Model Selection in Multiple Linear Regression

D.N. Kashid[1*]

[1] *Department of Statistics, Shivaji University, Kolhapur, India(MS), 416004; dnk_stats@unishivaji.ac.in*
[*] *Presenting author*

Keywords. *Outlier; Multicollinearity; M-estimator;Robust criterion.*

Model selection is the most persuasive problem in multiple linear regression. The performance of least squares parameter estimation based model selection methods are reasonably not well in the presence of outlier, multicollinearity, leverage, etc. due to non-robustness of least squares estimator. To overcome problems in the data, various alternative parameter estimation procedures like M-estimator, Rank-estimator, Ridge, Liu, Liu-M, JRM, LRM estimators are proposed by the researchers. Using some alternative parameter estimation methods, researchers have proposed robust model selection methods (**??**). However, only a few model selection methods cope with simultaneous occurrence of two or more problem in the data (**?**). In this article, we study the existing methods and propose a new model selection method resistant to simultaneously occurrence of outlier and multicollinearity. A superiority of a proposed method is illustrated through simulation study.

References

Ertas, H., Kaciranlar, S. & Guler, H. (2015). Robust Liu-type Estimator for Regression Based on M-estimator. *Communications in Statistics-Simulation and Computation*, DOI: 10.1080/03610918.2015.1045077.

Jadhav, N. H., Kashid, D. N. & Kulkarni, S. (2014). Subset selection in multiple linear regression in the presence of outlier and multicollinearity. *Statistical Methodology*, **19**, 44–59.

Ronchetti, E. & Staudte, R.G. (1994). A Robust Version of Mallows' C_p. *Journal of the American Statistical Association*, **89(426)**, 550–559.

Sommer, S. & Huggins, R. M. (1996). Variables Selection Using the Wald Test and a Robust C_p. *Applied Statistics*, **45(1)**, 15–29.