香港浸會大學
HONG KONG BAPTIST UNIVERSITY

Department of Mathematics
Institute of Computational and Theoretical Studies
Statistics Research and Consultancy Centre

Distinguished Lecture Series

Latent Graphical Model for Mixed Data

Professor Jianqing Fan

Professor of Statistics
Frederick L. Moore '18 Professor of Finance
Chair, Operations Research and Financial Engineering, Princeton University
Fellow of American Association for the Advancement of Science
Fellow of Institute of Mathematical Statistics
Fellow of American Statistical Association
COPSS Presidents' Award, 2000
Humboldt Research Award, 2006
Morningside Gold Medal of Applied Mathematics, 2007
Fellow of Guggenheim, 2009
Academician from Academia Sinica, 2012
Pao-Lu Hsu Prize, 2013
Guy Medal in Silver, 2014
Past President of the Institute of Mathematical Statistics, 2006-2009
Past President of the International Chinese Statistical Association, 2008-2010

Date: **20 January 2015 (Tuesday)**

Time: **11:30 am - 12:30 pm** (Preceded by Reception at 11:00 am)

Venue: **1/F Shiu Pong Hall, Ho Sin Hang Campus, Hong Kong Baptist University**

Abstract

Graphical models are commonly used tools for modeling multivariate random variables. While there exist many convenient multivariate distributions such as Gaussian distribution for continuous data, mixed data with the presence of discrete variables or a combination of both continuous and discrete variables poses new challenges in statistical modeling. In this paper, we propose a semiparametric model named latent Gaussian copula model for binary and mixed data. The observed binary data are assumed to be obtained by dichotomizing a latent variable satisfying the Gaussian copula distribution or the nonparanormal distribution. The latent Gaussian model with the assumption that the latent variables are multivariate Gaussian is a special case of the proposed model. A novel rank-based approach is proposed for both latent graph estimation and latent principal component analysis. Theoretically, the proposed methods achieve the same rates of convergence for both precision matrix estimation and eigenvector estimation, as if the latent variables were observed. Under similar conditions, the consistency of graph structure recovery and feature selection for leading eigenvectors is established. The performance of the proposed methods is numerically assessed through simulation studies, and the usage of our methods is illustrated by a genetic dataset.

This is a joint work with Han Liu, Yang Ning, and Hui Zou.

All are welcome

For enquires please contact Ms. Claudia Chui, 3411 2348.

http://www.math.hkbu.edu.hk/