On Comparison of Some Ridge Parameters in Ridge Regression

Ashok V. Dorugade

Y C Mahavidyalaya Halkarni, Tal-Chandgad, Kolhapur, Maharashtra, India
Corresponding Author: adorugade@rediffmail.com

Received: October 10, 2013 / Revised: January 25, 2014 / Accepted: February 17, 2014

ABSTRACT

In this article, a new approach to obtain the ridge parameter introduces for the multiple linear regression model suffers from the problem of multicollinearity. Furthermore, we compare the proposed ridge parameter with the other well-known ridge-parameters through ridge estimators evaluated elsewhere in terms of mean squares error (MSE) criterion. Finally, a numerical example and simulation study has been conducted to illustrate the optimality of the proposed ridge parameter.

Keywords: Multicollinearity, Ridge parameter, Mean squared error matrix, Simulation**.**

1. Introduction

Consider the linear regression model

$$Y = X\beta + \varepsilon\ , \tag{1}$$

where Y is a n×1 random vector of response variables, ***X*** is a known n×p matrix with full column rank, ε is a n×1 vector of errors with E(ε) = 0 and Cov(ε) = $\sigma^2 I_n$. β is a p×1 vector of unknown regression parameters and σ^2 is the unknown variance parameter. For the sake of convenience, we assume that the matrix X and response variable Y are standardized in such a way that $X^{'}X$ is a non-singular correlation matrix and $X^{'}Y$ is the correlation between X and Y. The most common estimator for β is the ordinary least squares estimator

$$\hat{\beta}_{OLS} = (X^{'}X)^{-1}X^{'}Y$$

Let $\wedge$ and T be the matrices of eigen values and eigen vectors of $X^{'}X$, respectively, satisfying $T^{'}X^{'}XT = \wedge$ = diagonal ($\lambda_1, \lambda_2, ..., \lambda_p$), where λi being the i[th] eigen value of $X^{'}X$ and $T^{'}T = TT^{'}$ = I_p. We obtain the equivalent model

$$Y = Z\alpha + \varepsilon, \tag{2}$$

where Z = XT , which implies that $Z^{'}Z = \wedge$, and α = $T^{'}\beta$ (Montgomery et al., [25])

Generally, we use the ordinary least squares technique to estimate the regression parameter α because of its simplicity and easiness of computation.

Then OLS estimator of α is given by

$$\hat{\alpha}_{OLS} = (Z^{'}Z)^{-1}Z^{'}Y = \wedge^{-1}Z^{'}Y. \tag{3}$$

Therefore, OLS estimator of β is given by

$$\hat{\beta}_{OLS} = T\hat{\alpha}_{OLS}.$$

However, in some situations the regressors are nearly or perfectly linearly related, the problem of multicollinearity is said to exit and in such cases the usual inference based on such models will become erroneous. Multicollinearity also tends to produce OLS estimates that are unstable and large in absolute values. Ordinary Ridge regression (RR) introduced by Hoerl and Kennard [15] is intended to overcome the problem of multicollinearity by adding a positive constant (or ridge parameter) k, normally lies between 0 and 1, to the diagonal elements of the least square estimator. It is given as:

$$\hat{\alpha}_{RR} = \left[I - k(\wedge + kI)^{-1}\right]\hat{\alpha}_{OLS}. \tag{4}$$

Therefore, RR estimator of β is given by

$$\hat{\beta}_{RR} = T\hat{\alpha}_{RR},$$

and mean square error of $\hat{\alpha}_{RR}$ is

MSE ($\hat{\alpha}_{RR}$) = Variance ($\hat{\alpha}_{RR}$) + [Bias ($\hat{\alpha}_{RR}$)] ^{2}

$$= \hat{\sigma}^2 \sum_{i=1}^{p} \lambda_i / (\lambda_i + k)^2 + k^2 \sum_{i=1}^{p} \hat{\alpha}_i^{\,2} / (\lambda_i + k)^2, \tag{5}$$

where, $\hat{\alpha}_i$ is the i[th] element of $\hat{\alpha}_{OLS}$, $i = 1,2,\ldots,p$, λ_i is the i[th] eigen value of $X'X$ and $\hat{\sigma}^2$ is the OLS estimator of σ^2 i.e. $\hat{\sigma}^2 = (Y'Y - \hat{\alpha}_{OLS}'Z'Y)/(n-p-1)$.

We observe that, when $k = 0$ in (5), MSE of OLS estimator of α is recovered. Hence

$$\text{MSE}(\hat{\alpha}_{OLS}) = \hat{\sigma}^2 \sum_{i=1}^{p} 1/\lambda_i ,$$

Numerous papers have been written, for estimating the α in the presence of multicollinearity. Some of the selected well known methods used for estimation are listed below.

Since ridge estimators can have a serious bias, Singh and Chaubey [30] introduced the Jackknifed ridge regression (JRR) estimator by using the jackknife procedure to reduce the bias of the generalized ridge regression (GRR) estimator. It is given as:

$$\hat{\alpha}_{JRR} = \left[I - k^2(\wedge + kI)^{-2}\right]\hat{\alpha}_{OLS} , \tag{6}$$

and mean square error of $\hat{\alpha}_{JRR}$ is

$$\text{MSE}(\hat{\alpha}_{JRR}) = \hat{\sigma}^2 \sum_{i=1}^{p} \frac{(\lambda_i + 2k)^2}{(\lambda_i + k)^4} + k^4 \sum_{i=1}^{p} \frac{\hat{\alpha}_i^{\,2}}{(\lambda_i + k)^4} .$$

Batah et al. [4] suggested a new ridge estimator namely Modified Jackknifed ridge regression (MJR) estimator by combining idea of GRR and JRR. They established the MSE superiority over both the GRR and JRR estimators. It is given as:

$$\hat{\alpha}_{MJR} = \left[I - k^2(\wedge + kI)^{-2}\right]\left[I - k(\wedge + kI)^{-1}\right]\hat{\alpha}_{OLS} . \tag{7}$$

Also, mean square error of $\hat{\alpha}_{MJR}$ is

$$\text{MSE}(\hat{\alpha}_{MJR}) = \hat{\sigma}^2 \sum_{i=1}^{p} \frac{\left[\lambda_i(\lambda_i + 2k)\right]^2}{(\lambda_i + k)^6} + k^2 \sum_{i=1}^{p} \frac{\left[(\lambda_i + k)^2 + k\lambda_i\right]^2 \hat{\alpha}_i^{\,2}}{(\lambda_i + k)^6} .$$

Recently, Fallah and Salam, [7] introduce an alternative shrinkage estimator, called modified unbiased ridge (MUR) estimator. This estimator is obtained from Unbiased Ridge Regression (URR), and stated that the MUR estimator is more efficient and more reliable than OLS, RR, and URR estimators based on Matrix Mean Squared Error (MMSE). It is given as:

$$\hat{\alpha}_{MUR} = \left[I - k(\wedge + kI)^{-1}\right](\wedge + kI)^{-1}\left(Z'Y + kJ\right) , \tag{8}$$

where $J = \sum_{i=1}^{p} \hat{\alpha}_{iOLS} \Big/ p$ and mean square error of $\hat{\alpha}_{MUR}$ is

$$\text{MSE}\,(\hat{\alpha}_{MUR}) = \hat{\sigma}^2 \sum_{i=1}^{p} \frac{\lambda_i^{\,2}}{(\lambda_i + k)^3} + k^2 \sum_{i=1}^{p} \frac{(\lambda_i + k)\hat{\alpha}_i^{\,2}}{(\lambda_i + k)^3}.$$

Much of the discussions on ridge regression concern the problem of finding good empirical value of k . Many different techniques for estimating k have been proposed or suggested by various researchers. To mention a few, Hoerl and Kennard [15], Hoerl et al. [16], McDonald and Galarneau [24] , Hocking et al. [14], Lawless and Wang [20], Kibria [18], Dempster et al. [5], Gunst and Mason [11], Wichern and Churchill [33], Hemmerle and Brantle [13], Lawless [21], Golub et al. [9], Gibbons [8], Nordberg [28], Saleh and Kibria [29], Haq and Kibria [12], Singh and Tracy [31], Wencheko [32], Kibria [19], Khalaf and Shukur [17], Zhang and Ibrahim [32], Alkhamisi et al. [3], Alkhamisi Shukur [2], Mardikyan and Cetin [23], Batah et al. [4], Muniz and Kibria [26], Mansson et al. [22], Dorugade and Kashid [6], Al-Hassan [1] and very recently, Muniz et al. [27] among others.

Some of the well known methods for choosing ridge parameter value are listed below.

(1) $k_1 = \dfrac{p\hat{\sigma}^2}{\hat{\alpha}'\hat{\alpha}}$ (Hoerl, Kennard and Baldwin, [16]) (9)

(2) $k_2 = \dfrac{p\hat{\sigma}^2}{\sum_{i=1}^{p} \lambda_i \hat{\alpha}_i^{\,2}}$ (Lawless and Wang, [20]) (10)

(3) $k_3 = (\lambda_{\max}\hat{\sigma}^2) \big/ ((n-p-1)\hat{\sigma}^2 + \lambda_{\max}\hat{\alpha}^2{}_{\max})$ (Khalaf and Shukur, [17]) (11)

where, $\lambda_{\max} = \max(\lambda_i)$ and $\hat{\alpha}_{\max} = \max(\hat{\alpha}_i)$.

(4) $k_4 = \max\left(0, \dfrac{p\hat{\sigma}^2}{\hat{\alpha}'\hat{\alpha}} - \dfrac{1}{n(VIF_j)_{\max}}\right)$ (Dorugade and Kashid, [6]) (12)

where $VIF_j = \dfrac{1}{1 - R_j^{\,2}}$ $j = 1,2,\ldots,p$ is variance inflation factor of jth regressor.

(5) $k_5 = \hat{\sigma}^2 \sum_{i=1}^{p} \left(\lambda_i \hat{\alpha}_i^{\,2}\right) \Big/ \left[\sum_{i=1}^{p} \left(\lambda_i \hat{\alpha}_i^{\,2}\right)\right]^2$ Hocking et al, [14]) (13)

(6) $k_6 = \left\{\hat{\sigma}^2\lambda_{\max}\sum_{i=1}^{p}\left(\lambda_i\,\hat{\alpha}_i^{\,2}\right)+\left[\sum_{i=1}^{p}\left(\lambda_i\,\hat{\alpha}_i^{\,2}\right)\right]^2\right\}\Big/\lambda_{\max}\sum_{i=1}^{p}\left(\lambda_i\,\hat{\alpha}_i^{\,2}\right)$ (Al- Hassan, [1]) (14)

(7) $k_7 = p\hat{\sigma}^2\Big/\sum_{i=1}^{p}\left\{\hat{\alpha}_i^{\,2}\Big/\left[\left[\left(\hat{\alpha}_i^{\,4}\lambda_i^2/4\hat{\sigma}^2\right)+\left(6\hat{\alpha}_i^{\,4}\lambda_i/\hat{\sigma}^2\right)\right]^{1/2}-\left(\hat{\alpha}_i^{\,2}\lambda_i/2\hat{\sigma}^2\right)\right]\right\}$

(Batah et al., [4]) (15)

(8) $k_8 = \frac{1}{p}\sum_{i=1}^{p}\frac{\hat{\sigma}^2}{\hat{\alpha}_i^2}$ (Kibria, [19]) (16)

(9) $k_9 = Median\left(\frac{\hat{\sigma}^2}{\hat{\alpha}_i^{2'}}\right)$ $i=1,2,\ldots,p$ (Kibria, [19]) (17)

(10) $k_{10} = \frac{\hat{\sigma}^2}{\left(\prod_{i=1}^{p}\hat{\alpha}_i^2\right)^{1/p}}$ (Kibria, [19]) (18)

(11) $k_{11} = \frac{\hat{\sigma}^2}{\hat{\alpha}_{\max}^2}$ (Hoerl and Kennard, [15]) (19)

(12) $k_{12} = \frac{\hat{\sigma}^2}{\max(\hat{\alpha}_i)}$ (Hoerl and Kennard, [15]) (20)

(13) $k_{13} = \frac{\hat{\sigma}^2}{\sum_{i=1}^{p}\hat{\alpha}_i^{\,2}}$ (Hoerl and Kennard, [15]) (21)

(14) $k_{14} = \max\left(\frac{\hat{\sigma}^2}{\hat{\alpha}_i^2}+\frac{1}{\lambda_i}\right)$ $i=1,2,\ldots,p$ (Alkhamisi and Shukur, [2]) (22)

All these methods of estimating ridge parameter are used in Section 3. The primary aim in this article is to introduce a new estimation method for ridge parameter and hence provide an alternative estimation method for unknown regression parameters to overcome multicollinearity in the linear regression. Then well known ridge estimators reviewed in this article are compared. The organization of the paper is as follows. We propose the new ridge parameter in Section 2. A numerical example and simulation study has been conducted to study the performance of proposed ridge parameter by comparing it with the other ridge parameters. Performance of various ridge parameters through different ridge estimators in terms of mean square error (MSE) criterion are given in Section 3. Finally some concluding remarks are given in Section 4.

2. Proposed estimator for ridge parameter

It is well known that $\frac{\hat{\sigma}^2}{\hat{\alpha}_{\max}^2}$ is the upper bound of ridge parameter used in ridge regression stated by Hoerl and Kennard [15]. Based on the same upper bound Alkhamisi and Shukur [2] present a new method to estimate the ridge parameter k, which is presented below:

$$k_{AS} = \frac{\hat{\sigma}^2}{\hat{\alpha}_{\max}^2} + \frac{1}{\lambda_{\max}}$$

Furthermore, by taking the maximum of both terms as an alternative to the above method they derived the following estimator of the ridge parameter

$$k_{14} = \max\left(\frac{\hat{\sigma}^2}{\hat{\alpha}_i^2} + \frac{1}{\lambda_i}\right) \qquad i = 1,2,\ldots,p$$

In this article, we present a new method to estimate the ridge parameter k. Initially, we suggest the modification by taking the minimum instead of taking the maximum of the term used in above (22), that is $\min\left(\frac{\hat{\sigma}^2}{\hat{\alpha}_i^2} + \frac{1}{\lambda_i}\right)$. Also, it is well known that $\hat{\sigma}^2$ plays an important role in most of the ridge parameters suggested by various researchers. We suggest our ridge parameter by multiplying $\hat{\sigma}^2$ to the reciprocal of $\min\left(\frac{\hat{\sigma}^2}{\hat{\alpha}_i^2} + \frac{1}{\lambda_i}\right)$, which is presented below:

$$k_{\sigma} = \frac{\hat{\sigma}^2}{\min\left(\frac{\hat{\sigma}^2}{\hat{\alpha}_i^2} + \frac{1}{\lambda_i}\right)} \qquad i = 1,2,\ldots,p \tag{23}$$

where, $\hat{\alpha}_i$ is the ith element of $\hat{\alpha}_{OLS}$, $i = 1,2,\ldots,p$, λ_i is the ith eigen value of $X'X$ and $\hat{\sigma}^2$ is the OLS estimator of σ^2 i.e. $\hat{\sigma}^2 = (Y'Y - \hat{\alpha}_{OLS}'Z'Y)/(n-p-1)$.

3. Performance of the proposed ridge parameter

3.1. Numerical Example

To illustrate the performance of our proposed estimator of k, we give a numerical example to investigate the estimators discussed in the dataset which was discussed in Gruber [10]. Condition number of the dataset is 146.4233, which clearly indicates that data exists with multicollinearity. Data shows Total National Research and Development Expenditures as a Percent of Gross National Product by Country: 1972–1986. It represents the relationship between the dependent variable Y the percentage spent by the United States and the four other independent variables X_1, X_2, X_3 and X_4.

It is convenient to make comparison among proposed ridge parameter and other ridge parameters given in (9) to (22). We compute estimated MSE values of well known ridge regression estimator $\hat{\alpha}_{RR}$ using (5) at different ridge parameters, and the values are reported in Table 1.

Table 1: Values of MSE ($\hat{\alpha}_{RR}$) at k

k	MSE($\hat{\alpha}_{RR}$)	k	MSE($\hat{\alpha}_{RR}$)	k	MSE($\hat{\alpha}_{RR}$)
k_1	0.1255	k_6	0.2397	k_{11}	0.1437
k_2	0.1335	k_7	0.1555	k_{12}	0.1727
k_3	0.1447	k_8	0.2458	k_{13}	0.1803
k_4	0.1269	k_9	0.1262	k_{14}	0.6117
k_5	0.1459	k_{10}	0.1492	k_σ	0.1255

From Table 1, we can see that the estimated MSE value of $\hat{\alpha}_{RR}$ at the proposed ridge parameter k_σ is equivalent to the MSE value of $\hat{\alpha}_{RR}$ at the well known ridge parameter k_1 proposed by Hoerl et al. [16]. But, at the same time, we have to note that the estimated MSE value of $\hat{\alpha}_{RR}$ at the k_σ is always smaller than at the remaining ridge parameters.

3.2. Simulation Study

Part A:

We are now ready to illustrate the behavior of the proposed ridge parameter via a simulation study. The simulation is carried out under different degrees of multicollinearity. Proposed estimator for estimating ridge parameter k is then compared in the sense of MSE criterion to the other ridge parameters reviewed in this article. We consider the true model as $Y = X\beta + \varepsilon$. Following McDonald and Galerneau [24] the explanatory variables are generated by

$$x_{ij} = (1-\rho^2)^{1/2} u_{ij} + \rho u_{ip}, \quad i = 1,2,...,n \qquad j = 1,2,..., p.$$

where, u_{ij} are independent standard normal pseudo-random numbers and ρ is specified so that the theoretical correlation between any two explanatory variables is given by ρ^2.

To examine the robustness of all parameters under consideration, random numbers are generated for the error terms (ε) from each of the normal, t, F, Chi-square and exponential distributions respectively. In this study, to investigate the effects of different degrees of multicollinearity on the estimators, we consider two different correlations, ρ = 0.95, 0.99. β parameter vectors are chosen arbitrarily such as $\beta = (4, 1, 2, 5)'$ and $\beta = (2, 1, 2)'$ for p = 4 and 3, respectively. We assumed samples of size of 20 and 60. We compute estimated MSE values of $\hat{\alpha}_{RR}$ using (5) for the ridge parameters given in (9) to (22) including proposed ridge parameter k_σ . The experiment is repeated 1000 times and obtained the average MSE (AMSE) of estimators and values are reported in Tables 2 and 3. We consider the method that leads to the minimum AMSE to the best from the MSE point of view.

From Tables 2 and 3 (See Appendix), we observe that performance of our proposed ridge parameter k_σ is equivalent but slightly better than ridge parameter k_1 proposed by Hoerl et al. [16]. However, table values clearly shown that k_σ is more efficient in terms of MSE than rest of the ridge parameters used in the simulation study for various values of triplet (ρ , n, distribution of ε).

Part B:

In part A, of the simulation study we compared proposed and other ridge parameters among themselves in the sense of MSE criterion. Here, we demonstrate the performance different estimation methods alternate to OLS estimator which are used to estimate unknown regression parameters in the presence of multicollinearity. Obviously, in the present comparative study we used different well known ridge estimators computed using different ridge parameters including proposed ridge parameter k_σ . We consider a simulated data on two different models I and II. For model I and II we have generated random samples respectively from N_4 (0, Σ_1) and N_3 (0, Σ_2).

where $\Sigma_1 = \begin{bmatrix} 1 & 0.2290 & -0.8240 & -0.2450 \\ 0.2290 & 1 & -0.139 & -0.973 \\ -0.8240 & -0.139 & 1 & -0.030 \\ -0.2450 & -0.973 & -0.030 & 1 \end{bmatrix}$ and $\Sigma_2 = \begin{bmatrix} 1 & 0.67 & 0.99 \\ 0.67 & 1 & 0.698 \\ 0.99 & 0.698 & 1 \end{bmatrix}$.

For both models, we have generated observations on a response variable Y with the following specifications on number of predictors (p), sample size (n) and distribution of the error term (ε).

Model	Model Specification	P	N	ε
I	$Y = 2X_1 + X_2 + X_3 + \varepsilon$	4	25	$\varepsilon \sim N(0,5)$
II	$Y = 20 + 2X_2 + 5X_3 + \varepsilon$	3	60	$\varepsilon \sim N(0,3)$

Obviously, multicollinearity is present in the simulated data generated from both the models. Using this data, we compute different estimators given in (4) and (6) to (8) by using each ridge parameters given in (9) to (22) including proposed ridge parameter k_σ. The experiment is repeated 1000 times and obtained the average MSE (AMSE) of estimators using the following expression.

$$AMSE(\hat{\alpha}) = \frac{1}{1000}\sum_{i=1}^{p}\sum_{j=1}^{1000}(\hat{\alpha}_{ij} - \alpha_i)^2 .$$

where, $\hat{\alpha}_{ij}$ denote the estimator of the ith parameter in the jth replication and α_i, $i = 1,2,\ldots,p$ are the true parameter values. The computed values of AMSE are reported in Table 4 for both the models I and II.

From Table 4 (See Appendix), for both models I and II we made the following observations:

1. AMSE values of $\hat{\alpha}_{RR}$, $\hat{\alpha}_{JRR}$, $\hat{\alpha}_{MJR}$ and $\hat{\alpha}_{MUR}$ at $k = k_\sigma$ are less than AMSE ($\hat{\alpha}_{OLS}$)
2. AMSE values of $\hat{\alpha}_{RR}$, $\hat{\alpha}_{JRR}$, $\hat{\alpha}_{MJR}$ and $\hat{\alpha}_{MUR}$ at $k = k_\sigma$ are less than AMSE values of $\hat{\alpha}_{RR}$, $\hat{\alpha}_{JRR}$, $\hat{\alpha}_{MJR}$ at $k = k_i$ $\quad i = 1,2,\ldots,14$
3. AMSE($\hat{\alpha}_{MUR}$) $\leq$ AMSE($\hat{\alpha}_{MJR}$) $\leq$ AMSE($\hat{\alpha}_{RR}$) $\leq$ AMSE($\hat{\alpha}_{JRR}$) at $k = k_\sigma$

It clearly indicates that suggested ridge parameter k_σ is the proper choice of ridge parameter among all the ridge parameters reviewed in this article. Whereas $\hat{\alpha}_{MUR}$ computed at $k = k_\sigma$ is the best alternative estimator to the OLS estimator in the presence of multicollinearity.

4. Conclusion

In this article we suggest a new approach to obtain an estimator of the ridge parameter k. New estimator is evaluated and compared with well-known existing estimators in terms of MSE criterion. Also, well-known existing estimators of the unknown regression parameters are evaluated using different ridge parameters including proposed ridge parameter. The investigation has been carried out using Monte Carlo simulations. Here we try to suggest the best estimation method for estimating unknown regression parameters based on proper choice of ridge parameter. Finally, from a numerical example and the simulation study, we found that the performance of the proposed estimator of the ridge parameter k is satisfactory in the presence of multicollinearity over the other estimators reviewed in this article.

Acknowledgements

The author is very grateful to the reviewers and the editor for their valuable comments and constructive suggestions which certainly improved the quality of the paper in the present version.

References

1. Y. Al-Hassan. Performance of new ridge regression estimators. Journal of the Association of Arab Universities for Basic and Applied Science, 9: 23–26, 2010.
2. M.A. Alkhamisi and G. Shukur. A Monte carlo study of recent ridge parameters. Commun. Statist. Simulation and Computation, 36 (3):535–547, 2007, DOI: 10.1080/03610910701208619
3. M.A. Alkhamisi, G. Khalaf and G. Shukur. Some modifications for choosing ridge parameter. Communications in Statistics, 35:1–16, 2006.
4. F.S. Batah,T. Ramnathan and S.D. Gore. The efficiency of modified jackknife and ridge type regression estimators: a comparison. Surveys in Mathematics and its Applications, 24(2): 157–174, 2008.
5. A. P. Dempster, M. Schatzoff and N. Wermuth. A simulation study of alternatives to ordinary least squares. Journal of the American Statistical Association, 72:77–91, 1977 DOI: 10.1080/01621459.1977.10479910
6. A.V. Dorugade and D. N. Kashid. Alternative method for choosing ridge parameter for regression. International Journal of Applied Mathematical Sciences, 4(9): 447 – 456, 2010.
7. M. E. Fallah and A.E.Salam. A Modification of the Ridge Type Regression Estimators. American Journal of Applied Sciences, 8 (1): 97–102, 2011, DOI: 10.3844/ajassp.2011.97.102
8. D. G. Gibbons. A simulation study of some ridge estimators. Journal of the American Statistical Association, 76:131–139, 1981, DOI: 10.1080/01621459.1981.10477619
9. G. H. Golub, M. Health and G. Wahaba. Generalized cross validation as a method for choosing a good ridge parameter. Technometrics, 21:215–223, 1979, DOI: 10.1080/00401706.1979.10489751
10. M.H.J Gruber. Improving efficiency by shrinkage the James–Stein and Ridge regression estimators. NewYork: Marcell Dekker, 1998.
11. R. F. Gunst and Mason R. L . Biased estimation in regression: An evaluation using mean squared error. Journal of the American Statistical Association, 72:616–628, 1977, DOI: 10.1080/01621459.1977.10480625

12. M. S.Haq and B. M. G. Kibria. A shrinkage estimator for the restricted linear regression model: ridge regression approach. Journal of Applied Statistical Science, 3(4):301–316, 1996.

13. W. J. Hemmerle and Brantle, T. F. Explicit and Constraint generalized ridge estimator. Technometrics, 2:109–120, 1978.R.R., DOI: 10.1080/00401706.1978.10489634

14. Hocking , F.M. Speed and M.J. Lynn. A Class of biased estimators in linear regression. Technometrics, 18 (4): 425–437,1976, DOI: 10.1080/00401706.1976.10489474

15. A.E. Hoerl and R.W. Kennard. Ridge regression: Biased estimation for non orthogonal problems. Technometrics, 12: 55–68, 1970a, DOI:10.1080/00401706.1970.10488634

16. A.E. Hoerl, R.W. Kennard and K.F. Baldwin. Ridge regression: Some Simulations. Commun. Statist, 4: 105–123, 1975, DOI: 10.1080/03610927508827232

17. G. Khalaf and G. Shukur. Choosing ridge parameter for regression problem. Commun. Statist –Theory and Methods, 34: 1177–1182, 2005.

18. B. M. G. Kibria. On preliminary test ridge regression estimators for linear restriction in a regression model with non-normal disturbances. Communications in Statistics, A25:2349–2369, 1996, DOI: 10.1080/03610929608831843

19. B.M.G. Kibria. Performance of some new ridge regression estimators. Commun. Statist – Simulation and Computation, 32 (2): 419–435, 2003.

20. J.F. Lawless and P. Wang. A Simulation study of ridge and other regression estimators. Commun. Statist –Theory and Methods, 14: 1589–1604, 1976.

21. J.F. Lawless. Ridge and related estimation procedure. Communications in Statistics, A7:139–164, 1978.

22. K. Mansson, G. Shukur and B. M. G. Kibria. A simulation study of some ridge regression estimators under different distributional assumptions. Communications in Statistics-Simulations and Computations, 39(8): 1639 – 1670, 2010, DOI: 10.1080/03610918.2010.508862

23. S. Mardikyan, and E. Cetin. Efficient Choice of Biasing Constant for Ridge Regression. Int.J.Contemp.Math.Sciences, 3(11): 527-547, 2008.

24. G.C. McDonald and D.I. Galarneau. A Monte Carlo evaluation of some ridge-type estimators. J. Amer. Statist. Assoc., 20: 407–416, 1975, DOI: 10.1080/01621459.1975.10479882

25. D.C. Montgomery, E.A. Peck. and G.G. Vining. Introduction to linear regression analysis. New York: Wiley, 2006.

26. G. Muniz and B. M. G. Kibria. On some ridge regression estimators: An Empirical Comparisons. Communications in Statistics-Simulation and Computation, 38: 621- 630, 2009, DOI: 10.1080/03610910802592838

27. G. Muniz and B. M. G. Kibria, M. Kristofer and G. Shukur. On Developing Ridge Regression Parameters: A Graphical investigation. SORT-Statistics and Operations Research Transactions, 36 (2):115-138, 2012.

28. L. Nordberg . A procedure for determination of a good ridge parameter in linear regression. Communications in Statistics, A11:285–309, 1982.

29. A. K. Md. Saleh and B. M. G. Kibria. Performances of some new preliminary test ridge regression estimators and their properties. Communications in Statistics—Theory and Methods, 22:2747–2764, 1993, DOI: 10.1080/03610929308831183

30. B. Singh and Y.P. Chaubey. On some improved ridge estimators. Statist. Papers, 28: 53–67, 1987, DOI: 10.1007/BF02932590

31. S. Singh and D. S. Tracy. Ridge-regression using scrambled responses. Metrika, 147–157, 1999.

32. E. Wencheko. Estimation of the signal-to-noise in the linear regression model. Statistical Papers, 41:327–343, 2000, DOI: 10.1007/BF02925926

33. D. W. Wichern and G. A. Churchill. A comparison of ridge estimators. Technometrics, 20:301–311, 1978, DOI: 10.1080/00401706.1978.10489675

34. J. Zhang and M. Ibrahim. A simulation study on SPSS ridge regression and ordinary least squares regression procedures for multicollinearity data. Journal of Applied Statistics, 32, 571-588, 2005,
DOI: 10.1080/02664760500078946

Appendix

Table 2: Values of AMSE of $\hat{\alpha}_{RR}$ at k (p = 4, $\beta = (4,1,2,5)'$ and ρ = 0.95)

k	n = 20					n = 60				
	distribution of error term (ε)					distribution of error term (ε)				
	$N(0,1)$	$t_{(4)}$	$F(2,3)$	$\chi^2_{(5)}$	$Exp(mean=6)$	$N(0,1)$	$t_{(4)}$	$F(2,3)$	$\chi^2_{(5)}$	$Exp(mean=6)$
k_1	0.0554	0.1225	0.2586	0.3479	0.0681	0.1140	0.0408	0.3164	0.1478	0.1435
k_2	0.0460	0.1070	0.2610	0.4403	0.0618	0.0701	0.0426	0.3685	0.1359	0.1305
k_3	0.0629	0.2762	0.3365	0.3541	0.0775	0.2060	0.0420	0.3416	0.1904	0.1789
k_4	0.0677	0.1235	0.2598	0.3437	0.0786	0.1288	0.0481	0.3160	0.1588	0.1555
k_5	0.0568	0.1700	0.2685	0.3404	0.0703	0.1741	0.0409	0.3002	0.1544	0.1507
k_6	1.1105	0.6116	0.8892	1.2165	1.1651	0.5027	1.3328	0.5783	1.1917	1.3129
k_7	0.0475	0.1141	0.2741	0.5186	0.0700	0.0298	0.0622	0.3735	0.1615	0.1533
k_8	0.0665	1.2901	1.7076	0.5344	0.0843	0.2701	0.0439	1.2561	0.2341	0.2127
k_9	0.0429	0.1082	0.2599	0.4115	0.0622	0.0687	0.0406	0.3815	0.1356	0.1307
k_{10}	0.0471	0.1094	0.2592	0.3750	0.0624	0.0654	0.0407	0.3452	0.1367	0.1297
k_{11}	0.0629	0.2692	0.3303	0.3538	0.0775	0.2058	0.0420	0.3174	0.1903	0.1789
k_{12}	0.0615	0.2666	0.3312	0.3494	0.0761	0.1925	0.0418	0.3399	0.1811	0.1703
k_{13}	0.0638	0.4064	0.4818	0.4129	0.0798	0.2184	0.0429	0.3852	0.2040	0.1901
k_{14}	2.3012	1.5099	1.6636	2.2821	2.1018	1.8089	2.3138	1.0531	2.2434	2.4342
k_σ	0.0415	0.1064	0.2573	0.35225	0.0515	0.0544	0.0384	0.2992	0.1322	0.1265
MSE($\hat{\alpha}_{OLS}$)	0.0674	2.5617	4.8188	0.6030	0.0859	0.2916	0.0443	6.8640	0.2467	0.2218

Table 3: Values of AMSE of $\hat{\alpha}_{RR}$ at k (p = 3, $\beta = (2,1,2)'$ and ρ = 0.99)

k	n = 20					n = 60				
	distribution of error term (ε)					distribution of error term (ε)				
	$N(0,1)$	$t_{(4)}$	$F(2,3)$	$\chi^2_{(5)}$	$Exp(mean = 6)$	$N(0,1)$	$t_{(4)}$	$F(2,3)$	$\chi^2_{(5)}$	$Exp(mean = 6)$
k_1	0.2077	0.1919	0.5923	0.1042	0.0887	0.1003	0.0873	39.8956	0.0827	0.0869
k_2	0.1854	0.1732	0.6162	0.0844	0.1015	0.0820	0.0660	121.3510	0.0840	0.0660
k_3	0.3194	0.1731	0.6329	0.1395	0.0921	0.1300	0.1210	114.6880	0.0900	0.1230
k_4	0.2160	0.1999	0.5923	0.1163	0.0930	0.1129	0.1024	37.2535	0.0893	0.0993
k_5	0.2386	0.2087	0.5923	0.1210	0.0896	0.1110	0.1010	118.2620	0.0850	0.1120
k_6	0.9878	1.0979	0.6195	0.7860	1.0096	0.9310	0.8290	121.7670	1.1120	0.5720
k_7	0.2021	0.1917	0.6170	0.0790	0.1632	0.0784	0.0382	74.6223	0.1215	0.0656
k_8	0.5057	0.4402	0.8817	0.1708	0.1021	0.1522	0.1428	29.1109	0.1004	0.1504
k_9	0.1854	0.1753	0.5924	0.0797	0.0892	0.0780	0.0780	121.0980	0.0820	0.0640
k_{10}	0.1868	0.2068	0.5945	0.0867	0.0903	0.0850	0.0690	116.4570	0.0820	0.0500
k_{11}	0.3192	0.1731	0.6087	0.1395	0.0921	0.1300	0.1210	116.1240	0.0900	0.1230
k_{12}	0.2935	0.1930	0.6317	0.1360	0.0916	0.1260	0.1180	115.5180	0.0890	0.1210
k_{13}	0.3705	0.3299	0.6503	0.1502	0.0966	0.1369	0.1274	25.7538	0.0945	0.1322
k_{14}	2.1475	2.0030	0.8580	1.8594	2.2004	1.9700	1.7700	122.7690	2.1040	1.6660
k_σ	0.1783	0.1636	0.5859	0.0765	0.0317	0.0690	0.0630	36.0660	0.0800	0.0480
MSE($\hat{\alpha}_{OLS}$)	0.5738	0.4946	46.4675	0.1789	0.1044	0.1581	0.1487	32.4962	0.1027	0.1575

Table 4: Values of AMSE for OLS estimator and various ridge estimators at k

k	Model I					Model II				
	$\hat{\alpha}_{RR}$	$\hat{\alpha}_{JRR}$	$\hat{\alpha}_{MJR}$	$\hat{\alpha}_{MUR}$	$\hat{\alpha}_{OLS}$	$\hat{\alpha}_{RR}$	$\hat{\alpha}_{JRR}$	$\hat{\alpha}_{MJR}$	$\hat{\alpha}_{MUR}$	$\hat{\alpha}_{OLS}$
k_1	1.1742	1.5932	0.9930	0.9490	3.3543	1.5627	1.9270	1.5000	1.4930	2.4256
k_2	1.2415	1.3658	1.2120	1.1360	3.3543	1.4410	1.5040	1.4000	1.4860	2.4256
k_3	1.6045	2.2941	1.3510	1.1390	3.3543	1.9123	2.2740	1.8000	1.7690	2.4256
k_4	1.1841	1.6120	1.0000	0.9520	3.3543	1.5863	1.9600	1.5000	1.5000	2.4256
k_5	1.2243	1.2360	1.1580	1.0810	3.3543	1.2751	1.5390	1.4000	1.4660	2.4256
k_6	1.1960	1.0284	1.2100	2.2080	3.3543	1.4710	1.5210	1.6000	2.3340	2.4256
k_7	1.1230	1.2302	1.0710	1.1490	3.3543	1.4012	1.5560	1.4000	1.5610	2.4256
k_8	2.6651	3.1776	2.5680	2.2640	3.3543	2.2775	2.4130	2.3000	2.1890	2.4256
k_9	1.1628	1.5058	1.0150	0.9820	3.3543	1.3901	1.4500	1.3000	1.4650	2.4256
k_{10}	1.1403	1.4137	0.9800	0.9910	3.3543	1.3938	1.6190	1.4000	1.4680	2.4256
k_{11}	1.8333	2.2135	1.8060	1.6500	3.3543	1.8978	2.2670	1.8000	1.7670	2.4256
k_{12}	1.9861	2.6327	1.3460	1.7660	3.3543	1.8887	2.3480	1.7000	1.7230	2.4256
k_{13}	1.8622	2.5831	1.6450	1.3720	3.3543	1.9887	2.3130	1.9000	1.8060	2.4256
k_{14}	30.3672	27.1970	33.4940	32.9580	3.3543	28.0911	22.7160	33.4000	35.1260	2.4256
k_{σ}	1.1141	1.4053	0.9690	0.8950	3.3543	1.3016	1.4280	1.2900	1.4210	2.4256