

综合报告会!

Lecture

国家数学与交叉科学中心

Time: 10:00am, August 24

Venue: Hall

Sparse Modeling for High Dimensional Data



Speaker: Prof. Bin Yu
University of California at Berkeley

Abstract:

Information technology has enabled the collection of massive amounts of data in science, engineering, social science, finance and beyond. Statistics is the science of data and indispensable for extracting useful information from high-dimensional data. After broad successes of statistical machine learning on prediction through regularization, interpretability is gaining attention and sparsity is used as its proxy. With the virtues of both regularization and sparsity, L1 penalized Least Squares (e.g. Lasso) has been intensively studied by researchers from statistics, applied mathematics and signal processing. Lasso is a special case of sparse modeling and has also been the focus of compressive sensing lately. In this talk, I would like to give an overview of both theory and practice of Lasso and its extensions. First, I will review theoretical results of Lasso and present an insightful unified M-estimation theory with decomposable penalties under sparse high dimensional statistical models. Second, I will present on-going collaborative research with the Gallant Neuroscience Lab at Berkeley on human understanding visual pathway. In particular, we estimate meaningful nonlinear sparse models to relate natural images to fMRI responses in human primary visual cortex area V1 with non-trivial improvements on prediction and decoding over linear sparse models.

Brief CV :

Bin Yu is Professor of Statistics and Electrical Engineering & Computer Science at UC Berkeley. She is currently the chair of department of Statistics, and a founding co-director of the Microsoft Lab on Statistics and Information Technology at Peking University, China. Her Ph.D. is in Statistics from UC Berkeley, and she has published over 80 papers in a wide range of research areas including empirical process theory, information theory (MDL), MCMC methods, signal processing, machine learning, high dimensional data inference (boosting and Lasso and sparse modeling in general), bioinformatics, and remote sensing. Her current research interests include statistical machine learning for high dimensional data and solving data problems from remote sensing, neuroscience, and text documents. She was a 2006 Guggenheim Fellow, and is a Fellow of AAAS (American Association for the Advancement of Science), IEEE, IMS (Institute of Mathematical Statistics) and ASA (American Statistical Association). She is a co-chair of National Scientific Committee of SAMSI (an US NSF center for statistical and mathematical sciences) and on the Board of Mathematical Sciences and Applications of the US National Academy of Sciences. She was a Chancellor's Distinguished Professor from 2006 to 2011 at UC Berkeley, and a Chang-Jiang Visiting Professor at Peking University since 2005.