International Journal of Scientific Research in Mathematical and Statistical Sciences **Research Paper**
Vol.6, Issue.4, pp.70-76, August (2019) **E-ISSN:** 2348-4519
DOI: https://doi.org/10.26438/ijsrmss/v6i4.7076

The Performance of a Ridge Estimator Based on Harmonic Mean

Satish Bhat[1], R. Vidya[2*]

[1,2]Dept. of Statistics, Yuvaraja's College, University of Mysore, Mysuru, Karnataka, India – 570005

Corresponding Author: drvidyaraju02@gmail.com, Mob. No.: +91-9448977945

Available online at: www.isroset.org

Received: 10/Jul/2019, Accepted: 12/Aug/2019, Online: 31/Aug/2019

Abstract - Linear dependence between predictors is one of the serious issues in regression analysis. Due to near linear dependence (or multicollinearity) between any two or more predictors, ordinary least squares (OLS) method will yield unstable estimates to the regression coefficients. In the literature, several techniques like Ridge regression, Principal component regression, Partial least squares regression, Liu method of regression etc., have been developed to overcome problem of multicollinearity. Among them Ridge regression is one of the most widely used methods, which will yield more stable estimate's as compared to OLS estimator. Here we propose a new ridge estimator based on Harmonic mean method. Performance of the ridge estimators is evaluated both theoretically and empirically under a wide range of degree of multicollinearity and error variances. Both methods have indicated that the performance of the suggested estimator is slightly more stable than some existing estimators, which are considered under study with respect to various degrees of multicollinearity, sample size, and error variance.

Key words: *Multiple linear regression (MLR), Multicollinearity, Ridge regression, and MSE.*

I. INTRODUCTION

Multicollinearity is one of the severe issues in regression analysis. To overcome the problem of multicollinearity several techniques such as, ridge regression (RR), principal component analysis (PCR), partial least squares regression (PLSR) etc., have been defined in the literature. Ridge regression is the one which is most widely used techniques among the above methods, especially in the areas of sentiment analysis [1], satellite imagery [2], genetics, forestry etc. Ridge regression is an alternative method to ordinary least squares (OLS) regression. Before we study the ridge regression, consider the standard form of multivariate linear regression (MLR) model, defined as

$$\boldsymbol{y} = X\boldsymbol{\beta} + \boldsymbol{u} \tag{1}$$

Where X is a ($n \times p$) data matrix, $\boldsymbol{y}$ is a ($n \times 1$) vector of response, $\boldsymbol{\beta}$ is ($p \times 1$) vector of regression coefficients and $\boldsymbol{u}$ is a ($n \times 1$) vector of random errors which are iid with zero mean and variance σ^2. If X has full rank, the ordinary least squares (OLS) method, will yield the estimate for $\boldsymbol{\beta}$ as

$$\hat{\boldsymbol{\beta}}_{OLS} = (X'X)^{-1} X'\boldsymbol{y}. \tag{2}$$

OLS estimator yields unstable estimates to the regression coefficients because sometimes inverse of $X'X$ may not exist due to the near linear dependence between predictors, and thus to overcome the problem of singularity, X is standardized such that $X'X$ is in the form of a near correlation matrix.

For simplicity in computation, we express the model defined in equation (1) in canonical form. Let W be a matrix of order ($p \times p$), such that its columns are normalized eigen vectors of $X'X$. Suppose $Z = XW$, then $Z'Z = W'X'XW$, where $Z'Z = D = diag(\lambda_1, \lambda_1, \ldots, \lambda_p)$, and $\lambda_j{}'s$ are the j^{th} eigen value of $X'X$, then the equation (1) can be written as

$$\boldsymbol{y} = Z\boldsymbol{\gamma} + \boldsymbol{u}, \tag{3}$$

where $\boldsymbol{\gamma} = W'\boldsymbol{\beta}$. The OLS estimator for $\boldsymbol{\gamma}$ is now given by

$$\hat{\boldsymbol{\gamma}}_{OLS} = (Z'Z)^{-1} Z'\boldsymbol{y} = D^{-1} Z'\boldsymbol{y}. \tag{4}$$

Since $\boldsymbol{\gamma} = W'\boldsymbol{\beta}$, implies $\hat{\boldsymbol{\beta}} = W\hat{\boldsymbol{\gamma}}$.

The outline of the present article is as follows. In section 2, the concept of ridge regression is discussed. Section 3, deals some ridge estimators. Under some stated conditions, performance of the proposed estimator is verified theoretically in section 4. In section 5, simulation study is conducted and results obtained are tabulated to see the performance of the suggested and few existing estimators in terms of their MSE. Some remarks are made on simulation results in section 6.

II. RIDGE ESTIMATION

There are number of techniques have been proposed in the literature to overcome the shortcomings of OLS, viz., Ridge regression, Principal component regression(PCR), Partial least squares regression(PLSR), etc. Among the above, ridge regression [3], [4] is the one, which is widely used technique in the regression analysis when there is a linear dependence between any two predictors. To overcome the problem of singularity, a constant $k(>0)$ is to be added to every j^{th} element of the diagonal of the matrix $X'X$, so that ridge estimator becomes more robust to the problem of singularity and it was shown that MSE of the ridge estimator is minimum compared to OLS, [5].
The ordinary ridge estimator for the regression coefficients is given by

$$\hat{\boldsymbol{\gamma}}_R = (Z'Z + kI)^{-1} Z'\boldsymbol{y} = A^{-1}Z'\boldsymbol{y} \qquad (5)$$

Where $A = D + kI$, and $X = ZW'$. Using equation (4), we write equation (5) as

$$\hat{\boldsymbol{\gamma}}_R = (I - A^{-1}k)\hat{\boldsymbol{\gamma}}_{OLS}\,. \qquad (6)$$

It was proved that the ridge estimator is biased and its bias-square is continuous and monotonically increasing function of k, and for $0 \leq k \leq \frac{\sigma^2}{\hat{\boldsymbol{\gamma}}^2_{max}}$, the MSE ($\hat{\boldsymbol{\gamma}}_R$) is minimum, where $\hat{\boldsymbol{\gamma}}^2_{max}$ is the largest element of $\hat{\boldsymbol{\gamma}}^2_R$ and σ^2 is replaced by its estimate $\hat{\sigma}^2 = \frac{\boldsymbol{y}'\boldsymbol{y} - \hat{\boldsymbol{\gamma}}'_{OLS} Z'\boldsymbol{y}}{n-p-1}$, [4]. Another estimate of σ^2 was suggested and it is defined as $\hat{\sigma}^2 = \frac{\boldsymbol{y}'\boldsymbol{y} - \hat{\boldsymbol{\gamma}}'_{OLS} Z'\boldsymbol{y}}{n-p}$, [6]. Above two estimators of σ^2, may yield negative estimates to the residual mean square, if $n < p$ and to overcome this, an estimate of σ^2 was suggested in [7] which makes use of 'hat matrix', H such that degrees of freedom for the error are $n - tr(2H - HH') = n - p$, as it avoids negative estimates to σ^2, when $n < p$.

III. SOME RIDGE ESTIMATORS

Several authors have been suggested various methods of estimating the ridge parameter k, say. Some of the well known methods in the literature are due to [8, 9, 10, 11, 12, 15, 16, 19, 20, 21, 22, 23, 24] etc.
Following are some of the well-known methods for estimating the ridge parameter k. Firstly, the estimator defined as

$$k_{HKB} = \frac{p\hat{\sigma}^2}{\hat{\boldsymbol{\gamma}}'\hat{\boldsymbol{\gamma}}} = k_1\,, \qquad (7)$$

This is due to [5]. It is observed that estimator due to [5] seems to be over shrunken the estimator towards zero, and it does not perform well if predictors are more than the sample size in the model.
The ordinary ridge estimator due to [5] was then modified [8], by making use of Eigen values and is given by

$$k_{LW} = \frac{p\hat{\sigma}^2}{\sum_{j=1}^{p} \lambda_j \hat{\boldsymbol{\gamma}}^2_j} = k_2\,, \qquad (8)$$

It relies on least-squares estimates of the parameters, and like the estimator due to [5], it shows poor performance when the error variance is small and the degree of correlation is very high. Another estimator which is defined [9] as

$$k_N = \frac{p\hat{\sigma}^2}{\sum_{j=1}^{p}\left\{\hat{\boldsymbol{\gamma}}^2_j / [1 + (1 + \lambda_j [\hat{\boldsymbol{\gamma}}^2_j / \hat{\sigma}^2]^{1/2})]\right\}} = k_3 \qquad (9)$$

This estimator which shrinks less as compared to the estimator defined in [5], and it is observed that it results in increased total variance of regression coefficients when number of observations is more than the predictors and thereby, far from true parameter value and shows little better performance when predictors is more than the sample size. Estimator due to [10], as

$$k_{KS} = \frac{\lambda_{max}\hat{\sigma}^2}{(n-p-1)\hat{\sigma}^2 + \lambda_{max}\hat{\boldsymbol{\gamma}}^2_{max}} = k_4\,. \qquad (10)$$

Where λ_{max} is the largest eigen value of $X'X$. It shrinks less as compared to estimator due to [5] but, like the estimator due to [9], it shows greater variability in total variance for large error variances. It is observed that the estimator

$$k_{DK} = Max\left(0, \frac{p\hat{\sigma}^2}{\hat{\boldsymbol{\gamma}}'\hat{\boldsymbol{\gamma}}} - \frac{1}{n(VIF_j)_{max}}\right) = k_5\,, \qquad (11)$$

was proposed by [11], which shrinks less and tend towards true parameter value, where $VIF_j = \frac{1}{(1-R_j^2)}, j = 1,2,\ldots,p;$ is the variance inflation factor of the j^{th} regressor. It is more stable than the estimators defined in equations (7) to (10) in both the cases i.e., when predictors are larger than the sample size and also when the sample size is more than the predictors. However, when sample size is more than the predictors it is observed that it does not dominate but coincides it with the estimator due to [5]. Estimator due to [12] is defined by

$$k_D(HM) = \frac{2p}{\lambda_{max}} \sum_{j=1}^{p} \frac{\hat{\sigma}^2}{\hat{\gamma}_j^2} = k_6, \quad (12)$$

It performs better when number of predictors is more than the sample size, and when predictors are highly collinear, but it becomes unstable when n is large.
Estimators due to [13, 14] are defined by

$$k_{SV} = \frac{p\hat{\sigma}^2}{\hat{\gamma}'\hat{\gamma}} + \frac{1}{n\lambda_{max}} = k_{HKB} + \frac{1}{n\lambda_{max}} = k_7, \quad (13)$$

$$k_{SV_1} = \frac{p\hat{\sigma}^2}{\hat{\gamma}'\hat{\gamma}} + \frac{1}{\lambda_{max}\hat{\gamma}'\hat{\gamma}} = k_1 + \frac{1}{\lambda_{max}\hat{\gamma}'\hat{\gamma}} = k_8 \quad (14)$$

$$k_{SV_2} = \frac{p\hat{\sigma}^2}{\hat{\gamma}'\hat{\gamma}} + \frac{1}{2m^2} = k_1 + \frac{1}{2m^2} = k_9, \quad (15)$$

where $m = \sqrt{\lambda_{max} / \lambda_{min}}$ is called the condition number [15]. Higher the value of m, higher is the degree of multicollinearity. If (30 < m < 100) means a moderate to strong multicollinearity, and if m > 100 suggests severe multicollinearity [16].

Few modified ridge estimators were suggested in [13, 14], and are obtained by modifying the estimator due to [5], and therefore they also seem to be over shrunken the estimator due to [5], i.e., more often the estimator tend towards zero and therefore they seem to be unstable like [5] when number of predictors exceeds the number of sample observations. Also, when error follows normal distribution, the estimators SV_1 and SV_2 [14], deviate a little away from the true parameter value as compared to estimators due to [5, 11].
Ordinary ridge estimator (SV_3) was suggested [17], which is obtained by taking the geometric mean of the estimators due to [5, 11] and is defined by

$$k_{SV_3} = GM(k_1, k_5) = \sqrt{k_1 \times k_5} = k_{10}, \quad (16)$$

Above estimator seems to be more stable in both the cases i.e., when the sample size *n,* is either more or less than the number of predictors.

IV. PROPOSED ESTIMATOR AND ITS PERFORMANCE

Here we suggest an ordinary ridge estimator say, SV_4 which is obtained by taking the harmonic mean of the estimators due to [4, 11], and is defined by

$$k_{SV_4} = HM(k_1, k_5) = 2k_1 \times k_5 / (k_1 + k_5) = k_{11} \quad (17)$$

It is noted that the estimators defined in equations (7) to (14) are verified under very high degree ($\rho \geq 0.9$) of multicollinearity between the predictors [18], whereas the performance of the estimators due to [13, 14, 17] are investigated under various degree of multicollinearity viz., low, moderate and high degree of multicollinearity.
The following are the results which prove that under certain general conditions, the proposed estimator is superior to the other estimators considered in this study.

Theorem 1: When $n < p$ for the linear regression model with homoscedastic, k_{11} is superior to k_1 in the MSE sense. That is,
$\Delta = MSE(\hat{\gamma}_{HKB}) - MSE(\hat{\gamma}_{SV_4}) \geq 0$, if

$$\sigma^2 \leq \frac{\sum_{j=1}^{p} \gamma_j^2 [(\hat{k}_1 + \hat{k}_{11})\lambda_j + 2\hat{k}_1\hat{k}_{11}] c_j^*}{\sum_{j=1}^{p} [\hat{k}_1 + \hat{k}_{11} + 2\lambda_j] c_j^*}, \text{ where}$$

$$c_j^* = \frac{\lambda_j}{(\lambda_j + \hat{k}_1)^2 (\lambda_j + \hat{k}_{11})^2}.$$

Proof: Since k_{11} is the harmonic mean of k_1 land k_5, and $k_1 > k_5$ implies,

$$k_{11} < k_1. \quad (18)$$

Accordingly, it is trivial that
$MSE(\hat{\gamma}_{SV_4}) \leq MSE(\hat{\gamma}_{HKB})$.

Or alternatively, since $k_{11} < k_1$, consider,

$$MSE(\hat{\gamma}_{SV_4}) = \sum_{j=1}^{p} \frac{\sigma^2\lambda_j + \hat{k}_{11}^2\gamma_j^2}{(\lambda_j + \hat{k}_{11})^2},$$

And $$MSE(\hat{\gamma}_{HKB}) = \sum_{j=1}^{p} \frac{\sigma^2\lambda_j + \hat{k}_1^2\gamma_j^2}{(\lambda_j + \hat{k}_1)^2} \quad (19)$$

Then,

$$\Delta = MSE(\hat{\gamma}_{HKB}) - MSE(\hat{\gamma}_{SV_4})$$

$$= \sum_{j=1}^{p} \left\{ \frac{\sigma^2\lambda_j + \hat{k}_1^2\gamma_j^2}{(\lambda_j + \hat{k}_1)^2} - \frac{\sigma^2\lambda_j + \hat{k}_{11}^2\gamma_j^2}{(\lambda_j + \hat{k}_{11})^2} \right\}$$

On simplification,

$$\Delta = \sum_{j=1}^{p} \left\{ \frac{\sigma^2\lambda_j(\hat{k}_1 + \hat{k}_{11} + 2\lambda_j) - \lambda_j\gamma_j^2[(\hat{k}_1 + \hat{k}_{11})\lambda_j + 2\hat{k}_1\hat{k}_{11}}{(\lambda_j + \hat{k}_1)^2(\lambda_j + \hat{k}_{11})^2} \right\}(\hat{k}_{11} - \hat{k}_1) \quad (20)$$

Since $(\hat{k}_{11} - \hat{k}_1) \le 0$, then $\Delta \ge 0$, if

$$\sigma^2 \le \frac{\sum_{j=1}^{p}\gamma_j^2[(\hat{k}_1 + \hat{k}_{11})\lambda_j + 2\hat{k}_1\hat{k}_{11}]c_j^*}{\sum_{j=1}^{p}[\hat{k}_1 + \hat{k}_{11} + 2\lambda_j]c_j^*}, \text{where}$$

$$c_j^* = \frac{\lambda_j}{(\lambda_j + \hat{k}_1)^2(\lambda_j + \hat{k}_{11})^2}$$

Hence the theorem is proved.

Theorem 2: For, $0 < k_j \le 1$ the linear regression model with homoscedastic k_{11} is superior to k_2 in the MSE sense. That is,

$$\Delta = MSE(\hat{\gamma}_{LW}) - MSE(\hat{\gamma}_{SV_4}) \ge 0, \text{ if}$$

$$\sigma^2 \le \frac{\sum_{j=1}^{p}\gamma_j^2[(\hat{k}_2 + \hat{k}_{11})\lambda_j + 2\hat{k}_2\hat{k}_{11}]c_j^*}{\sum_{j=1}^{p}[\hat{k}_2 + \hat{k}_{11} + 2\lambda_j]c_j^*}, \text{where}$$

$$c_j^* = \frac{\lambda_j}{(\lambda_j + \hat{k}_2)^2(\lambda_j + \hat{k}_{11})^2}.$$

Proof:

Case 1: First to prove $k_{11} < k_2$.

Consider, $$k_{LW} = \frac{p\hat{\sigma}^2}{\sum_{j=1}^{p}\lambda_j\hat{\gamma}_j^2} = k_2$$

and $k_{SV_4} = 2k_{HKB} \times k_{DK} / (k_{HKB} + k_{DK}) = k_{11}$.

Since $k_2 < k_1$, and suppose $k_5 \le k_2$, then

$$k_2 + k_5 \le k_1 + k_5 \Leftrightarrow k_2 \le k_1 + k_5$$
$$\Rightarrow 1/k_2 \ge 1/(k_1 + k_5)$$
$$\Rightarrow 2k_1k_5/k_2 \ge 2k_1k_5/(k_1 + k_5) = k_{11}$$
$$\Leftrightarrow k_{11} \le k_2 \quad (21)$$

Consider,

$$MSE(\hat{\gamma}_{SV_4}) = \sum_{j=1}^{p} \frac{\sigma^2\lambda_j + \hat{k}_{11}^2\gamma_j^2}{(\lambda_j + \hat{k}_{11})^2} \quad \text{And}$$

$$MSE(\hat{\gamma}_{SV_4}) = \sum_{j=1}^{p} \frac{\sigma^2\lambda_j + \hat{k}_2^2\gamma_j^2}{(\lambda_j + \hat{k}_2)^2} \quad (22)$$

Then,

$$\Delta = MSE(\hat{\gamma}_{LW}) - MSE(\hat{\gamma}_{SV_4})$$

$$= \sum_{j=1}^{p} \left\{ \frac{\sigma^2\lambda_j + \hat{k}_2^2\gamma_j^2}{(\lambda_j + \hat{k}_2)^2} - \frac{\sigma^2\lambda_j + \hat{k}_{11}^2\gamma_j^2}{(\lambda_j + \hat{k}_{11})^2} \right\}$$

On simplification,

$$\Delta = \sum_{j=1}^{p} \left\{ \frac{\sigma^2\lambda_j(\hat{k}_2 + \hat{k}_{11} + 2\lambda_j) - \lambda_j\gamma_j^2[(\hat{k}_2 + \hat{k}_{11})\lambda_j + 2\hat{k}_2\hat{k}_{11}}{(\lambda_j + \hat{k}_2)^2(\lambda_j + \hat{k}_{11})^2} \right\}(\hat{k}_{11} - \hat{k}_2) \quad (23)$$

Since $(\hat{k}_{11} - \hat{k}_2) \le 0$, then $\Delta \ge 0$, if

$$\sigma^2 \le \frac{\sum_{j=1}^{p}\gamma_j^2[(\hat{k}_2 + \hat{k}_{11})\lambda_j + 2\hat{k}_2\hat{k}_{11}]c_j^*}{\sum_{j=1}^{p}[\hat{k}_2 + \hat{k}_{11} + 2\lambda_j]c_j^*}, \text{where}$$

$$c_j^* = \frac{\lambda_j}{(\lambda_j + \hat{k}_2)^2(\lambda_j + \hat{k}_{11})^2}$$

Case 2: Suppose $k_5 \ge k_2$ then as in above, it can be shown that

$$k_{11} \ge k_2 \quad (24)$$

Thus $\Delta \ge 0$, if

$$\sigma^2 \le \frac{\sum_{j=1}^{p}\gamma_j^2[(\hat{k}_2 + \hat{k}_{11})\lambda_j + 2\hat{k}_2\hat{k}_{11}]c_j^*}{\sum_{j=1}^{p}[\hat{k}_2 + \hat{k}_{11} + 2\lambda_j]c_j^*}, \text{where}$$

$$c_j^* = \frac{\lambda_j}{(\lambda_j + \hat{k}_2)^2(\lambda_j + \hat{k}_{11})^2}.$$

Hence the theorem is proved.

It is observed that in a similar approach one could compare MSE ($SV_4 = k_{11}$) with that of MSE of the remaining other estimators, which are considered under study.

V. SIMULATION STUDY

The simulation study was conducted for various values of n, the sample size; p the number of predictors, residual variance σ^2, and ρ, the degree of correlation, in the presence of low, moderate and a high degree of multicollinearity. The results were obtained by generating a random data matrix X of size ($n \times p$) using the relation:

$$x_{ij} = (1-\rho^2)^{1/2}\xi_{ij} + \rho\xi_{ip},\ i = 1,2,\ldots,n;\ j = 1,2,\ldots,p;$$

where ξ_{ij} 's are independent standard normal pseudo-random numbers, ρ is fixed such that ρ^2 is the degree of correlation between any two predictors. These variables are standardized such that $X'X$ takes up correlation matrix form, and to generate $\boldsymbol{y}$ we have assumed vector $\boldsymbol{\beta}$ as

$$\boldsymbol{\beta} = (0.03,0.5,0.03,0.91,0.59,0.74,0.3,0.95,0.83,0.9,0.5,0.4,0.3,0.5,0.3,0.9)'$$

The performance of the suggested estimators was evaluated for various sample size n = 10, 25, 100 and 1000; number of predictors $p = 15$, and the variance of the residual term σ^2: 5, 25, 100, and 1000; and the degree of correlation ρ = 0.3, 0.5, 0.7, 0.9, 0.99 and 0.9999. The experiment was replicated 1000 times each and the average of mean square error (AMSE) was computed using the relation,

$$AMSE(\hat{\boldsymbol{\beta}}^*) = \frac{1}{1000}\sum_{j=1}^{1000}(\hat{\boldsymbol{\beta}}^*_{(j)} - \boldsymbol{\beta})'(\hat{\boldsymbol{\beta}}^*_{(j)} - \boldsymbol{\beta}),$$

where, $\hat{\boldsymbol{\beta}}^*$ is any estimator that was used in this study. Ridge estimates were computed by considering the different estimators of the ridge parameter k, defined in equations (7) to (15). The results of the simulation are presented in Table 1. Here, the estimators leading to the maximum ratio of AMSE of OLS over AMSE of other ridge estimators were considered to be the best in terms of MSE.

VI. CONCLUSION

The simulation study indicates that the suggested estimator yields more stable estimates as compared to all the other estimators which are considered under study in terms of ratio of AMSE over OLS. It was noticed that (Table 1), when the sample size n is small (i.e., $n < p$), and for small error variance (σ^2=5), the estimators due to [9, 11, 12] have yielded more unstable estimates for the ridge parameter. Estimators due to [5, 13, 14, 17] gives more stable estimates to the regression coefficients, but these estimates over shrinks when ($n < p$), and thereby deviated slightly from the true parameter value. In this context, the suggested estimator has yielded more stable estimate's as compared to all the other estimators which are considered under study for a wide range of sample size(n), degree of correlation(ρ), and error variance(σ^2). Since the performance of the suggested estimator was verified empirically under various values of n, ρ and σ^2, and also theoretically, we conclude that the performance of the suggested estimator is better, satisfactory and comparable to all the other estimators which are considered under this study. Further in real life situations there are possibility for further research in the area of studying inaccuracy of estimates, testing for the significance of the estimators, and presence of outliers.

Table 1: AMSE ratio of OLSE over different Ridge estimator's when error (u) ~ N (0, $\sigma^2 I$), for n=10.

σ^2	AMSE	ρ = 0.3	ρ = 0.5	ρ = 0.7	ρ = 0.9	ρ = 0.99	ρ = 0.9999
5	k_1	3.5052	3.9484	4.1245	4.7121	4.7880	4.4881
	k_2	3.4992	3.8991	4.0960	4.6656	3.7379	1.0154
	k_3	2.3015	2.8246	3.3689	6.4354	14.5040	18.1223
	k_4	1.9289	2.3246	2.1865	2.2704	2.3591	2.2819
	k_5	7.7626	9.0196	9.2655	10.4305	9.0393	7.3488
	k_6	8.6582	10.6234	10.7924	15.8184	18.6218	31.3335
	k_7	3.5119	3.9517	4.1266	4.7167	4.8415	7.2092
	k_8	3.5054	3.9485	4.1246	4.7122	4.7881	4.4881
	k_9	3.5146	3.9547	4.1285	4.7154	4.7915	4.4917
	k_{10}	3.4295	3.8941	4.0948	4.6636	3.8207	1.0178
	k_{11}	3.4285	3.8911	4.0940	4.6624	3.7395	1.0166
25	k_1	4.5073	4.7924	4.4934	4.6348	4.9861	4.7326
	k_2	4.2380	4.6868	4.3512	4.4955	3.3945	1.0184
	k_3	2.8982	3.1859	3.7799	6.8787	15.3765	28.9510
	k_4	2.2514	2.3540	2.2857	2.2965	2.2678	2.3069
	k_5	13.3060	13.2277	12.0312	10.8080	9.3626	7.9695
	k_6	24.2855	19.4778	15.7593	18.3605	13.9564	22.9745
	k_7	4.5145	4.7958	4.4982	4.6416	5.0413	7.5719
	k_8	4.5073	4.7924	4.4934	4.6348	4.9861	4.7326
	k_9	4.5208	4.8006	4.5000	4.6394	4.9895	4.7359
	k_{10}	4.3514	4.7384	4.4171	4.5634	3.6023	1.0218
	k_{11}	4.3233	4.7357	4.4086	4.5602	3.5569	1.0201
100	k_1	4.8054	4.9950	4.7052	4.8531	5.1445	4.7719
	k_2	4.6798	4.9065	4.5967	4.6430	3.8365	1.0163
	k_3	3.0156	3.2552	3.9979	7.3391	17.0845	39.3947
	k_4	2.2558	2.4095	2.2541	2.2848	2.4057	2.4421
	k_5	13.9265	14.4947	12.7387	11.1766	9.8476	7.9595
	k_6	27.7249	21.8505	16.0573	21.0466	25.3750	30.7331
	k_7	4.8096	4.9979	4.7092	4.8618	5.2192	7.5731
	k_8	4.8054	4.9950	4.7052	4.8531	5.1445	4.7719
	k_9	4.8192	5.0031	4.7112	4.8573	5.1477	4.7754
	k_{10}	4.7411	4.9505	4.6500	4.7431	4.0339	1.0190
	k_{11}	4.7379	4.9497	4.6480	4.7311	4.0102	1.0177
1000	k_1	4.6484	4.6762	4.8552	4.6926	4.8502	4.9056
	k_2	4.5085	4.3757	4.7434	4.4808	3.4806	1.0142
	k_3	2.8849	3.3872	3.8610	7.0808	14.4073	38.0209
	k_4	2.4028	2.1886	2.3576	2.2841	2.2890	2.3098
	k_5	13.8205	13.5604	12.7824	10.8553	9.1894	8.3496
	k_6	26.8844	22.4698	16.0421	19.5978	25.2019	52.9182
	k_7	4.6532	4.6841	4.8596	4.7012	4.9151	7.9087
	k_8	4.6484	4.6762	4.8552	4.6926	4.8502	4.9056
	k_9	4.6617	4.6856	4.8612	4.6971	4.8539	4.9092
	k_{10}	4.5770	4.5102	4.7951	4.5793	3.6948	1.0167
	k_{11}	4.5729	4.4815	4.7918	4.5670	3.6483	1.0155

Table 2: AMSE ratio of OLSE over different Ridge estimator's when error (u) ~ N (0, $\sigma^2 I$), for n=25.

σ^2	AMSE	ρ = 0.3	ρ = 0.5	ρ = 0.7	ρ = 0.9	ρ = 0.99	ρ = 0.9999

σ^2=5	k_1	1.0966	1.2336	1.5706	2.5119	3.8460	4.0615
	k_2	1.0963	1.2334	1.5704	2.5113	3.8376	3.1381
	k_3	0.6172	0.6854	0.8168	1.1948	5.2660	15.0136
	k_4	0.6708	0.7702	0.9627	1.4223	1.8313	1.9037
	k_5	2.2634	2.6651	3.4762	6.3740	12.0762	8.5925
	k_6	1.0752	1.1985	1.4460	2.2302	12.0186	28.4818
	k_7	1.0966	1.2336	1.5706	2.5119	3.8461	4.0808
	k_8	1.0966	1.2336	1.5706	2.5119	3.8460	4.0615
	k_9	1.0976	1.2342	1.5711	2.5125	3.8471	4.0627
	k_{10}	1.0965	1.2335	1.5705	2.5116	3.8418	3.5400
	k_{11}	1.0965	1.2335	1.5705	2.5116	3.8418	3.4857
σ^2=25	k_1	3.3770	3.5243	3.6864	3.8376	4.1507	4.0701
	k_2	3.3756	3.5232	3.6856	3.8364	4.1408	3.1486
	k_3	1.0593	1.0903	1.1533	1.4294	5.8949	16.3765
	k_4	1.2111	1.3492	1.5839	1.7441	1.9226	1.9033
	k_5	28.1625	28.8995	28.4109	23.4998	14.3555	18.6029
	k_6	3.6256	2.9249	2.7010	3.2274	13.9508	14.4779
	k_7	3.3770	3.5243	3.6864	3.8376	4.1509	4.0891
	k_8	3.3770	3.5243	3.6864	3.8376	4.1507	4.0701
	k_9	3.3818	3.5272	3.6883	3.8389	4.1519	4.0712
	k_{10}	3.3763	3.5237	3.6860	3.8370	4.1458	3.5545
	k_{11}	3.3763	3.5237	3.6860	3.8370	4.1458	3.5008
σ^2= 100	k_1	3.7285	3.7172	4.0023	4.0213	4.0709	4.0479
	k_2	3.7269	3.7159	4.0014	4.0200	4.0617	3.1546
	k_3	1.0957	1.1147	1.1781	1.4547	5.6495	26.0630
	k_4	1.2480	1.3882	1.6132	1.8003	1.8784	1.8742
	k_5	51.3975	45.5635	41.3902	27.0046	14.1050	8.5569
	k_6	4.3790	3.1160	2.7706	3.2837	14.2973	19.6255
	k_7	3.7285	3.7172	4.0023	4.0213	4.0711	4.0666
	k_8	3.7285	3.7172	4.0023	4.0213	4.0709	4.0479
	k_9	3.7341	3.7205	4.0044	4.0226	4.0721	4.0490
	k_{10}	3.7277	3.7166	4.0018	4.0206	4.0663	3.5498
	k_{11}	3.7277	3.7166	4.0018	4.0206	4.0663	3.5021
σ^2= 1000	k_1	3.7276	3.8183	3.8662	3.9286	4.0728	4.2448
	k_2	3.7259	3.8170	3.8653	3.9273	4.0634	3.3299
	k_3	1.0984	1.1184	1.1738	1.4417	5.6510	94.8065
	k_4	1.2482	1.3969	1.5920	1.7785	1.9168	1.9233
	k_5	53.8947	48.7379	39.9739	26.6138	14.1935	9.1099
	k_6	4.4434	3.2807	2.7367	3.2672	14.0924	15.9370
	k_7	3.7276	3.8183	3.8662	3.9286	4.0729	4.2639
	k_8	3.7276	3.8183	3.8662	3.9286	4.0728	4.2448
	k_9	3.7332	3.8216	3.8683	3.9299	4.0739	4.2459
	k_{10}	3.7267	3.8176	3.8658	3.9279	4.0681	3.7388
	k_{11}	3.7267	3.8176	3.8658	3.9279	4.0681	3.6936

Table 3: AMSE ratio of OLSE over different Ridge estimator's when error (u) ~ N (0, $\sigma^2 I$), for n=100.

σ^2	AMSE	$\rho=$ 0.3	$\rho=$ 0.5	$\rho=$ 0.7	$\rho=$ 0.9	$\rho=$ 0.99	$\rho=$ 0.9999
σ^2=5	k_1	0.5371	0.6112	0.8020	1.5711	3.4197	3.8714
	k_2	0.5371	0.6112	0.8020	1.5710	3.4183	3.7001
	k_3	0.3819	0.4298	0.5268	0.8186	2.3471	12.7263
	k_4	0.3898	0.4449	0.5628	0.9798	1.7384	1.8272
	k_5	0.9844	1.1736	1.5587	3.3054	11.9150	8.9183
	k_6	0.5245	0.5674	0.6988	1.2482	5.1468	16.1818
	k_7	0.5371	0.6112	0.8020	1.5711	3.4197	3.8731
	k_8	0.5371	0.6112	0.8020	1.5711	3.4197	3.8714
	k_9	0.5374	0.6114	0.8021	1.5713	3.4203	3.8721
	k_{10}	0.5371	0.6112	0.8020	1.5710	3.4190	3.7841
	k_{11}	0.5371	0.6112	0.8020	1.5710	3.4190	3.7828
σ^2=25	k_1	2.9121	3.0700	3.3203	3.6810	3.9848	3.9387
	k_2	2.9118	3.0698	3.3201	3.6808	3.9831	3.7715
	k_3	0.9589	0.9817	1.0211	1.1384	2.5636	18.8428
	k_4	1.0030	1.0876	1.2503	1.5647	1.8747	1.8224
	k_5	9.9761	22.9204	27.6244	31.7710	19.5080	9.1642
	k_6	2.0828	1.7809	1.7654	2.0874	6.1912	17.1139
	k_7	2.9121	3.0700	3.3203	3.6811	3.9848	3.9404
	k_8	2.9121	3.0700	3.3203	3.6810	3.9848	3.9387
	k_9	2.9150	3.0716	3.3213	3.6818	3.9856	3.9395
	k_{10}	2.9119	3.0699	3.3202	3.6809	3.9839	3.8536
	k_{11}	2.9119	3.0699	3.3202	3.6809	3.9839	3.8524
σ^2= 100	k_1	3.5072	3.6338	3.6792	3.8057	3.9890	3.9478
	k_2	3.5068	3.6336	3.6790	3.8054	3.9872	3.7781
	k_3	1.0263	1.0367	1.0552	1.1486	2.5771	20.1797
	k_4	1.0754	1.1540	1.2967	1.5742	1.8178	1.8641
	k_5	9.4578	18.0158	17.0645	16.1660	15.0556	9.1875
	k_6	2.4985	1.9162	1.8441	2.0322	6.2564	18.2907
	k_7	3.5072	3.6338	3.6792	3.8057	3.9890	3.9495
	k_8	3.5072	3.6338	3.6792	3.8057	3.9890	3.9478
	k_9	3.5110	3.6360	3.6805	3.8065	3.9898	3.9485
	k_{10}	3.5070	3.6337	3.6791	3.8055	3.9881	3.8614
	k_{11}	3.5070	3.6337	3.6791	3.8055	3.9881	3.8602
σ^2= 1000	k_1	3.6794	3.6787	3.7625	3.8500	3.9495	3.9542
	k_2	3.6790	3.6785	3.7623	3.8498	3.9477	3.7821
	k_3	1.0338	1.0402	1.0591	1.1523	2.5472	20.3159
	k_4	1.0828	1.1598	1.3017	1.6043	1.8166	1.8695
	k_5	11.6633	17.4302	8.9516	9.4455	9.7090	9.1825
	k_6	2.4830	1.9109	1.7848	2.1669	6.0644	17.4241
	k_7	3.6794	3.6787	3.7625	3.8500	3.9495	3.9560
	k_8	3.6794	3.6787	3.7625	3.8500	3.9495	3.9542
	k_9	3.6834	3.6809	3.7638	3.8509	3.9502	3.9549
	k_{10}	3.6792	3.6786	3.7624	3.8499	3.9486	3.8666
	k_{11}	3.6792	3.6786	3.7624	3.8499	3.9486	3.8653

Table 4: AMSE ratio of OLSE over different Ridge estimator's when error (u) ~ N (0, $\sigma^2 I$), for n=1000.

σ^2	AMSE	$\rho=$ 0.3	$\rho=$ 0.5	$\rho=$ 0.7	$\rho=$ 0.9	$\rho=$ 0.99	$\rho=$ 0.9999
σ^2=5	k_1	0.0608	0.0711	0.0963	0.2266	1.3529	3.8387
	k_2	0.0608	0.0711	0.0963	0.2266	1.3529	3.8373
	k_3	0.0584	0.0686	0.0922	0.2024	0.7374	26.1321
	k_4	0.0585	0.0687	0.0924	0.2043	0.8593	1.7892
	k_5	0.0944	0.1152	0.1591	0.3409	2.4975	11.9403
	k_6	0.0607	0.0715	0.0961	0.2154	1.0178	53.7723
	k_7	0.0608	0.0711	0.0963	0.2266	1.3529	3.8387
	k_8	0.0608	0.0711	0.0963	0.2266	1.3529	3.8387
	k_9	0.0608	0.0711	0.0963	0.2266	1.3529	3.8388
	k_{10}	0.0608	0.0711	0.0963	0.2266	1.3529	3.8380
	k_{11}	0.0608	0.0711	0.0963	0.2266	1.3529	3.8380
σ^2=25	k_1	0.9818	1.0931	1.3614	2.2360	3.6231	4.0034
	k_2	0.9818	1.0931	1.3614	2.2360	3.6231	4.0020
	k_3	0.5635	0.6135	0.6970	0.8694	1.0874	13.1024
	k_4	0.5646	0.6175	0.7103	0.9443	1.4909	1.8210
	k_5	1.9272	2.3914	3.5856	8.2576	13.7715	12.9191
	k_6	0.6522	0.6925	0.7893	1.0235	1.7037	7.1502
	k_7	0.9818	1.0931	1.3614	2.2360	3.6231	4.0034
	k_8	0.9818	1.0931	1.3614	2.2360	3.6231	4.0034
	k_9	0.9819	1.0932	1.3614	2.2361	3.6232	4.0035
	k_{10}	0.9818	1.0931	1.3614	2.2360	3.6231	4.0027
	k_{11}	0.9818	1.0931	1.3614	2.2360	3.6231	4.0027
σ^2= 100	k_1	3.0546	3.1144	3.3425	3.6371	3.9738	3.9670
	k_2	3.0546	3.1144	3.3425	3.6371	3.9737	3.9656
	k_3	0.9564	0.9663	0.9830	1.0051	1.1108	9.7520
	k_4	0.9591	0.9751	1.0081	1.1054	1.5295	1.8197
	k_5	9.9624	16.8922	12.8025	19.8380	17.2263	12.7779
	k_6	1.2049	1.1201	1.1393	1.2160	1.8140	6.1615
	k_7	3.0546	3.1144	3.3425	3.6371	3.9738	3.9670
	k_8	3.0546	3.1144	3.3425	3.6371	3.9738	3.9670
	k_9	3.0551	3.1147	3.3426	3.6372	3.9739	3.9671
	k_{10}	3.0546	3.1144	3.3425	3.6371	3.9738	3.9663
	k_{11}	3.0546	3.1144	3.3425	3.6371	3.9738	3.9663
σ^2= 1000	k_1	3.5234	3.6244	3.5727	3.8064	3.8899	3.8633
	k_2	3.5234	3.6244	3.5727	3.8064	3.8899	3.8619
	k_3	1.0040	1.0053	1.0074	1.0146	1.1095	8.1315
	k_4	1.0069	1.0147	1.0338	1.1167	1.5239	1.8253
	k_5	9.1010	7.8982	6.8085	9.6437	7.0370	12.2899
	k_6	1.2339	1.1768	1.1730	1.2371	1.7592	9.2876
	k_7	3.5234	3.6244	3.5727	3.8064	3.8899	3.8633
	k_8	3.5234	3.6244	3.5727	3.8064	3.8899	3.8633
	k_9	3.5240	3.6247	3.5729	3.8066	3.8900	3.8634
	k_{10}	3.5234	3.6244	3.5727	3.8064	3.8899	3.8626
	k_{11}	3.5234	3.6244	3.5727	3.8064	3.8899	3.8626

ACKNOWLEDGMENT

I am very grateful to Dr. C. Raju, Professor in Quantitative Methods and Operations Management, Kozhikode, Kerala, India, for his continuous support and invaluable suggestions in writing this article.

REFERENCES

[1] R. Nair, V. Jain, A. Bhagat, R. Agarwal, "An Efficient approach for Sentiment Analysis Using Regression analysis Technique", International Journal of Computer Sciences and Engineering, Vol. 7, issue 3, pp. 161-165, 2019.

[2] T. Sur, A. Nath, "An innovative Method to Calculate the Economic Development Index of Important Cities in West Bengal From Satellite Imagery", International Journal of Computer Sciences and Engineering, Vol. 7, issue 4, pp. 411-417, 2019.

[3] A. E. Hoerl, "Application of Ridge Analysis to Regression Problems", Chemical Engineering Progress, Vol. 58, pp. 54–59, 1962.

[4] A.E. Hoerl, R. W. Kennard, "Ridge Regression: Biased Estimation for Non Orthogonal Problems", Technometrics, Vol. 12, pp. 69–82, 1970.

[5] A. E. Hoerl, R.W. Kennard, K. F. Baldwin, "Ridge Regression: Some Simulations", Communications in Statistics-Theory and Methods, Vol. 4, No., 2, pp.105–123, 1975.

[6] T. Hastie, R. Tibshirani, "Generalised Additive Models", Statistical Science, Vol. 1, No., 3, pp. 297–318, 1986.

[7] A. M. Halawa, M.Y. El Bassiouni, "Tests of Regression Coefficients under Ridge Regression Models", Journal of Statistical Computation and Simulation, Vol. 65, pp. 341–356, 2000.

[8] J. F. Lawless, P. Wang, "A Simulation Study of Ridge and Other Regression Estimators", Communications in Statistics-Theory and Methods, Vol. 5, No., 4, pp. 307–323, 1976.

[9] M. Nomura, "On the Almost Unbiased Ridge Regression Estimator", Communications in Statistics – Simulation and Computation, Vol. 17, No., 3, pp. 729–743, 1988.

[10]G. Khalaf, G. Shukur, "Choosing Ridge Parameter for Regression Problems", Communications in Statistics-Theory and Methods, Vol. 34, pp. 1177–1182, 2005.

[11]A.V. Dorugade, D. N. Khashid, "Alternative Method for Choosing Ridge Parameter for Regression", Applied Mathematical Sciences, Vol .4, No., 9, pp. 447–456, 2010.

[12]A. V. Dorugade, "New Ridge Parameters for Ridge Regression", Journal of the Association of Arab Universities for Basic and Applied Sciences. Vol. 15, pp. 94–99, 2014.

[13]S. Bhat, R. Vidya, "Improvement on Ridge Regression Estimator", In the Proceedings of the National Conference on Statistical Methods and Data Analysis, Yuvaraja's College, Mysore, India, ISBN No. 81-7877-093-8, pp. 65–71, 2014.

[14]S. Bhat, R. Vidya, "A Comparative Study on the Performance of New Ridge Estimators", Pakistan Journal of Statistics and Operations Research (PJSOR), Vol. 12, No. 2, pp. 317–325, 2016.

[15]H. D. Vinod, A. Ullah, "Recent Advances in Regression Methods", Marcel Dekker, New York, 1981.

[16]K. Liu, "Using Liu Type Estimator to Combat Collinearity", Communications in Statistics-Theory and Methods. Vol. 32, pp. 1009–1020, 2003.

[17]S. Bhat, R. Vidya, "A New Ridge Estimator and its Performance", In the Proceedings of the National Conference on Operations Management, Analytics and Statistical Methods(NOMAS-2017), Mysore, India, ISBN No. 978-9383302-20-8, 2017a.

[18]S. Bhat, "Contributions to the study of Partial Least Squares Regression(PLSR)", Ph. D. Thesis, Submitted to University of Mysore, Mysuru, India, chapter 3, pp. 71–105, 2017b.

[19]G. Khalaf, "A Proposed Ridge Parameter to Improve the Least Square Estimator", Journal of Modern Applied Statistical Method, Vol. 11, No., 2, pp. 443–449, 2012.

[20]B. M. Kibria, "Performance of Some Ridge Regression Estimators", Communications in Statistics-Simulation and Computation, Vol. 32, pp. 419–435, 2003.

[21]S. Mardikyan, E. Cetin, "Efficient Choice of Biasing Constant for Ridge Regression", International Journal of Contemporary Mathematical Sciences, Vol. 3, pp. 527–547, 2008.

[22]D. C. Montgomery, E. A. Peck, G. G. Vining, "Introduction to linear regression Analysis", 3rd Edition, John Wiley and Sons (Asia), Singapore, 2003.

[23]G. Muniz, B. M. Kibria, "On Some Ridge Regression Estimators-An Empirical Comparisons", Communications in Statistics –Simulation and Computation, Vol. 38, No., 3, pp. 621–630, 2009.

[24]M. Yazid Al-Hassan, "Performance of a New Ridge Regression Estimator", Journal of the Association of Arab Universities for Basic and Applied Sciences, Vol. 9, 23–26, 2010.

AUTHORS PROFILE

Satish Bhat, has obtained his B.Sc., M.Sc., M.Phil from Karnatk University, Dharwad and Ph. D. in Statistics from University of Mysore (UOM), Mysuru, in 1992, 1994, 1996 and 2017. He is currently working as Assistant Professor, in the Department of Statistics, Yuvaraja's College, a Constituent-Auotonomous College of UOM, since 2007. He is a life member of The Indian Science Congress Association, Kolkatta, since 2015. He has published 10 research papers in reputed International Journal including Communications in Statistics-Simulation and Computation and Conferences and it's also available online. Area of interest includes regression analysis, inferential statistics, and count regression. He has over 20 years of teaching experience including Pre University, B.Sc., MCA, MBA, MPED, M.Sc. (Mathematics, Molecular Biology, Food Science) and about 8 years of research experience.

Vidya, R., obtained her B.Sc., M.Sc., from Mangalore University, Mangalore, and Ph D. in Statistics from University of Mysore (UOM), Mysuru, in 1984, 1986, and 1999. She is currently working as Associate Professor, in the Department of Statistics, Yuvaraja's College, a Constituent-Autonomous College of UOM, since 1999. She is the member of The Indian Science Congress Association, Kolkatta. She has published more than 15 research articles in reputed National / International Journals including Communications in Statistics-Simulation and Computation and Conferences. Area of interest include Statistical quality control, Inferential Statistics, Regression analysis, count regression etc. She has over 25 years of teaching experience including B.Sc., M.Sc. (Statistics, Molecular Biology & Food Science) and about 15 years of research experience.