PRINCETON UNIVERSITY
OPERATIONS RESEARCH & FINANCIAL ENGINEERING SEMINAR

Nonparametric Graph Estimation

Han Liu
Johns Hopkins University

Abstract

The graphical model has proven to be a useful abstraction in statistics and machine learning. The starting point is the graph of a distribution. While often the graph is assumed given, we are interested in estimating the graph from data. In this talk we present new nonparametric and semiparametric methods for graph estimation. One approach is a nonparametric extension of the Gaussian graphical models that allows arbitrary graphs. Another approach is to restrict the family of allowed graphs to spanning forests, enabling the use of fully nonparametric density estimation in high dimensions. These two approaches can both be viewed as special cases of the more general log-density ANOVA models and reflect an interesting tradeoff of model flexibility with structural complexity. In terms of function estimation, these methods achieve the minimax optimal rates of convergence. In terms of graph estimation, these methods even achieve the optimal parametric rates of convergence. Therefore, the extra flexibility gained by nonparametric modeling comes at very low cost. In terms of computing, we provide currently the most scalable software package that is several times faster than the state-of-the-art softwares implementing the standard parametric methods. The performance of these methods is illustrated and compared on several real and simulated examples.
Joint work with Fang Han, John Lafferty, Larry Wasserman, and Tuo Zhao.

Thursday, February 16, 2012
Sherrerd Hall, Room 101
4:30 PM