Department of Statistics

DISSERTATION PRESENTATION AND DEFENSE

ZHE LIU

Department of Statistics

The University of Chicago

High-Dimensional Graph Estimation and Density Estimation

WEDNESDAY, April 27, 2016, at 11:00 AM

Jones 304, 5747 S. Ellis Avenue

ABSTRACT

Graphical models have become a common tool in many fields and a useful way of modeling probability distributions. In this talk, we investigate approaches for graph estimation and density estimation problems in high dimensions.

We first combine the ideas behind trees and Gaussian graphical models to form a new nonparametric family of graphical models. Our approach is to attach nonparanormal blossoms, with arbitrary graphs, to a collection of nonparametric trees. The tree edges are chosen to connect variables that most violate joint Gaussianity. The non-tree edges are partitioned into disjoint groups, and assigned to tree nodes using a nonparametric partial correlation statistic. A nonparanormal blossom is then grown for each group using established methods. The result is a factorization with respect to the union of the tree branches and blossoms, defining a high-dimensional joint density that can be efficiently estimated and evaluated on test points.

Moreover, we present a nonparametric method for estimating scale-free graphical models. To avoid the usual Gaussian assumption, we restrict the graph to be a forest and build on the work of forest density estimation. The method is motivated from a Bayesian perspective and is equivalent to finding the maximum spanning tree of a weighted graph with a log degree penalty. We solve the optimization problem via a minorize-maximization procedure with Kruskal's algorithm.

Finally, we study robust methods for Gaussian graphical models in the presence of possible outliers and corrupted data. We consider the neighborhood selection and graphical lasso algorithms and show that the robust counterparts obtained using the trimmed inner product or the nonparanormal give stronger performance guarantees.

Our analysis and experimental results indicate that the newly proposed methods can be powerful alternatives to standard approaches for graph estimation and density estimation.

For information about building access for persons with disabilities, please contact Laura Rigazzi at 773.702-0541 or send an email to lrigazzi@galton.uchicago.edu. If you wish to subscribe to our email list, please visit the following web site: https://lists.uchicago.edu/web/arc/statseminars.