Computational Concerns in Statistical Interference and Learning for Network Data Analysis

Abstract:
Network data analysis has wide applications in computational social science, computational biology, online social media, and data visualization. For many of these network inference problems, the brute-force (yet statistically optimal) methods involve combinatorial optimization, which is computationally prohibitive when we are faced with large scale networks. Therefore, it is important to understand the effect of computational constraints on statistical inference.

In this talk, we will discuss three closely related statistical models for different network inference problems. These models answer inference questions on cliques, communities, and ties, respectively. For each particular model, we will describe the statistical model, propose new computationally efficient algorithms, and study the theoretical properties and numerical performance of the algorithms. Further, we will quantify the computational optimality through describing the intrinsic barrier for certain efficient algorithm classes, and investigate the computational-to-statistical gap theoretically. A key feature shared by our studies is that, as the parameters of the model changes, the problems exhibit different phases of computational difficulty.

Host: Lily Xu
Monday, January 9, 2017
3:00 PM
AP&M 6402