A Model Averaging Approach for High-Dimensional Regression

Tomohiro Ando [(1)*] and Li Ker-Chau [(2)]

[(1)]Keio University, Japan
andoh@kbs.keio.ac.jp
[(2)] University of California Los Angeles
kcli@stat.ucla.edu

Abstract

We considered high-dimensional regression problems where the number of regressors exceeds the sample size. A new regression model averaging procedure was developed. Unlike most works in variable selection which concerns mainly about the identification of the true predictors, this paper focused on the accuracy in predicting the true conditional mean of response given the set of predictors. Our method consists of two steps. The first step is to construct a class of regression models, each with a smaller number of regressors to avoid the degeneracy of the information matrix. The second step is to find suitable model weights for averaging. Theoretical results were derived to justify our procedure. Simulation studies were conducted to illustrate the merits of the proposed method over the existing competitors. Our method is computationally feasible even there are thousands of covariates.

Keywords: Asymptotic optimality; High-dimensional regression models.