Calibrated Shrinkage Ridge Estimation for High Dimensional Data Analysis

Xiaoli Gao, Oakland University and S. Ejaz Ahmed, Brock University

sahmed5@brocku.ca

In high-dimensional data settings where p >> n, many penalized regularization approaches were studied for simultaneous variable selection and estimation. However, with the existence of covariates with weak effect, many existing variable selection methods, including Lasso and its generations, cannot distinguish covariates with weak and no contribution. Thus, prediction based on a subset model of selected covariates only can be inefficient. In this talk, we propose a high-dimensional shrinkage estimation strategy to improve the prediction performance of a subset model. Such a high-dimensional (positive) shrinkage estimator is constructed by shrinking a weighted ridge estimator in the direction of a pre-defined candidate subset. Under an asymptotic distributional quadratic risk criterion, its prediction performance is explored analytically. We show that the proposed high-dimensional (positive) shrinkage estimator performs better than the weighted ridge estimator. More importantly, it improves the prediction performance of any candidate subset model generated from most existing Lasso-type variable selection methods significantly. The relative performance of the proposed high-dimensional shrinkage strategy is demonstrated by both simulation studies and the real data analysis.

Key words: Asymptotic risk, Lasso, ridge regression, shrinkage estimation, sparse model.