Stochastic Matching Pursuit for Bayesian Variable Selection

Ray-Bing Chen, National University of Kaohsiung

Chi-Hsiang Chu, National University of Kaohsiung

Te-You Lai, National University of Kaohsiung

Ying Nian Wu, University of California, Los Angeles

Abstract

This article proposes a stochastic version of the matching pursuit algorithm for Bayesian variable selection in linear regression models. In the Bayesian formulation, the prior distribution of each coefficient is assumed to be a mixture of a point mass at 0 and a normal distribution with zero mean and a large variance. The proposed stochastic matching pursuit algorithm is designed for sampling from the posterior distribution of the coefficients for the purpose of variable selection. The proposed algorithm combines the efficiency of the matching pursuit algorithm and the rigorous Bayesian formulation with well defined prior distributions on coefficients. The algorithm is a Metropolis scheme with a pair of reversible moves. One is the addition move, which adds a new variable into the existing set of selected variables, where the variables with larger correlations with the residuals are assigned higher probabilities of being added, in a fashion that is very similar to the original matching pursuit algorithm. The other move is the deletion move, which deletes a variable from the existing set of selected variables. Several simulated and real examples for cases of large n small p and small n large p are used to illustrate the proposed algorithm. These examples show that the algorithm is efficient in screening and selecting variables.