Department of Applied and Computational Mathematics and Statistics Colloquium

Arash Amini
Department of Statistics
University of California, Los Angeles

Optimal Bipartite Network Clustering

The analysis of network data, where the object of study is a graph of edges between a collection of nodes, has been an active area of research for many decades and has seen a substantial expansion in the past 15 years. The interest has been fueled in recent years by the plethora of network data coming from a variety of sources in science, engineering, and society. Among these data are the large-scale and rapidly-growing online social networks, various forms of biological networks (protein and gene interaction networks, gene-regulatory networks, neurological networks, ecological networks, etc.), computer and communication networks, and so on. Many of these networks also exhibit a bipartite structure.

Community detection has emerged in recent years as one of the fundamental problems of network analysis. Informally, one seeks to partition the network into cohesive groups of nodes, or communities, that reveal its large-scale connective structure. In this talk, we explore community detection in the setting of bipartite networks. We consider extensions of the so-called Stochastic Block Model (SBM) to the bipartite setting and show how these models can be efficiently fit using a combination of spectral and likelihood-based approaches. In particular, we show how simple pseudo-likelihood updates can be used to boost the performance of spectral clustering and achieve optimal rates of misclassification. We also present some new results on the nature of these optimal rates in the bipartite setting, sharpening existing estimates.

Monday, October 7, 2019
4:30 PM – 5:30 PM
127 Hayes-Healy Center

Colloquium Tea 4:00 PM to 4:30 PM 101A Crowley Commons Room