surprisingly suggest that the CX decomposition is capable of accurately retaining the leading singular values even for very few features. Using synthetic and real data matrices, we report numerical tests to reaffirm the efficiency and accuracy of our algorithms.