

Abstract: High dimensional data analysis has become a major issue in a range of different scientific fields. Wide big data is often characterized by a huge number of predictor variables measured on a limited number of observations. Moreover, variables are often correlated following complicated patterns. In these situations, it turns out to be difficult to find a method that provides an optimum between few computational operations and high statistical reliability. Tuning of hyper-parameters in regularized regression tends to be difficult in the frequentist paradigm, and Bayesian methods are often computationally intractable because of the need to iteratively cycle through a large number of variables. The first part of the talk compares some recent Bayesian methods in terms of statistical efficiency with focus on variable selection. Secondly, a new two-stage approach denoted SIS-BEN is presented. It is shown that this method is computationally efficient and provides very low error rates regarding variable selection.