2010 – 04

Minimizing Variable Selection Criteria by Markov Chain Monte Carlo

Ting-Li Chen, Yen-Shiu Chin

August 16, 2010

Minimizing Variable Selection Criteria by Markov Chain Monte Carlo

Ting-Li Chen[a,*], Yen-Shiu Chin[a]

[a]*Institute of Statistical Science, Academia Sinica, Taipei 115, Taiwan*

Abstract

Regression models with a large number of predictors arise in diverse fields of social sciences and natural sciences. For their proper interpretation, we often would like to identify a smaller subset of the variables that shows the strongest information. In such large size of candidate predictors setting, one will encounter a computationally cumbersome search in practice by optimizing some criteria for selecting variables, such as AIC, C_p and BIC, over all possible subsets. Instead of the forward stepwise selection, backward stepwise selection or Lasso method, we present an efficient algorithm via Markov chain Monte Carlo (MCMC) approach for searching the optimum value of the corresponding criterion. We will exhibit that our proposed MCMC algorithm obtains better solutions in terms of minimizing those criteria on simulated examples as well as a real data set.

Keywords: Variable selection; Markov chain Monte Carlo method; Lasso

*Corresponding author

Email address: tlchen@stat.sinica.edu.tw (Ting-Li Chen)