

An Algorithm for Bayesian Variable Selection in High-dimensional Generalized Linear Models

Vitara Pungpapong

Department of Statistics, Faculty of Commerce and Accountancy
Chulalongkorn University, Bangkok, Thailand vitara@cbs.chula.ac.th

Inspired by analysis of genomic data, the primary quest is to identify associations between studied traits and genetic markers where number of markers is typically much larger than sample size. Bayesian variable selection methods with Markov chain Monte Carlo (MCMC) are extensively applied to analyze such high-dimensional data. However, MCMC is often slow to converge with large number of candidate predictors. In this talk, we examine the empirical Bayes variable selection with a sparse prior on the unknown coefficients. An iterated conditional modes/medians (ICM/M) algorithm is proposed for implementation by iteratively minimizing a conditional loss function in high-dimensional linear regression model. Attention is then directed to extend the algorithm to a generalized linear model. The performances of our approach along with some of the competitors are evaluated through simulation study.

Key Words: Bayesian inference, high-dimensional data, sparse variables, generalized linear models