Big data analysis: Big biases

In high-dimensional data settings where the number of variables is greater than observations, many penalized regularization approaches were studied for simultaneous variable selection and estimation. With the existence of covariates with weak effect, however, many existing variable selection methods, including Lasso and its generations, cannot distinguish covariates with weak and no contribution. Thus, prediction based on a submodel of selected covariates may not be trustworthy. In this talk, a high-dimensional shrinkage estimation strategy to improve the prediction performance of a submodel is proposed. Such a high-dimensional shrinkage estimator (HDSE) is constructed by shrinking a weighted ridge estimator in the direction of a predefined candidate submodel. Under an asymptotic distributional quadratic risk criterion, its prediction performance is explored analytically. It is shown that the proposed HDSE performs better than the weighted ridge estimator. More importantly, it significantly improves the prediction performance of any candidate submodel generated from most existing Lasso-type variable selection methods. The relative performance of the proposed HDS strategy is demonstrated by both simulation studies and the real data analysis.

Biography

S Ejaz Ahmed is Professor and Dean of Faculty of Maths and Science. Before joining Brock, he was a Professor and Head of Mathematics at the University of Windsor and University of Regina. Prior to that, he had a Faculty position at the University of Western Ontario. He is a Fellow of American Statistical Association. His area of expertise includes statistical inference, shrinkage estimation, big data analysis and asymptotic theory. He has published more than 130 articles in scientific journals, both collaborative and methodological. Further, he has written several books, co-edited two volumes and organized many invited sessions and workshops. He made several invited scholarly presentations in various countries. He serves on the editorial board of many statistical journals and book review editor for Technometrics. He served as a Board of Director and Chairman of the Education Committee of the Statistical Society of Canada. He was VP communication for ISBIS. He served as a member of an Evaluation Group, Discovery Grants and the Grant Selection Committee, Natural Sciences and Engineering Research Council of Canada.

sahmed5@brocku.ca