



Temple University

ANNOUNCES A
COLLOQUIUM

Dr. Runze Li

*Department of Statistics
Penn State University*

will speak on

Feature Screening via Distance Correlation Learning

Time: 3:00 – 4:00 PM

Date: Friday, September 16, 2011

Place: Alter Hall 746

Abstract

This paper is concerned with screening features in ultrahigh dimensional data analysis, which has become increasingly important in diverse scientific fields. We develop a sure independence screening procedure based on the distance correlation (DC-SIS, for short). The DC-SIS can be implemented as easily as the sure independence screening procedure based on the Pearson correlation (SIS, for short) proposed by Fan and Lv (2008). However, the DC-SIS can significantly improve the SIS. Fan and Lv (2008) established the sure screening property for the SIS based on linear models, but the sure screening property is valid for the DC-SIS under more general settings including linear models. Furthermore, the implementation of the DC-SIS does not require model specification (e.g., linear model or generalized linear model) for responses or predictors. This is a very appealing property in ultrahigh dimensional data analysis. Moreover, the DC-SIS can be used directly to screen grouped predictor variables and for multivariate response variables. We study the theoretic properties of the DC-SIS, establish its ranking consistency and sure screening properties, and conduct simulations to examine its finite sample performance. Numerical comparison indicates that the DC-SIS performs much better than the SIS in various models. We illustrate the DC-SIS through a real data example.