Statistical and Computational Limits for Submatrix Localization and Sparse Matrix Detection

Tony Cai

Department of Statistics, University of Pennsylvania, USA
Email: tcai@wharton.upenn.edu

In the conventional statistical framework, the goal is developing optimal inference procedures, where optimality is understood with respect to the sample size and parameter space. When the dimensionality of the data becomes large as in many contemporary applications, the computational concerns associated with the statistical procedures come to the forefront. A fundamental question is: Is there a price to pay for statistical performance if one only considers computable (polynomial-time) procedures? After all, statistical methods are useful in practice only if they can be computed within a reasonable amount of time.

In this talk, we discuss the interplay between statistical accuracy and computational efficiency in two specific problems: submatrix localization and sparse matrix detection based on a noisy observation of a large matrix. The results show some interesting phenomena that are quite different from other high-dimensional problems studied in the literature.