Concave group selection in high-dimensional models

Jian Huang

Department of Statistics and Actuarial Science
University of Iowa, Iowa City, IA 52242, USA

Abstract

Grouping structures arise naturally in many high-dimensional data analysis problems. In this talk, we present a class of group selection methods that respect such structures in parameter estimation and variable selection. These methods use the composition of a concave function and the Euclidean norm of the coefficients in each group as the penalty for selection and estimation. Under certain sparsity and regularity conditions, they possess an oracle property, meaning that with high probability they yield solutions that are equal to the oracle estimator under the unknown true model. This result holds even when the number of groups exceeds the sample size. We derive a group coordinate descent algorithm for computing the solution paths of group estimators. This algorithm takes advantage of the closed form expressions of the estimators for a single group model and is efficient in high-dimensional settings. We also discuss the applications of group selection in several statistical modeling and data analysis problems.