Variable selection beyond linear regression model

by
Wenxuan Zhong
University of Illinois at Urbana Champaign
725 South Wright Street, Champaign, IL 61820
ADDRESS HERE
wenxuan@illinois.edu

Abstract

In this talk, I will present the Correlation Pursuit (COP) method, which can be viewed as a generalization of the conventional stepwise regression method from linear models to semi-parametric regression models. Unlike the conventional stepwise, COP selects variables that maximize the correlation between a transformed response and a linear combination of the predictors. A sequential selection strategy is used to select variables on multiple linear combinations. Asymptotic properties of the COP algorithm will be discussed both theoretically and empirically. The variable selection consistency for COP algorithm will also be addressed under certain conditions. A usage of the COP procedure in helping solve a gene regulation problem based on gene sequence and microarray expression data will be discussed.