

MEAN SQUARE ERROR COMPARISON BETWEEN RIDGE AND LASSO ESTIMATORS IN THE MISSPECIFIED LINEAR REGRESSION MODEL

M. Kayanan^{1, 2, *} and P. Wijekoon³

¹*Postgraduate Institute of Science, University of Peradeniya, Peradeniya, Sri Lanka,*

²*Department of Physical Science, Vavuniya Campus of the University of Jaffna, Vavuniya, Sri Lanka,*

³*Department of Statistics and Computer Science, Faculty of Science, University of Peradeniya, Peradeniya, Sri Lanka.*

* *mgayanan@vau.jfn.ac.lk*

It is well-known that the Ridge Estimator has been used as an alternative estimator for Ordinary Least Squared Estimator (OLSE) to handle multicollinearity problem in the linear regression model. However, it introduces heavy bias when the number of predictors is high, and it may shrink irrelevant regression coefficients, but they are still in the model. To tackle this problem, Least Absolute Shrinkage and Selection Operator (LASSO) has been used to make the variable selection and shrinking the regression coefficients simultaneously. Further, the model misspecification due to excluding relevant explanatory variable in the linear regression model is considered as a severe problem in statistical research, and it will lead to bias and inconsistent parameter estimation. The performance of Ridge and LASSO estimators under a correctly specified regression model was well studied in the literature. This study intends to compare the performance of Ridge and LASSO estimators in root mean square error (RMSE) criterion under the misspecified regression model using a Monte-Carlo simulation study. In addition to that, a real-world example was employed to support the results. The analysis revealed that LASSO outperformed Ridge estimator in both cases such as correctly specified model and misspecified model. However, LASSO failed to outperform Ridge estimator in some situation under high multicollinearity.

Keywords: Ridge Estimator, LASSO, Misspecified model, Multicollinearity, Variable selection, Mean square error