Journal of the Association of Arab Universities for Basic and Applied Sciences (2014) **15**, 94–99

University of Bahrain
Journal of the Association of Arab Universities for Basic and Applied Sciences
www.elsevier.com/locate/jaaubas
www.sciencedirect.com

معاملات تراجع جديدة لتراجع القمة

A.V. Dorugade

Y C Mahavidyalaya Halkarni, Tal-Chandgad, Kolhapur 416552, Maharashtra, India

الملخص:

في عام (1970م) قام هورل و كنارد (Hoerl and Kennard) بتقديم مقدر تراجع القمة كبديل لتقدير المربعات الصغرى الاعتيادية (OLS) عند وجود علاقات خطية متداخلة متعددة. عند تراجع القمة، فان معاملات التراجع تلعب دور مهم في تقدير المعاملات. في هذا البحث، تم اقتراح طريقة جديدة لتقدير معاملات القمة في حالتي تراجع القمة العادي (ORR) والعام (GRR). ان دراسة المحاكاة قدمت تقييم لاداء التقدير المقترح بناء على معيار متوسط الخطأ المربعي (MSE)، وتبين أنه تحت شروط معينة فان المقدر المقترح ذا اداء جيد بالمقارنة مع طريقة تقييم المربعات الصغرى الاعتيادية، وطرق أخرى معروفة للمقدرات التي تم استعراضها في هذا البحث.

Journal of the Association of Arab Universities for Basic and Applied Sciences (2014) 15, 94–99

University of Bahrain
Journal of the Association of Arab Universities for Basic and Applied Sciences
www.elsevier.com/locate/jaaubas
www.sciencedirect.com

REVIEW ARTICLE

New ridge parameters for ridge regression

A.V. Dorugade *

Y C Mahavidyalaya Halkarni, Tal-Chandgad, Kolhapur 416552, Maharashtra, India

Received 8 October 2012; revised 9 February 2013; accepted 31 March 2013
Available online 10 May 2013

KEYWORDS

Ridge regression;
Ridge parameter;
Multicollinearity

Abstract Hoerl and Kennard (1970a) introduced the ridge regression estimator as an alternative to the ordinary least squares (OLS) estimator in the presence of multicollinearity. In ridge regression, ridge parameter plays an important role in parameter estimation. In this article, a new method for estimating ridge parameters in both situations of ordinary ridge regression (ORR) and generalized ridge regression (GRR) is proposed. The simulation study evaluates the performance of the proposed estimator based on the mean squared error (MSE) criterion and indicates that under certain conditions the proposed estimators perform well compared to OLS and other well-known estimators reviewed in this article.

1. Introduction

In the presence of multicollinearity OLS estimator yields regression coefficients whose absolute values are too large and whose signs may actually reverse with negligible changes in the data (Buonaccorsi, 1996). Whenever the multicollinearity presents in the data, the OLS estimator performs 'poorly'. The method of ridge regression, proposed by Hoerl and Kennard (1970a) is one of the most widely used tools to the problem of multicollinearity. In a ridge regression an additional parameter, the ridge parameter (k) plays an vital role to control the bias of the regression toward the mean of the response variable. In addition, they proposed the generalized ridge regression (GRR) procedure that allows separate ridge parameter for each regressor. Using GRR, it is easier to find optimal values of ridge parameter, i.e., values for which the MSE of the ridge estimator is minimum. In addition, if the optimal values for biasing constants differ significantly from each other then this estimator has the potential to save a greater amount of MSE than the OLS estimator (Stephen and Christopher, 2001). In both ORR and GRR as 'k' increases from zero and continues up to infinity, the regression estimates tend toward zero. Though these estimators result in biased, for certain value of k, they yield a minimum mean squared error (MMSE) compared to the OLS estimator (see Hoerl and Kennard, 1970a). Ridge parameter 'k' proposed by Hoerl et al. (1975) performs fairly well.

Much of the discussions on ridge regression concern the problem of finding good empirical value of k. Recently, many researchers have suggested various methods for choosing ridge parameter in ridge regression. These methods have been suggested by Hoerl and Kennard (1970a), Hoerl et al. (1975), McDonald and Galarneau (1975), Hocking et al. (1976), Lawless and Wang (1976), Gunst and Mason (1977), Lawless (1978), Nomura (1988), Heath and co-workers (1979), Nordberg (1982), Saleh and Kibria (1993), Haq and Kibria (1996), Kibria (2003), Pasha and Shah (2004), Khalaf and

* Tel./fax: +91 02320 239076.
E-mail address: adorugade@rediffmail.com.

Peer review under responsibility of University of Bahrain.

ELSEVIER Production and hosting by Elsevier

1815-3852 © 2013 Production and hosting by Elsevier B.V. on behalf of University of Bahrain.
http://dx.doi.org/10.1016/j.jaubas.2013.03.005

Shukur (2005), Norliza et al. (2006), Alkhamisi and Shukur (2007), Mardikyan and Cetin (2008), Dorugade and Kashid (2010) and Al-Hassan (2010) to mention a few. The objective of the article is to investigate some of the existing popular techniques that are available in the literature and to make a comparison among them based on mean square properties. Moreover, we suggested some methods for estimating ridge parameters in ORR and GRR which produce ridge estimators that yield minimum MSE than other estimators. The organization of the article is as follows.

In this article, we introduce alternative ordinary and generalized ridge estimators and study their performance by means of simulation techniques. Comparisons are made with other ridge-type estimators evaluated elsewhere, and the estimators to be included in this study are described in Section 2. In Section 3, we propose some new methods for estimating the ridge parameter. In Section 4, we illustrate the simulation technique that we have adopted in the study and related results of the simulations appear in the tables and figures. In Section 5, we give a brief summary and conclusion.

2. Model and estimators

Consider, a widely used linear regression model

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

where Y is a $n \times 1$ vector of observations on a response variable. β is a $p \times 1$ vector of unknown regression coefficients, X is a matrix of order $(n \times p)$ of observations on 'p' predictor (or regressor) variables and ε is an $n \times 1$ vector of errors with $E(\varepsilon) = 0$ and $V(\varepsilon) = \sigma^2 I_n$. 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 paper is concerned with data exhibited with multicollinearity leading to a high MSE for β meaning that $\hat{\beta}$ is an unreliable estimator of β.

Let $\wedge$ and T be the matrices of eigen values and eigen vectors of $X'X$, respectively, satisfying $T'X'XT = \wedge =$ diagonal $(\lambda_1, \lambda_2, \ldots, \lambda_p)$, where λi being the ith eigen value of $X'X$ and $T'T = TT' = I_p$ we obtain the equivalent model

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

where $Z = XT$, it implies that $Z'Z = \wedge$, and $\alpha = T'\beta$ (see Montogomery et al. (2006)) 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

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

2.1. Generalized ridge estimator (GRR)

The GRR estimator of α is defined by

$$\hat{\alpha}_{GR} = (I - KA^{-1})\hat{\alpha}_{OLS}, \tag{4}$$

where $K =$ diagonal$(k_1, k_2 \ldots k_p)$, $k_i \geqslant 0$, i = 1,2,…,p be the different ridge parameters for different regressors and $A = \wedge + K$.

Hence GRR estimator for β is $\hat{\beta}_{GR} = T\hat{\alpha}_{GR}$.
and mean square error of $\hat{\alpha}_{GR}$ is

$$\mathrm{MSE}(\hat{\alpha}_{GR}) = \mathrm{Variance}(\hat{\alpha}_{GR}) + [\mathrm{Bias}(\hat{\alpha}_{GR})]^2$$
$$= \hat{\sigma}^2 \sum_{i=1}^{p} \lambda_i/(\lambda_i + k_i)^2 + \sum_{i=1}^{p} k_i^2 \hat{\alpha}_i^2/(\lambda_i + k_i)^2 \tag{5}$$

In case of GRR, various methods are available in the literature to determine the separate ridge parameter for each regressor. Among these, well known methods for determination of ridge parameter which are used in the further study are given below.

(1) Hoerl and Kennard (1970a) have proposed the following ridge parameter

$$k_i(\mathrm{HK}) = \frac{\hat{\sigma}^2}{\hat{\alpha}_i^2}, \quad i = 1, 2, \ldots, p \tag{6}$$

(2) Nomura (1988) proposed a ridge parameter and it is given by

$$k_i(\mathrm{HMO}) = \frac{\hat{\sigma}^2}{\hat{\alpha}_i^2}\left\{1 + \left[1 + \lambda_i(\hat{\alpha}_i^2/\hat{\sigma}^2)^{1/2}\right]\right\}, \quad i = 1, 2, \ldots, p \tag{7}$$

(3) Troskie and Chalton (1996) proposed a ridge parameter and it is given by

$$k_i(\mathrm{TC}) = \lambda_i\hat{\sigma}^2/\left(\lambda_i\hat{\alpha}_i^2 + \hat{\sigma}^2\right), \quad i = 1, 2, \ldots, p \tag{8}$$

(4) Firinguetti (1999) proposed a ridge parameter and it is given by

$$k_i(F) = \lambda_i\hat{\sigma}^2/\left[\lambda_i\hat{\alpha}_i^2 + (n-p)\hat{\sigma}^2\right], \quad i = 1, 2, \ldots, p \tag{9}$$

(5) Batah et al. (2008) proposed a ridge parameter and it is given by

$$k_i(\mathrm{FG}) = \hat{\sigma}^2\left\{\left[\left(\hat{\alpha}_i^4\lambda_i^2/4\hat{\sigma}^2\right) + \left(6\hat{\alpha}_i^4\lambda_i/\hat{\sigma}^2\right)\right]^{1/2} - \left(\hat{\alpha}_i^2\lambda_i/2\hat{\sigma}^2\right)\right\}/\hat{\alpha}_i^2 \quad i = 1, 2, \ldots, p \tag{10}$$

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

2.2. Ordinary ridge estimator (ORR)

Setting $k_1 = k_2 = \ldots = k_p = k$ and $k \geqslant 0$, the Ordinary ridge regression (ORR) estimator of β is

$$\hat{\beta}_{RR} = T\hat{\alpha}_{RR} = T\left[I - kA_k^{-1}\right]\hat{\alpha}, \quad \text{where } A_k = (\wedge + kI_p) \tag{11}$$

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

$$\mathrm{MSE}(\hat{\alpha}_{RR}) = \mathrm{Variance}(\hat{\alpha}_{RR}) + [\mathrm{Bias}(\hat{\alpha}_{RR})]^2$$
$$= \hat{\sigma}^2 \sum_{i=1}^{p} \lambda_i/(\lambda_i + k)^2 + k^2 \sum_{i=1}^{p} \hat{\alpha}_i^2 \Big/ (\lambda_i + k)^2 \tag{12}$$

We observe that, when $k = 0$ in (12), MSE of OLS estimator of α is recovered. Hence

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

Hoerl et al. (1975) suggested that, the value of 'k' is chosen small enough, for which the mean squared error of ridge estimator, is less than the mean squared error of OLS estimator.

In case of ORR also, many researchers have suggested different ways of estimating the ridge parameter. Some of the well known methods for choosing the ridge parameter value are listed below.

(1) $$k_1 = \frac{p\hat{\sigma}^2}{\hat{\alpha}'\hat{\alpha}} \qquad \text{(Hoerl et al. (1975))} \tag{13}$$

(2) $$k_2 = \frac{p\hat{\sigma}^2}{\sum_{i=1}^{p} \lambda_i \hat{\alpha}_i^2} \qquad \text{(Lawless and Wang (1976))} \tag{14}$$

(3) $$k_3 = p\hat{\sigma}^2 \Big/ \sum_{i=1}^{p} \left\{ \hat{\alpha}_i^2 \Big/ \left[1 + \left(1 + \lambda_i \left(\hat{\alpha}_i^2/\hat{\sigma}^2\right)^{1/2}\right) \right] \right\}$$
Nomura (1988) (15)

(4) $$k_4 = (\lambda_{max}\hat{\sigma}^2)\big/\left((n-p-1)\hat{\sigma}^2 + \lambda_{max}\hat{\alpha}_{max}^2\right)$$
Khalaf and Shukur (2005) (16)

(5) $$k_5 = \max\left(0, \frac{p\hat{\sigma}^2}{\hat{\alpha}'\hat{\alpha}} - \frac{1}{n(\mathrm{VIF}_j)_{max}}\right)$$
Dorugade and Kashid (2010) (17)

where $\mathrm{VIF}_j = \frac{1}{1-R_j^2}$ $j = 1, 2, \ldots, p$ is the variance inflation factor of jth regressor.

(6) $$k_6 = \hat{\sigma}^2 \sum_{i=1}^{p} \left(\lambda_i \hat{\alpha}_i^2\right) \Big/ \left[s \sum_{i=1}^{p} \left(\lambda_i \hat{\alpha}_i^2\right) \right]^2$$
Montogomery et al. (2006) (18)

(7) $$k_7 = \left\{ \hat{\sigma}^2 \lambda_{max} \sum_{i=1}^{p} \left(\lambda_i \hat{\alpha}_i^2\right) + \left[\sum_{i=1}^{p} \left(\lambda_i \hat{\alpha}_i^2\right) \right]^2 \right\} \Big/ \lambda_{max} \sum_{i=1}^{p} \left(\lambda_i \hat{\alpha}_i^2\right)$$
Norliza et al. (2006) (19)

(8) $$k_8 = p\hat{\sigma}^2 / \sum_{i=1}^{p} \left\{ \hat{\alpha}_i^2 / \left[\left[\left(\hat{\alpha}_i^4 \lambda_i^2 / 4\hat{\sigma}^2\right) + \left(6\hat{\alpha}_i^4 \lambda_i / \hat{\sigma}^2\right) \right]^{1/2} - \left(\hat{\alpha}_i^2 \lambda_i / 2\hat{\sigma}^2\right) \right] \right\}$$
Batah et al. (2008) (20)

(9) $$k_9 = \frac{1}{p} \sum_{i=1}^{p} \frac{\hat{\sigma}^2}{\hat{\alpha}_i^2} \qquad \text{Kibria (2003)} \tag{21}$$

(10) $$k_{10} = \mathrm{Median}\left(\frac{\hat{\sigma}^2}{\hat{\alpha}_i^{2\prime}}\right) \qquad i = 1, 2, \ldots, p$$
Kibria (2003) (22)

(11) $$k_{11} = \frac{\hat{\sigma}^2}{\left(\prod_{i=1}^{p} \hat{\alpha}_i^2\right)^{1/p}} \qquad \text{Kibria (2003)} \tag{23}$$

(12) $$k_{12} = \frac{\hat{\sigma}^2}{\hat{\alpha}_{max}^2} \qquad \text{Hoerl and Kennard (1970a)} \tag{24}$$

(13) $$k_{13} = \frac{\hat{\sigma}^2}{\max(\hat{\alpha}_i)} \qquad \text{Hoerl and Kennard (1970a)} \tag{25}$$

(14) $$k_{14} = \frac{\hat{\sigma}^2}{\sum_{i=1}^{p} \hat{\alpha}_i^2} \qquad \text{Hoerl and Kennard (1970a)} \tag{26}$$

(15) $$k_{15} = \max\left(\frac{\hat{\sigma}^2}{\hat{\alpha}_i^2} + \frac{1}{\lambda_i}\right) \qquad i = 1, 2, \ldots, p$$
Alkhamisi and Shukur (2007) (27)

All the methods of estimating ridge parameter are used in Section 4.

3. Proposed ridge parameter

Hoerl and Kennard (1970a) conclude that, bias and total variance of the parameter estimates are, respectively monotonically increasing and decreasing functions of ridge parameters. They also suggested a value of ith ridge parameter k_i(HK) used in GRR given in (6). Hoerl et al. (1975) suggests the modification in k_i(HK) used in ORR which performs fairly well. Lawless and Wang (1976) suggest the modification in k_i(HK) to reduce the bias by multiplying the ith eigen value λ_i to the denominator of (6) to keep the variation depends on the strength of the multicollinearity. Their estimator reduces the bias but results in greater total variance of the parameter estimates.

In this article, we suggest a estimator that takes a little bias than estimator given by Hoerl et al. (1975) and substantially reduces the total variance of the parameter estimates than the total variance using estimator given by Lawless and Wang (1976), thereby improving the mean square error of estimation and prediction. We suggest the modification by multiplying $\lambda_{max}/2$ to the denominator of (6). The suggested estimator is:

$$k_i(\mathrm{AD}) = \frac{2\hat{\sigma}^2}{\lambda_{max}\hat{\alpha}_i^2} \quad i = 1, 2, \ldots, p \tag{28}$$

where λ_{max} is the largest eigen value of $X'X$.

This leads to the denominator of the alternative estimator given by (28) being greater than that of Hoerl and Kennard (1970a) by $\lambda_{max}/2$. Hence, we can write

$$\frac{\hat{\sigma}^2}{\hat{\alpha}_i^2} \geqslant \frac{2\hat{\sigma}^2}{\lambda_{max}\hat{\alpha}_i^2} \geqslant \frac{\hat{\sigma}^2}{\lambda_i\hat{\alpha}_i^2} \qquad i = 1, 2, \ldots, p$$

It clearly indicates that our suggested estimator lies in between the estimators given by Hoerl et al. (1975) and Lawless and Wang (1976). Kibria (2003) suggested optimal ridge parameters by proposing new ridge parameters by modifying the quantity $k_i(\mathrm{HK}) = \hat{\sigma}^2/\hat{\alpha}_i^2$. By adopting algorithms outlined in Kibria (2003), we propose new methods to determine ridge parameters in case of ORR for the ridge parameter k as below,

$$k_1(\mathrm{AD}) = \text{Arithmatic Mean}[k_i(\mathrm{AD})] = \frac{2}{p\lambda_{max}} \sum_{i=1}^{p} \frac{\hat{\sigma}^2}{\hat{\alpha}_i^2} \tag{29}$$

$$k_2(\mathrm{AD}) = \text{Median}[k_i(\mathrm{AD})] = \text{Median}\left(\frac{2\hat{\sigma}^2}{\lambda_{max}\hat{\alpha}_i^2}\right) \quad i = 1, 2, \ldots, p \tag{30}$$

$$k_3(\mathrm{AD}) = \text{Geometric Mean}[k_i(\mathrm{AD})] = \frac{2\hat{\sigma}^2}{\lambda_{max}\left(\prod_{i=1}^{p} \hat{\alpha}_i^2\right)^{1/p}} \tag{31}$$

$$k_4(\mathrm{AD}) = \text{Harmonic Mean}[k_i(\mathrm{AD})] = \frac{2p}{\lambda_{max}} \sum_{i=1}^{p} \frac{\hat{\sigma}^2}{\hat{\alpha}_i^2} \tag{32}$$

Table 1 Ratio of AMSE of OLS over various ridge estimators for different 'k'.

n		20				50				100			
$\hat{\sigma}^2$		1	5	10	25	1	5	10	25	1	5	10	25
k	k_1	1.9591	1.9041	1.8936	1.6576	1.5648	2.5220	1.4035	2.9800	2.2481	1.8036	1.5990	3.1048
	k_2	1.6458	0.9895	0.4923	0.9257	0.9901	1.1969	0.7672	1.7184	1.7886	1.1572	0.6954	2.1013
	k_3	1.3152	1.3529	0.9294	1.2887	1.0035	1.7400	1.0407	2.4535	1.1916	1.2865	1.1165	2.6206
	k_4	1.7382	1.8200	2.1747	1.6185	1.5653	2.2507	1.6252	2.4887	1.9691	1.2809	1.4109	2.3090
	k_5	1.8347	1.9279	1.9896	1.6721	1.5275	2.5228	1.4097	2.9830	2.1999	1.8048	1.6031	3.1061
	k_6	1.6473	0.9834	0.4908	0.9277	0.9900	1.1914	0.7644	1.6984	1.7886	1.1534	0.6935	2.0706
	k_7	0.3247	0.9443	0.4901	0.9241	0.1079	0.9689	0.7559	1.7019	0.0466	0.7116	0.6531	2.0729
	k_8	1.6918	1.3932	0.9431	1.3282	1.2879	1.8878	1.0647	2.5851	1.7534	1.4120	1.1524	2.7445
	k_9	1.3903	1.2905	1.5517	1.3176	1.1746	1.4637	1.2562	1.5268	1.3655	1.3197	1.2091	1.5733
	k_{10}	1.4950	1.0041	0.5202	1.0275	1.2124	1.6613	0.8165	2.5551	1.3958	1.2047	1.2136	2.2630
	k_{11}	1.6073	1.0439	0.5224	0.9630	1.3523	1.6802	0.8153	2.1401	1.4944	1.2839	0.9872	2.2678
	k_{12}	1.0007	1.7249	2.2101	1.6536	0.8251	1.9723	1.6099	2.5894	0.5697	1.0581	1.3364	2.3380
	k_{13}	1.4461	0.0093	0.2329	0.0221	0.4099	0.6154	0.0489	0.6211	0.0860	0.2930	0.5761	0.0650
	k_{14}	1.8815	1.8245	2.1426	1.7185	1.4737	2.1611	1.5981	2.5000	1.8469	1.7491	1.5616	2.5921
	k_{15}	0.13589	0.7942	0.47573	0.8983	0.05383	0.67948	0.71321	1.5857	0.02639	0.44348	0.56721	1.89147
	k_1(AD)	1.2140	1.1574	1.3590	1.2027	1.0938	1.2634	1.1487	1.2918	1.2187	1.1850	1.1156	1.3133
	k_2(AD)	1.4810	1.0225	0.5435	1.1078	1.1589	1.7110	0.8607	2.7609	1.3960	1.2103	1.3920	2.3595
	k_3(AD)	1.6132	1.1010	0.5511	0.9986	1.4388	1.8322	0.8557	2.3839	1.6500	1.3221	1.1685	2.3874
	k_4(AD)	1.9872	2.0282	2.2600	1.8183	1.5990	2.5482	1.6182	2.9882	2.3328	1.8823	1.6926	3.2643

From 29, 30, 31, 30 $\hat{\alpha}_i$ is the ith element of $\hat{\alpha}_{OLS}, i = 1, 2, \ldots, p$ and $\hat{\sigma}^2$ is the OLS estimator of σ^2, i.e. $\hat{\sigma}^2 = \frac{Y'Y - \hat{\alpha}'Z'Y}{n-p-1}$.

Result 1.

$k_1 \geqslant k_4(\text{AD}) \geqslant k_2$

Result 2. If $\lambda_{\max}$ is close to 'p' then $k_4(\text{AD}) \cong 2k_{14}$, and if $\lambda_{\max}$ is close to '1' then $k_4(\text{AD}) \cong 2k_1$.

Hoerl et al. (1975) have shown that $k_1 \leqslant \frac{\sigma^2}{\hat{\alpha}^2_{\max}}$. Using this, inequality from result 1, $k_4(\text{AD}) \leqslant k_1 \leqslant \frac{\sigma^2}{\hat{\alpha}^2_{\max}}$. Hence k_4(AD) satisfies the upper bound of ridge parameter stated by Hoerl and Kennard (1970a).

Proposed estimator is examined by means of a simulation technique which we present in the next section.

4. Performance of the proposed ridge parameter

In this section, we examined the performance of the ridge estimator using the proposed ridge parameters in both ORR and GRR over the different ridge parameters (k) reviewed in this article. We examined the average MSE (AMSE) ratio of the ridge estimator using proposed ridge parameters and other ridge parameters over OLS estimator. Performances of new ridge estimators given in (28)–(32) are studied in two parts. In part A, performance for proposed ridge estimators is evaluated through simulation in case of ORR. Whereas, in part B a simulation study is carried out for evaluating the performance of proposed ridge estimators in case of GRR.

4.1. Part A

We consider the true model as $Y = X\beta + \varepsilon$. Here ε follows a normal distribution $N(0, \sigma^2 I_n)$ and the explanatory variables are generated (see Batah et al., 2008) from

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

where u_{ij} is an independent standard normal random number and ρ^2 is the correlation between x_{ij} and x'_{ij} for $j, j' < p$ and $j \neq j'$. $j, j' = 1,2,\ldots,p$. When j or $j' = p$, the correlation will be 'ρ'. Here we consider predictor variables $p = 4$ and $\rho = 0.9$. These variables are standardized such that $X'X$ is in the correlation form and it is used for the generation of Y with $\beta = (2,3,5,1)'$. We have simulated the data with sample sizes $n = 20$, 50 and 100. The variance of the error terms is taken as $\sigma^2 = 1$, 5, 10 and 25. Ridge estimates are computed using different ridge parameters given in (13)–(27) and (29)–(32). The MSE of such ridge regression parameters are obtained using (12). This experiment is repeated 2000 times and obtains the AMSE. Firstly, we computed the AMSE ratios (AMSE $(\hat{\alpha}_{\text{OLS}})$/AMSE $(\hat{\alpha}_{\text{RR}})$) of OLS estimator over different estimators for various values of triplet (ρ, n, σ^2) and reported in Table 1. We consider the method that leads to the maximum AMSE ratio to the best from the MSE point of view.

In the following figure (Fig. 1), we represent the same values reported in Table 1 Here we noted that values of AMSE ratios only for k_1,k_5,k_{14} and k_4(AD) are represented because these values for remaining choice of 'k' have less importance for the comparative study. Here input values are n, ρ and σ^2. These input values are ordered according to the increase of values. For fixed value of 'ρ' changes values of 'n' and for fixed values of (ρ, n) changes the values of σ^2. There are 12 sets of (ρ, n, σ^2) values. These are arranged as (0.9,20,1), (0.9,20,5),…,(0.9,100,25) and it is numbered as 1, 2,…,12, respectively.

Same procedure for another choice of $p = 3$ and $\beta = (3,1,5)'$ is done and AMSE ratios are computed and represented in Fig. 2.

From Table 1, Figs. 1 and 2, we observe that the performance of proposed ridge parameters k_1(AD),k_2(AD),k_3(AD) and k_4(AD) is better than OLS. Particularly k_4(AD) performs equivalently and is little better than ridge parameters proposed by Hoerl et al. (1975) and Dorugade and Kashid (2010) whereas, it gives better performance than other ridge parameters reviewed in this article for all combinations of correlation between predictors (ρ), sample size (n) and variance of the error term (σ^2) used in this simulation study.

4.2. Part B

Here we evaluate the performance of proposed ridge parameters in case of GRR. Here we generate the data which exhibit with multicollinearity using the procedure for the generation of y with $\beta = (2,3,5,1)'$ as discussed in part A. We consider predictor variables $p = 4$ and $\rho = 0.9$. We have simulated the data with sample sizes $n = 20$, 50 and 100. The variance of the error term is taken as $\sigma^2 = 1,5,10$ and 25. Generalized ridge estimators are computed using different ridge parameters given in (6)–(10) and (28). The MSE of such ridge regression parameters are obtained using (5). This experiment is repeated 2000 times and obtains the AMSE.

We computed the AMSE ratios (AMSE ($\hat{\alpha}_{OLS}$)/ AMSE ($\hat{\alpha}_{GRR}$)) of OLS estimator over different estimators for various values of triplet (ρ, n, σ^2). These ratios are reported in Table 2 and here we noted that values of AMSE ratios only for k_i(HK), k_i(FG) and k_i(AD) are represented in Fig. 3

Same procedure for another choice of $p = 3$ and $\beta = (3,1,5)'$ is done and AMSE ratios are represented in Fig. 4.

From Table 2, Figs. 3 and 4, we conclude that proposed ridge parameter k_i(AD) and k_i(HK), the ridge parameter proposed by Hoerl and Kennard (1970a) both perform equivalently. Whereas, the performance of k_i (AD) is better than OLS, k_i(HMO), k_i(FG),k_i(TC) and k_i(F) for all combinations of correlation between predictors (ρ), sample size (n) and variance of the error term (σ^2) used in this simulation study.

5. Conclusion

In this article we have proposed a new method for estimating the ridge parameter in the presence of multicollinearity. The performance of the proposed ridge parameter is evaluated through the simulation study, for different combinations of correlation between predictors (ρ), the number of explanatory

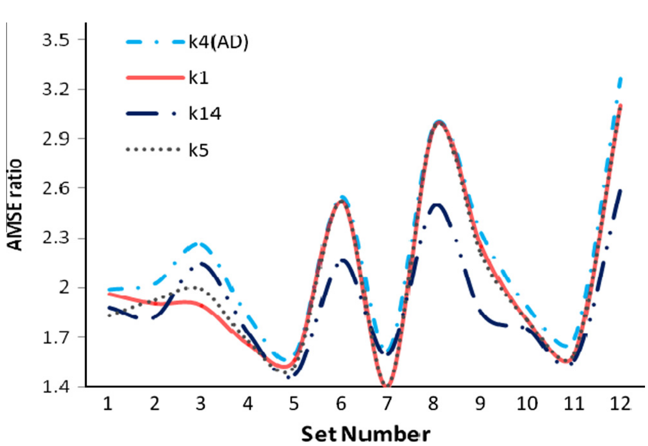

Figure 1 Ratio of AMSE of OLS over various ridge estimators for different 'k' ($p = 4$, $\beta = (2,3,5,1)'$ and $\rho = 0.9$).

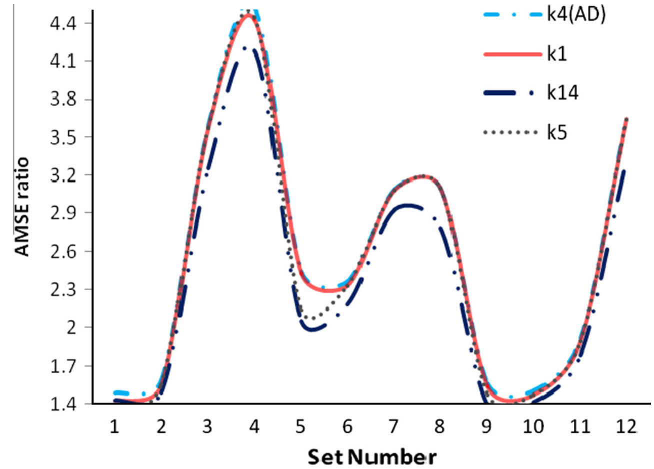

Figure 2 Ratio of AMSE of OLS over various ridge estimators for different 'k' ($p = 3$, $\beta = (3,1,5)'$ and $\rho = 0.9$).

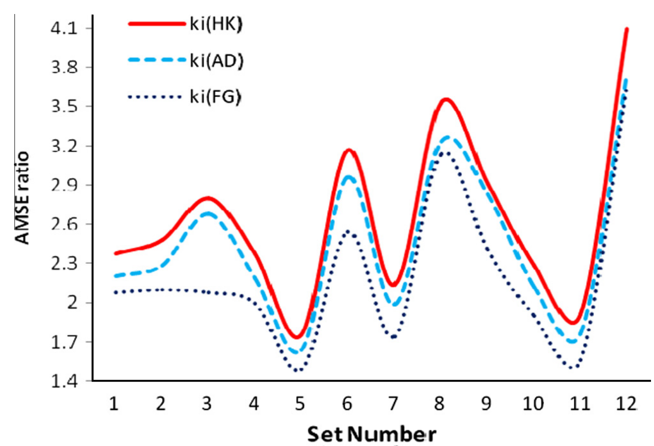

Figure 3 Ratio of AMSE of OLS over various ridge estimators for different 'k_i' ($p = 4$, $\beta = (2,3,5,1)'$ and $\rho = 0.9$).

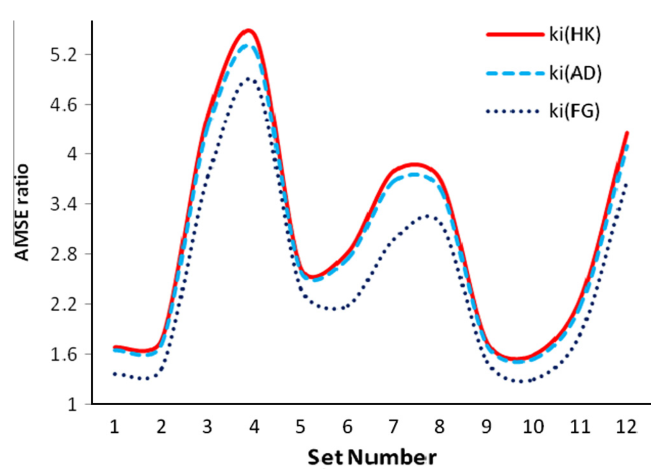

Figure 4 Ratio of AMSE of OLS over various ridge estimators for different 'k_i' ($p = 3$, $\beta = (3,1,5)'$ and $\rho = 0.9$).

Table 2 Ratio of AMSE of OLS over various ridge estimators for different 'k_i'.

n		20				50				100			
$\hat{\sigma}^2$		1	5	10	25	1	5	10	25	1	5	10	25
k_i	k_i(HK)	2.3771	2.4828	2.7997	2.3730	1.7549	3.1660	2.1404	3.5339	2.9098	2.2880	1.9120	4.1019
	k_i(HMO)	1.7969	2.0272	2.0637	1.9364	1.2744	2.4380	1.7089	3.0186	1.8384	1.7823	1.5184	3.5253
	k_i(FG)	2.0767	2.0997	2.0756	1.9922	1.4868	2.5389	1.7438	3.1379	2.4047	1.9007	1.5634	3.6587
	k_i(TC)	1.9325	1.9963	2.0460	1.9345	1.5955	2.1970	1.8157	2.3773	2.0415	1.8883	1.7068	2.5269
	$k_i(F)$	1.1077	1.1161	1.1112	1.1166	1.0371	1.0419	1.0420	1.0431	1.0178	1.0201	1.0203	1.0208
	k_i(AD)	2.2035	2.2866	2.6788	2.1872	1.6402	2.9614	1.9841	3.2328	2.8273	2.1192	1.7657	3.7458

variables (p), sample size (n) and variance of the error variable (σ^2). The evaluation of our estimator has been done by comparing the AMSE ratios of OLS estimator over the proposed estimator and the other estimators reviewed in this article. Finally, we found that the performance of the proposed estimator is satisfactory over the other estimators in the presence of multicollinearity.

Acknowledgments

The author is very grateful to the reviewers and the editor for so detailed comments and constructive suggestions which resulted in the present version. The present studies were supported in part by University Grants Commission, India, Project No. F.No.23-1743/10 (WRO), dated 22.09.10.

References

Al-Hassan, Y., 2010. Performance of new ridge regression estimators. Journal of the Association of Arab Universities for Basic and Applied Science 9, 23–26.

Alkhamisi, M.A., Shukur, G., 2007. A Monte Carlo study of recent ridge parameters. Communications in Statistics – Simulation and Computation 36 (3), 535–547.

Batah, F.S., Ramnathan, T., Gore, S.D., 2008. The efficiency of modified jackknife and ridge type regression estimators: a comparison. Surveys in Mathematics and its Applications 24 (2), 157–174.

Buonaccorsi, J.P., 1996. A modified estimating equation approach to correcting for measurement error in regression. Biometrika 83, 433–440.

Dorugade, A.V., Kashid, D.N., 2010. Alternative method for choosing ridge parameter for regression. International Journal of Applied Mathematical Sciences 4 (9), 447–456.

Firinguetti, L., 1999. A generalized ridge regression estimator and its finite sample properties. Communications in Statistics – Theory and Methods 28 (5), 1217–1229.

Golub, G.H., Heath, M., Wahba, G., 1979. Generalized cross-validation as a method for choosing a good ridge parameter. Technometrics 21 (2), 215–223.

Gunst, R.F., Mason, R.L., 1977. Biased estimation in regression: an evaluation using mean squared error. Journal of the American Statistical Association 72, 616–628.

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

Hocking, R.R., Speed, F.M., Lynn, M.J., 1976. A class of biased estimators in linear regression. Technometrics 18 (4), 425–437.

Hoerl, A.E., Kennard, R.W., 1970a. Ridge regression: biased estimation for nonorthogonal problems. Technometrics 12, 55–67.

Hoerl, A.E., Kennard, R.W., Baldwin, K.F., 1975. Ridge regression: some simulations. Communications in Statistics 4, 105–123.

Khalaf, G., Shukur, G., 2005. Choosing ridge parameter for regression problem. Communications in Statistics – Theory and Methods 34, 1177–1182.

Kibria, B.M., 2003. Performance of some new ridge regression estimators. Communications in Statistics – Simulation and Computation 32 (2), 419–435.

Lawless, J.F., Wang, P., 1976. A simulation study of ridge and other regression estimators. Communications in Statistics – Theory and Methods 14, 1589–1604.

Lawless, J.F., 1978. Ridge and related estimation procedure. Communications in Statistics – Theory and Methods 7 (2), 139–164.

Mardikyan, S., Cetin, E., 2008. Efficient choice of biasing constant for ridge regression. International Journal of Contemporary Mathematical Sciences 3 (11), 527–547.

Nomura, Masuo, 1988. On the almost unbiased ridge regression estimation. Communications in Statistics – Simulation and Computation 17 (3), 729–743.

McDonald, G.C., Galarneau, D.I., 1975. A Monte Carlo evaluation of some ridge-type estimators. Journal of the American Statistical Association 70 (350), 407–412.

Montogomery, D.C., Peck, E.A., Vining, G.G., 2006. Introduction to Linear Regression Analysis. John Wiley and Sons.

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

Norliza, A., Maizah, H.A., Robin, A., 2006. A comparative study on some methods for handling multicollinearity problems. Mathematika 22 (2), 109–119.

Pasha, G.R., Shah, M.A., 2004. Application of ridge regression to multicollinear data. Journal of Research (Science) 15 (1), 97–106.

Saleh, A.K., Kibria, B.M., 1993. Performances of some new preliminary test ridge regression estimators and their properties. Communications in Statistics – Theory and Methods 22, 2747–2764.

Stephen, G.W., Christopher, J.P., 2001. Generalized ridge regression and a generalization of the Cp statistic. Journal of Applied Statistics 28 (7), 911–922.

Troskie, C.G., Chalton, D.O., 1996. A Bayesian estimate for the constants in ridge regression. South African Statistical Journal 30, 119–137.