Bayesian Variable Selection with Related Predictors using the Gibbs Sampler*

V.A. Delouille

September 1998

Abstract

In data sets with many predictors, algorithms for identifying a good subset of predictors are often used. One approach that is being used by an increasing number of people is the stochastic search variable selection. This procedure uses probabilistic considerations for selecting promising subsets. A latent vector serves to identify one of two types of regression coefficients: those close to 0 and those not, thus identifying subset choices. In this framework the promising subsets of predictors can be identified as those with higher posterior probability. In this dissertation, a given stochastic search variable selection algorithm is modified to treat efficiently regression models where there exists some relationships between predictors. Simulations are carried out in order to assess the performance of the modified algorithm and a practical data set is analysed.

*Thesis submitted to the University of Kent at Canterbury in the subject of Statistics for a degree of Master of Science (M.Sc.)