

Bayesian sparse group selection

Ray-Bing Chen¹ (陳瑞彬), Chi-Hsiang Chu^{2*} (朱基祥)
and Ying Nian Wu³ (吳英年)

¹Department of Statistics, National Cheng-Kung University

²Institute of Statistics, National University of Kaohsiung

³Department of Statistics, University of California, Los Angeles

Abstract

This article proposes a Bayesian approach for the sparse group selection problem in regression models. In this problem, the variables or regressors are partitioned into different groups. It is assumed that only a small number of groups are active or important for explaining the response variable. It is further assumed that within each active group only a small number of variables are active. We adopt a Bayesian hierarchical formulation, where each candidate group is associated with a binary variable indicating whether the group is active or not. Within each group, each candidate variable is also associated with a binary variable, indicating whether the variable is selected or not. In this Bayesian formulation, the sparse group selection problem can be solved by sampling from the posterior distribution of the two layers of indicator variables as well as the coefficients of the selected variables. We adopt a group-wise Gibbs sampler for posterior sampling. We demonstrate the proposed method by simulation studies as well as a real example. The simulation results show that the proposed method is competitive with the sparse group Lasso in terms of selecting the active groups as well as identifying the active variables within the selected groups.

Keywords: Bayesian variable selection, group-wise Gibbs sampler, group sparsity, median probability criterion