Faster Subset Selection for Matrices and Applications

Haim Avron, IBM Corporation

We study subset selection for matrices defined as follows: given a matrix $\matX \in \R^{n \times m}$ ($m>n$) and an oversampling parameter k ($n \leq k \leq m$), select a subset of k columns from $\matX$ such that the pseudo-inverse of the subsampled matrix has as smallest norm as possible. In this work, we focus on the Frobenius and the spectral matrix norms. We describe several novel (deterministic and randomized) approximation algorithms for this problem with approximation bounds that are optimal up to constant factors. The study is motivated by the observation that the combinatorial problem of finding a low-stretch spanning tree in an undirected graph corresponds to subset selection. We discuss various implications of this reduction.

Joint work with Christos Boutsidis