Abstract No:IV-1-1

Model Averaging for High-dimensional Linear Regression Models with Dependent Observations

蔡恆修[1], 銀慶剛[2], 余定宏[3]

[1]Academia Sinica, [2]National Tsing Hua University, [3] University of Iowa

Abstract

We introduce the orthogonal greedy algorithm (OGA) to screen out the nested set of signal variables under a high-dimensional linear regression framework with dependent observations. To gain the prediction performance, we propose the high-dimensional Mallow model averaging (HDMMA) criteria to determine the weight for averaging these nested high-dimensional linear regression models. We further analyze rates of convergence of prediction error for the averaging model under different sparsity conditions. Our contribution has three folds. First, we show that our procedure, named OGA+HDMMA, can achieve optimal convergence rates of prediction error discussed in Ing (2019). Second, we use simulation to show that the out-sample prediction of OGA+HDMMA can outperform the MCV method proposed in Ando and Li (2014) when the covariates are highly correlated or contain time-series effects. Third, the out-sample prediction of OGA+HDMMA performs comparably or even better than many well-known high-dimensional variables selection methods in some scenarios.

Keyword: High-dimensional Mallow model averaging, orthogonal greedy algorithm, sparsity conditions, high-dimensional linear regression models, optimal rate of convergence