An EM Algorithm for Bayesian Variable Selection

Jin Wang, Feng Liang*
University of Illinois at Urbana-Champaign, IL, USA – e-mail address

Yuan Ji
University of Chicago, Chicago, IL, USA – e-mail address

Motivated by a recent work by Rockova and George (2014), we propose an efficient EM algorithm for Bayesian variable selection, which returns the MAP estimate of the relevant variable subset. An appealing feature of our EM algorithm is that we do not need to calculate the inverse of a large matrix in each iteration, which seems unavoidable in many other algorithms. Although our algorithm is shown to achieve variable selection consistency asymptotically, in practice, it could get stuck at a local optimal due to the large model space. Therefore we propose an ensemble version of our algorithm based on Bayesian bootstrap, where we repeatedly apply our EM algorithm on a subset of the samples with a subset of the covariates, and then aggregate the variable selection results across those bootstrap replicates. In addition to improved accuracy, the BB (Bayesian bootstrap) EM approach is also faster than the original EM, especially in the large $p$ small $n$ scenario, since only a subset of the covariates is considered in each bootstrap replicate. Empirical studies have shown that the performance of BBEM is competitive in comparison with other methods.

Keywords: Bayesian bootstrap; variable selection; EM algorithm.