THE UNIVERSITY OF CHICAGO

Department of Statistics

MASTER'S THESIS PRESENTATION

SIFU YUE

Department of Statistics
The University of Chicago

Instability and Selection Rate for Gaussian and Elliptical Data

TUESDAY, May 12, at 12:00 PM
Eckhart 110, 5734 S. University Avenue

ABSTRACT

Nowadays, most graph estimation methods rely on the Gaussian graphical model, in which the random vector X is assumed to be Gaussian: $X \sim N(\mu, \Sigma)$ Besides Gaussian models, recent research works propose an extension of Gaussian family named nonparanormal, which is able to deal with the weak normality and seminormality data, and the extension from elliptical distribution transelliptical graphical model, which can handle more heavy tailed data. Also I discuss the three different methods to estimate the precision matrix $\Theta = \Sigma^{-1}$by the CLIME using the Pearson sample correlation, the Kendall's tau and the nonparanormal correlation estimator. In this thesis, I aim to study the variability of the edges selected in the estimated precision matrix across bootstrapped samples, called instability, and to compare this measure with the selection rate, which is the proportion of selected non-diagonal edges. On simulated data, I compare instability measurement with the ROC curve (True Positive Rate and False Positive Rate) given the known precision matrix, which give a traditional way of measuring the performance of each method, but rely on knowing the true graph. On a real dataset, I we cannot calculate the ROC curve, So I explore the three different methods' instability and selection rate instead of the ROC curve to determine which method is more or less appropriate on the dataset.

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.