Paper Title and Its Main Contribution

First Author and Second Author

Abstract

Your abstract goes here. Variable selection is an important topic in linear regression analysis. In practice, a large number of predictors are introduced at the initial stage of modeling to enhance predictability and to avoid possible modeling biases. We propose a unified approach via penalized least squares. The penalty functions have to be singular at the origin to produce sparse solutions. Furthermore, the penalty function should be bounded by a constant to reduce bias.

KEY WORDS: Hard thresholding; LASSO; Nonparametric curve; Penalized likelihood; Oracle estimator.

Table 1: Median Relative Model Error

Method	Model 1	Model 2	Model 3
SCAD	0.35	0.42	0.38
LASSO	0.60	0.58	0.59
Hard	0.82	0.71	0.73

well as the oracle. This is very analogous to the superefficiency phenomenon in the Hodges estimator.

1 INTRODUCTION

Variable selection is an important topic in linear regression analysis. In practice, a large number of predictors usually are introduced at the initial stage of modeling to enhance possible modeling biases. On the other hand, to enhance predictability and to avoid overfitting, statisticians usually use stepwise deletion and subset selection. The penalty functions have to be singular at the origin to produce sparse solutions. Furthermore, with proper choice of regularization parameters, we show that the proposed estimators perform as well as the oracle procedure in variable selection; namely, they work as well as if the correct submodel were known. Our simulation shows that the newly proposed penalized likelihood compare favorably with other variable selection techniques.

2 METHODOLOGY

The penalized least squares idea can be extended naturally to likelihood-based models in various statistical contexts. Our approaches are distinguished from traditional methods (usually quadratic penalty) in that the penalty functions are symmetric, convex on $(0, \infty)$ (rather than concave for the negative quadratic penalty in the penalized likelihood situation), and possess singularities at the origin.

$$p_\lambda(|\beta_j|) = \lambda^2 - (|\beta_j| - \lambda)^2 I(|\beta_j| < \lambda) \tag{1}$$

Algorithm 1 SCAD Algorithm

1: Initialize $\beta^{(0)} =$ least squares estimate
2: **for** $k = 0, 1, 2, \dots$ until convergence **do**
3:     Set weights $w_j = p'_\lambda(|\beta_j^{(k)}|)/|\beta_j^{(k)}|$
4:     Minimize $\|Y - X\beta\|^2 + \sum_{j=1}^p w_j \beta_j^2$
5:     Set $\beta^{(k+1)} =$ the minimizer
6: **end for**

---

3 NUMERICAL STUDIES

We conduct some simulations to compare the performance of different variable selection procedures. We demonstrate how the rates of convergence for the penalized likelihood estimators depend on the regularization parameter.

4 CONCLUSION

The penalized likelihood estimators perform as well as the oracle procedure in terms of selecting the correct submodel when the regularization parameter is appropriately chosen. In short, the proposed procedures outperform the maximum likelihood estimator and perform as