Alena Skolkova
CERGE-EI

Model Averaging with Ridge Regularization

This paper contributes to the literature on model uncertainty and model averaging for prediction problems. When a model for determination of a specific variable is not precisely dictated by theory, one often faces a trade-off between parsimony of a model with few variables and sophisticatedness of a model with potentially high-dimensional sets of predictors. While a parsimonious model delivers the estimates with a low variance and large bias, a sophisticated model tends to do exactly the opposite. Therefore, combining models with different number of variables generally reduces the mean squared error of resulting predictions. Many methods for finding the optimal combination exist. A leading one is based on generalization of the Mallows (1973) model selection criterion to the Mallows model averaging criterion by B. Hansen (Hansen, 2007).

I propose a ridge-regularized Mallows model averaging estimator. The ridge model averaging estimator (RMA) ensures better finite-sample properties via ridge regularization of the design matrices corresponding to the models being averaged. In principle, ridge regression and model averaging serve a similar purpose, minimization of the mean squared error through shrinkage, though in different ways. While model averaging, e.g. as in Hansen (2007), reduces the asymptotic mean squared error, ridge regularization leads to finite-sample improvements. Therefore, combining model averaging with ridge regularization leads to an estimator that inherits asymptotic optimality, and, in addition, yields better finite-sample properties due to ridge regularization.

Specifically, I suggest ridge-based modifications of both Mallows model averaging (Hansen, 2007) and heteroskedasticity-robust Mallow model averaging (Liu and Okui, 2013). A tractable theoretical example with two models demonstrates that the relative reduction of the mean squared error is increasing with the strength of predictor correlatedness. Via a simulation study, I examine the finite-sample improvements obtained by replacing ordinary least-squares with a ridge regression for model averaging prediction. Ridge-based model averaging is shown to be superior when one deals with sets of moderately to highly correlated predictors as underlying ridge regressions accommodate correlated predictors without blowing up estimation variance. In addition, I also show the superiority of the ridge-regularized estimator modifications via empirical examples focused on wages and economic growth.