AN EFFICIENT METHOD FOR VARIABLE SELECTION IN LINEAR AND NON-LINEAR MODELS

ARNAB KUMAR MAITY AND SANJIB BASU
DIVISION OF STATISTICS
NORTHERN ILLINOIS UNIVERSITY
DEKALB
ILLINOIS -- 60115

ABSTRACT: Appropriate model selection is a fundamental problem in the field of statistics. Models with large number of possible explanatory variables require special attention due to infeasibility of huge model space. There are several suggestions available in the literature. Under the Bayesian approach, the classical way is to select the model with highest posterior probability. Using this fact the problem may be thought as a maximization problem over the model space where the objective function is the posterior probability of model and the maximization is taken place with respect to the models. We propose an efficient method for implementing this maximization and we illustrate its feasibility in high dimensional problem. By means of various simulation studies, this new approach has been shown to be efficient and to outperform other Bayesian methods namely median probability model and sampling method with frequency based estimators. Theoretical justification is provided.

KEY WORDS: Bayes Factor, Median Probability Model, Simulated Annealing, g-prior, Logistic Regression, Mixture of Weibull Regression, High Dimension, LPML