STATISTICS IN TRANSITION new series, June 2019
Vol. 20, No. 2, pp. 173–185, DOI 10.21307/stattrans-2019-021

EFFICIENT TWO-PARAMETER ESTIMATOR IN LINEAR REGRESSION MODEL

Ashok V. Dorugade[1]

ABSTRACT

In this article, two-parameter estimators in linear model with multicollinearity are considered. An alternative efficient two-parameter estimator is proposed and its properties are examined. Furthermore, this was compared with the ordinary least squares (OLS) estimator and ordinary ridge regression (ORR) estimators. Also, using the mean squares error criterion the proposed estimator performs more efficiently than OLS estimator, ORR estimator and other reviewed two-parameter estimators. A numerical example and simulation study are finally conducted to illustrate the superiority of the proposed estimator.

Key words: multicollinearity, ridge regression, two-parameter estimator, mean squared error.

1. Introduction

The ordinary least squares (OLS) method is one of the most important ways for estimating the parameters of the general linear model. Because of its simplicity and rationality, the results are obtained when specific assumptions are achieved. But if these assumptions are violated, OLS method does not assure the desirable results. Multicollinearity occurs when two or more than two explanatory variables are correlated with each other. To solve this problem, various biased estimators were put forward in the literature. The ordinary ridge regression (ORR) proposed by Hoerl and Kennard (1970a) is the most popular biased estimator. However, ORR estimator has some disadvantages; mainly it is a nonlinear function of the ridge parameter k. This leads to complicated equations when selecting k. To solve such difficulty, Liu (1993) then proposed the estimator called Liu estimator (LE). As seen, LE is a linear function of the ridge parameter d and thus it is more convenient to choose d than k. Liu (2003) suggested two Liu-type estimators and proved that these estimators have some superior properties over RR estimator under the mean squared error (MSE) criterion. However, it is difficult to determine which is better between them.

Some other popular numerical techniques to deal with multicollinearity are the ridge regression due to Singh and Chaubey (1987), Liu (1993), Akdeniz and

[1] Y C Mahavidyalaya Halkarni, Tal-Chandgad, Kolhapur, Maharashtra India – 416552, India. E-mail: adorugade@rediffmail.com.

Kaciranlar (1995), Crouse *et al.,* (1995), Kaciranlar *et al.*, (1999), Ozkale and Kaciranlar (2007), Yang and Chang (2010), Wu and Yang (2011), Dorugade (2014) and others.

In this paper, a new method for estimating the parameters in linear regression model with multicollinearity problem is proposed. The rest of this paper is organized as follows. The model and some well-known estimators are reviewed in Section 2. The efficient two parameter estimator is introduced in Section 3. Performances of the proposed estimator with respect to the scalar MSE criterion are discussed in Section 4. In Section 5, the methods of choosing the parameters were discussed. A simulation study to justify the superiority of the suggested estimator is given in Section 6. Some concluding remarks are given in Section 7.

2. Model and estimators

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 the vector of errors 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, it was assumed that the matrix X and the 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.

Let $\wedge$ and T be the matrices of eigen values and eigen vectors of $X'X$, respectively, satisfying $T'X'XT$ = $\wedge$ = diagonal ($\lambda_1, \lambda_2, ..., \lambda_p$), with λ_i 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 , it implies that $Z'Z = \wedge$, and α = $T'\beta$ (see Montgomery *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, LS estimator of β is given by

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

However, it is well-known that OLS estimator performs poorly when multicollinearity exists. In order to control the instability in least squares estimates, Hoerl (1962), Hoerl and Kennard (1968) and then Hoerl and Kennard (1970b) suggested an alternative estimate of the regression coefficients namely ridge regression as obtained by adding a positive constant (or ridge parameter) k to the diagonal elements of the ordinary least square estimator. It is given as:

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

Therefore, ORR estimator of β is given by

$$\hat{\beta}_{ORR} = T\hat{\alpha}_{ORR}$$

The literature has shown that some ridge estimators are based on a single ridge parameter while some are based on two ridge parameters. Some of the well-known methods used for estimation are listed below.

The Jackknifed ridge regression estimator introduced by Singh and Chaubey (1987) is defined by

$$\hat{\alpha}_{JRR} = \left[I - k^2(\wedge + kI)^{-2}\right]\hat{\alpha}_{OLS} \quad (5)$$

Liu (1993) introduced a biased estimator, which is defined by

$$\hat{\alpha}_{Liu} = (\wedge + I)^{-1}(\wedge + dI)\hat{\alpha}_{OLS} \quad (6)$$

The almost unbiased Liu estimator introduced by Akdeniz and Kaciranlar (1995) is defined by

$$\hat{\alpha}_{AUL} = \left[I - (\wedge + I)^{-2} + (1-d)^2\right]\wedge^{-1} X'Y . \quad (7)$$

Crouse et al., (1995) defined unbiased ridge estimator given by

$$\hat{\alpha}_{URR} = (\wedge + kI)^{-1}(Z'Y + kJ) \text{ where } J = \sum_{i=1}^{p}\hat{\alpha}_i / p . \quad (8)$$

Ozkale and Kaciranlar (2007) introduced a two-parameter estimator, which is defined by

$$\hat{\alpha}_{TP} = (\wedge + kI)^{-1}(\wedge + kdI)\hat{\alpha}_{OLS} . \quad (9)$$

The ridge parameter $k = p\hat{\sigma}^2 \Big/ \sum_{i=1}^{p}\hat{\alpha}_i^2$ given by Hoerl et al. (1975) performs fairly well and the well-known estimate of 'd'proposed by Liu (1993) is given as:

$$d = \sum_{i=1}^{p}(\hat{\alpha}_i^2 - \hat{\sigma}^2)/(\lambda_i + 1)^2 \Big/ \sum_{i=1}^{p}(\hat{\sigma}^2 + \lambda_i\hat{\alpha}_i^2)/(\lambda_i + 1)^2\lambda_i .$$

The above calculated values of k and d are used to determine estimators given in (5) to (9), where, $\hat{\alpha}_i$ is the i[th] element of $\hat{\alpha}_{OLS}$, $i = 1,2,..., p$ and $\hat{\sigma}^2$ is the OLS estimator of σ^2 i.e. $\hat{\sigma}^2 = (Y'Y - \hat{\alpha}'Z'Y)/(n - p - 1)$.

In the context of two-parameter estimator, Yang and Chang (2010), Wu and Yang (2011) have recently suggested two-parameter estimator's alternative to LS estimator in the presence of multicollinearity. These estimators are given as:

Yang and Chang (2010) suggested new two-parameter (NTP) estimator, given by

$$\hat{\alpha}_{NTP} = (\wedge + I)^{-1}(\wedge + dI)(\wedge + kI)^{-1} Z' Y \tag{10}$$

where, $k = p\hat{\sigma}^2 \Big/ \sum_{i=1}^{p} \hat{\alpha}_i^2$ and

$$d = \frac{\sum_{i=1}^{p} \left\{ \left[(k+1)\lambda_i + k \right] \lambda_i \hat{\alpha}_i^2 - \lambda_i{}^2 \hat{\sigma}^2 \right\} \Big/ \left[(\lambda_i + 1)^2 (\lambda_i + k)^2 \right]}{\sum_{i=1}^{p} (\hat{\sigma}^2 + \lambda_i \hat{\alpha}_i^2) \Big/ \left[(\lambda_i + 1)^2 (\lambda_i + k)^2 \right]} .$$

Also, MSE of $\hat{\alpha}_{NTP}$ is given as:

$$MSE(\hat{\alpha}_{NTP}) = \sigma^2 \sum_{i=1}^{p} \left[\frac{\lambda_i (\lambda_i + d)^2}{(\lambda_i + 1)^2 (\lambda_i + k)^2} \right] + \sum_{i=1}^{p} \left\{ \frac{\left[(k + 1 - d)\lambda_i + k \right]^2}{(\lambda_i + 1)^2 (\lambda_i + k)^2} \right\} \alpha_i^2 . \tag{11}$$

Wu and Yang (2011) introduced always unbiased two-parameter (AUTP) estimator, which is defined by

$$\hat{\alpha}_{AUTP} = \hat{\alpha}_{TP} + k(1-d)(\wedge + kI)^{-1} \hat{\alpha}_{TP} \tag{12}$$

where, $d < 1 - \min\left(\hat{\sigma} \Big/ \sqrt{\lambda_i \hat{\alpha}_i^2 + \hat{\sigma}^2} \right)$ and $k = \lambda_i \hat{\sigma} \Big/ \left[(1-d)\sqrt{\lambda_i \hat{\alpha}_i^2 + \hat{\sigma}^2} - \hat{\sigma} \right]$.

Also, MSE of $\hat{\alpha}_{AUTP}$ is given as:

$$MSE(\hat{\alpha}_{AUTP}) = \sigma^2 \sum_{i=1}^{p} \left\{ \frac{\left[\lambda_i (\lambda_i + 2k) + d\, k^2 (2-d) \right]^2}{\lambda_i (\lambda_i + k)^4} \right\} + \sum_{i=1}^{p} \left[\frac{k^4 (1-d)^4}{(\lambda_i + k)^4} \right] \alpha_i^2 . \tag{13}$$

Recently, Dorugade (2014) introduced a modified two-parameter (MTP) estimator, which is defined by

$$\hat{\alpha}_{MTP} = \left[I + k(1-d)(\Lambda + kdI)^{-1} \right] \left[I - kd(\Lambda + kdI)^{-1} \right] \hat{\alpha}_{OLS} \tag{14}$$

Where, $k = p\hat{\sigma}^2 \Big/ \sum_{i=1}^{p} \hat{\alpha}_i^2$ and $\hat{d} = \sum_{i=1}^{p} \left[\frac{(\lambda_i + k)(\hat{\sigma}^2 + \lambda_i \hat{\alpha}_i^2) - \lambda_i}{k \hat{\alpha}_i^2} \right]$.

Also, MSE of $\hat{\alpha}_{MTP}$ is given as:

$$MSE(\hat{\alpha}_{MTP}) = \sigma^2 \sum_{i=1}^{p} \left[\frac{\lambda_i (\lambda_i + k)^2}{(\lambda_i + kd)^4} \right] + \sum_{i=1}^{p} \left\{ \frac{k^2 \left[(1-2d)\lambda_i - kd^2 \right]^2}{(\lambda_i + kd)^4} \right\} \alpha_i^2 . \tag{15}$$

All the above methods of estimating α are used in Section 6.

3. Proposed estimator

Hoerl and Kennard (1962) first suggested that to control the inflation and general instability associated with the least squares estimates. The relationship of a ridge estimate to an ordinary estimate is given by the alternative form. It is observed that OLS is unbiased but has inflated variances under multicollinearity. Due to the complicated nature of the ridge parameter k, in the ridge regression method, Liu (2003) proposed a two-parameter estimator. In this article we introduce the efficient two-parameter estimator, which can be computed in two steps. Initially, following the similar method proposed by Liu (1993), Kaciranlar *et al.,* (1999) and Yang and Chang (2010), we introduce the two-parameter estimator as:

$$\hat{\alpha}^* = \left[\wedge + k(1-d)I\right]^{-1} Z'Y . \tag{16}$$

Hoerl and Kennard (1970a) pointed that ORR avoids inflating the variances at the cost of bias by pre-multiplying $\hat{\alpha}_{OLS}$ with the matrix $[I - k(\Lambda + kI)^{-1}]$ to reduce the inflated variances in OLS. The proposed estimator is based on this same logic by pre-multiplying $\hat{\alpha}^*$ with the matrix $[I - k(\Lambda + kI)^{-1}]$. This is defined in equation (17) as :

$$\hat{\alpha}_{ETP} = [I - k(\Lambda + kI)^{-1}]\hat{\alpha}^*, \tag{17}$$

or

$$\hat{\alpha}_{ETP} = \Lambda^2[(\Lambda + kI)^{-1}(\Lambda + k(1-d)I)^{-1}]\,\hat{\alpha}_{OLS} .$$

Equation (17) is termed as Efficient Two-Parameter (ETP) estimator of α.
Thus, the coordinate wise estimators can be written as

$$\hat{\alpha}_{iETP} = \left[\frac{\lambda_i^2}{(\lambda_i + k)(\lambda_i + k(1-d))}\right]\hat{\alpha}_i \quad i = 1,2,\dots, p \tag{18}$$

where $\hat{\alpha}_i$ are the individual components of $\hat{\alpha}_{OLS}$.

We can see that it is a general estimator which includes the OLS and ORR estimators as special cases:

at $(k = 0,\ d)$ $\quad \hat{\alpha}_{ETP} = \hat{\alpha}_{OLS}$ the OLS estimator,

at $(k,\ d = 1)$ $\quad \hat{\alpha}_{ETP} = [I - k(\wedge + kI)^{-1}]\hat{\alpha}_{OLS}$ the ORR estimator,

at $(k = d = 1)$ $\quad \hat{\alpha}_{ETP} = [I - (\wedge + I)^{-1}]\hat{\alpha}_{OLS}$,

at $(k,\ d = 0)$ $\quad \hat{\alpha}_{ETP} = [I - k(\wedge + kI)^{-1}]\hat{\alpha}_{ORR}$.

3.1. Bias, Variance and MSE of ETP estimator

It is clear that $\hat{\alpha}_{ETP}$ is biased estimator, the bias of the ETP estimator is given by:

$$Bias(\hat{\alpha}_{ETP}) = E[\hat{\alpha}_{ETP}] - \alpha$$

$$= [\Lambda^2(\Lambda + kI)^{-1}(\Lambda + k(1-d)I)^{-1} - I]\,\alpha$$

$$= \sum_{i=1}^{p} \left\{ \frac{-\left[k(2-d)\lambda_i + k^2(1-d)\right]}{(\lambda_i + k)(\lambda_i + k(1-d))} \right\} \hat{\alpha}_i \qquad (19)$$

$$V(\hat{\alpha}_{ETP}) = \sigma^2 V \wedge^{-1} V' \text{ where } V = \Lambda^2[(\Lambda + kI)^{-1}(\Lambda + k(1-d)I)^{-1}]$$

$$= \hat{\sigma}^2 \sum_{i=1}^{p} \left[\frac{\lambda_i^3}{(\lambda_i + k)^2(\lambda_i + k(1-d))^2} \right]. \qquad (20)$$

The MSE of ETP estimator is

$$MSE(\hat{\alpha}_{ETP}) = V(\hat{\alpha}_{ETP}) + [Bias(\hat{\alpha}_{ETP})]\,[Bias(\hat{\alpha}_{ETP})]'$$

$$MSE(\hat{\alpha}_{ETP}) = \hat{\sigma}^2 \sum_{i=1}^{p} \left[\frac{\lambda_i^3}{(\lambda_i + k)^2(\lambda_i + k(1-d))^2} \right] + \sum_{i=1}^{p} \left\{ \frac{\left[k(2-d)\lambda_i + k^2(1-d)\right]^2}{(\lambda_i + k)^2(\lambda_i + k(1-d))^2} \right\} \hat{\alpha}_i^2 . \qquad (21)$$

Setting $k = 0$ in (21), we obtain

$$MSE(\hat{\alpha}_{OLS}) = \hat{\sigma}^2 \sum_{i=1}^{p} \frac{1}{\lambda_i}. \qquad (22)$$

Also, setting $d = 1$ in (21), we obtain

$$MSE(\hat{\alpha}_{ORR}) = \hat{\sigma}^2 \sum_{i=1}^{p} \frac{\lambda_i}{(\lambda_i + k)^2} + k^2 \sum_{i=1}^{p} \frac{\hat{\alpha}_i^2}{(\lambda_i + k)^2}. \qquad (23)$$

4. Performance of the proposed estimator

This section compares the performance of $\hat{\alpha}_{ETP}$ with $\hat{\alpha}_{OLS}$ and $\hat{\alpha}_{ORR}$ using MSE criteria.

4.1. Comparison between the $\hat{\alpha}_{ETP}$ and $\hat{\alpha}_{OLS}$ using MSE criterion

The difference between $\hat{\alpha}_{OLS}$ and $\hat{\alpha}_{ETP}$ in the MSE sense is as follows:

$MSE(\hat{\alpha}_{OLS})$ - $MSE(\hat{\alpha}_{ETP})$

$$=\hat{\sigma}^2\sum_{i=1}^{p}\left[\frac{1}{\lambda_i}-\frac{\lambda_i^{\,3}}{(\lambda_i+k)^2(\lambda_i+k(1-d))^2}\right]-\sum_{i=1}^{p}\left\{\frac{\left[k(2-d)\lambda_i+k^2(1-d)\right]^2}{(\lambda_i+k)^2(\lambda_i+k(1-d))^2}\right\}\hat{\alpha}_i^2$$

$$=\sum_{i=1}^{p}\left\{\frac{\hat{\sigma}^2\left[(\lambda_i+k)^2(\lambda_i+k(1-d))^2-\lambda_i^{\,4}\right]-\lambda_i\hat{\alpha}_i^2\left[k(2-d)\lambda_i+k^2(1-d)\right]^2}{\lambda_i(\lambda_i+k)^2(\lambda_i+k(1-d))^2}\right\}.$$

From the above equation, it can be shown that the difference $MSE(\hat{\alpha}_{OLS})$-$MSE(\hat{\alpha}_{ETP})$ will be positive if and only if

$$\hat{\sigma}^2\left[(\lambda_i+k)^2(\lambda_i+k(1-d))^2-\lambda_i^{\,4}\right]\geq\lambda_i\hat{\alpha}_i^2\left[k(2-d)\lambda_i+k^2(1-d)\right]^2$$

Thus, $MSE(\hat{\alpha}_{OLS})\geq MSE(\hat{\alpha}_{ETP})$

if and only if

$$\hat{\sigma}^2\left[(\lambda_i+k)^2(\lambda_i+k(1-d))^2-\lambda_i^{\,4}\right]\geq\lambda_i\hat{\alpha}_i^2\left[k(2-d)\lambda_i+k^2(1-d)\right]^2. \qquad (24)$$

4.2. Comparison between the $\hat{\alpha}_{ETP}$ and $\hat{\alpha}_{ORR}$

The difference between $\hat{\alpha}_{ORR}$ and $\hat{\alpha}_{ETP}$ in the MSE sense is as follows:

$$MSE(\hat{\alpha}_{ORR})\text{-}MSE(\hat{\alpha}_{ETP})=\hat{\sigma}^2\sum_{i=1}^{p}\left[\frac{\lambda_i}{(\lambda_i+k)^2}-\frac{\lambda_i^{\,3}}{(\lambda_i+k)^2(\lambda_i+k(1-d))^2}\right]$$

$$+\sum_{i=1}^{p}\left\{\frac{k^2}{(\lambda_i+k)^2}-\frac{\left[k(2-d)\lambda_i+k^2(1-d)\right]^2}{(\lambda_i+k)^2(\lambda_i+k(1-d))^2}\right\}\hat{\alpha}_i^2$$

$$=\sum_{i=1}^{p}\left\{\frac{\hat{\sigma}^2k(1-d)\lambda_i\left[2\lambda_i+k(1-d)\right]+\hat{\alpha}_i^2\left\{k^2(\lambda_i+k(1-d))^2-\left[k(2-d)\lambda_i+k^2(1-d)\right]^2\right\}}{(\lambda_i+k)^2(\lambda_i+k(1-d))^2}\right\}$$

From above equation, it can be shown that the difference $MSE(\hat{\alpha}_{ORR})$-$MSE(\hat{\alpha}_{ETP})$ will be positive if and only if $k^2(\lambda_i+k(1-d))^2\geq\left[k(2-d)\lambda_i+k^2(1-d)\right]^2$

Thus, $MSE(\hat{\alpha}_{ORR})\geq MSE(\hat{\alpha}_{ETP})$

if and only if

$$k^2(\lambda_i+k(1-d))^2\geq\left[k(2-d)\lambda_i+k^2(1-d)\right]^2. \qquad (25)$$

5. Determination of ridge parameter k and d

A very important issue in the study of the ridge and Liu regression is how to find appropriate ridge and Liu parameters, k and d respectively. These parameters may either be nonstochastic or may depend on the observed data. The choice of values for these ridge parameters has been one of the most difficult problems confronting the study of the generalized ridge regression.

In order to determine and evaluate the performance of our proposed estimator $\hat{\alpha}_{ETP}$ as compare to OLS estimator and others, we will find the optimal values of ridge parameters k and d. Let $\hat{k}$ is the optimal value of the k determined by well-known method of determining the ridge parameter, the optimal value of the d can be considered to be this d that minimizes $MSE(\hat{\alpha}_{ETP})$.

Let $g(k,d) = MSE(\hat{\alpha}_{ETP})$

$$= \sigma^2 \sum_{i=1}^{p} \left[\frac{\lambda_i^3}{(\lambda_i + k)^2 (\lambda_i + k(1-d))^2} \right] + \sum_{i=1}^{p} \left\{ \frac{\left[k(2-d)\lambda_i + k^2(1-d) \right]^2}{(\lambda_i + k)^2 (\lambda_i + k(1-d))^2} \right\} \alpha_i^2 .$$

Then, by differentiating $g(k,d)$ w.r.t. d and equating to 0, we have

$$d = \sum_{i=1}^{p} \left\{ \frac{k\,\alpha_i^2 \lambda_i^2 \left[(\lambda_i + k)\, k + \lambda_i \right] - \sigma^2 \lambda_i^3}{k^2 \alpha_i^2 \lambda_i^2 (\lambda_i + k)} \right\}, \tag{26}$$

d is the function of k depends on the σ^2 and α_i. For practical purposes, they are replaced by their unbiased estimator $\hat{\sigma}^2$ and $\hat{\alpha}_i$. Hence,

$$\hat{d} = \sum_{i=1}^{p} \left\{ \frac{k\,\hat{\alpha}_i^2 \lambda_i^2 \left[(\lambda_i + k)\, k + \lambda_i \right] - \hat{\sigma}^2 \lambda_i^3}{k^2 \hat{\alpha}_i^2 \lambda_i^2 (\lambda_i + k)} \right\}. \tag{27}$$

6. Comparative study

6.1. A simulation study

The performance of the proposed estimator and the existing estimators is examined via a simulation study. The simulation is carried out under different degrees of multicollinearity. The average MSE (AMSE) ratios of the $\hat{\alpha}_{ETP}$ and other ridge estimators over OLS estimator are evaluated. The true model is considered as $Y = X\beta + \varepsilon$. Here ε follows a normal distribution $N(0, \sigma^2 I_n)$ and following McDonald and Galerneau (1975) the explanatory variables are generated by

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

where u_{ij} 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. 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 that

$\beta = (2, 1, 6, 2)'$, $\beta = (1, 1, 3)'$ for p = 4 and 3, respectively. The sizes of samples are 20, 50 and 100. The variance of the error terms is taken as σ^2 = 1, 5 and 10.

The well-known ridge parameter $k = p\hat{\sigma}^2 \Big/ \sum_{i=1}^{p} \hat{\alpha}_i^2$ suggested by Hoerl *et al.*, (1975) was used. d is determined as defined in equation (27). Efficient two-parameter estimator and estimators given in (3) to (5), (7) to (10), (12) and (14) are computed. The experiment is repeated 1000 times and the average MSE (AMSE) of estimators is obtained using the following expression:

$$AMSE(\hat{\alpha}) = \frac{1}{1000} \sum_{i=1}^{p} \sum_{j=1}^{1000} (\hat{\alpha}_{ij} - \alpha_i)^2 \qquad (28)$$

where $\hat{\alpha}_{ij}$ denote the estimator of the i^th^ parameter in the j^th^ replication and α_i, $i = 1,2,..., p$ are the true parameter values.

Firstly, we computed the AMSE ratios ($AMSE(\hat{\alpha}_{OLS}) / AMSE(\hat{\alpha})$) of OLS estimator over different estimators for various values of triplet (ρ, n, σ^2) and reported in Tables 1-4. We consider the method that leads to the maximum AMSE ratio to the best from the MSE point of view.

From Tables 1-4, we observe that the performance of our proposed efficient two-parameter estimator ($\hat{\alpha}_{ETP}$) is better than $\hat{\alpha}_{OLS}$. Also, $\hat{\alpha}_{ETP}$ is more efficient in terms of MSE than other biased estimators $\hat{\alpha}_{JRR}$, $\hat{\alpha}_{AUL}$, $\hat{\alpha}_{URR}$, $\hat{\alpha}_{TP}$, $\hat{\alpha}_{NTP}$, $\hat{\alpha}_{AUTP}$ and $\hat{\alpha}_{MTP}$ including $\hat{\alpha}_{ORR}$ for various values of triplet (ρ, n, σ^2). The results agree with our theoretical findings in Section 4.

Table 1. Ratio of AMSE of OLS over various ridge estimators
(p = 4, $\beta = (2, 1, 6, 2)'$ and ρ = 0.95)

Estimator	**n = 20**			**50**			**100**		
	$\hat{\sigma}^2 = \mathbf{1}$	**5**	**10**	**1**	**5**	**10**	**1**	**5**	**10**
$\hat{\alpha}_{ORR}$	1.0209	1.3800	1.4746	1.5603	1.7090	2.1274	2.1833	2.2644	2.3872
$\hat{\alpha}_{JRR}$	1.0019	1.3525	1.3624	1.5186	1.5745	1.8444	1.7952	1.4780	1.4779
$\hat{\alpha}_{AUL}$	1.0326	1.2966	1.0040	1.0793	1.2278	1.0487	1.0013	1.0125	1.0219
$\hat{\alpha}_{URR}$	1.4067	1.3184	1.3363	1.5621	1.7799	1.6437	1.8530	2.4037	3.2057
$\hat{\alpha}_{TP}$	0.9909	0.8698	0.9730	0.8516	0.7984	0.8615	0.9675	0.9390	1.0504
$\hat{\alpha}_{NTP}$	0.7030	0.3278	1.2536	1.5321	0.8943	1.9981	2.4686	2.0685	2.3089
$\hat{\alpha}_{AUTP}$	1.0004	1.0695	1.0017	1.0697	1.1131	1.0328	1.0014	1.0032	1.0017
$\hat{\alpha}_{MTP}$	1.0210	1.3800	1.4748	1.5600	1.7094	2.1270	2.1836	2.2648	2.3870
$\hat{\alpha}_{ETP}$	1.4388	1.4078	1.5967	1.6028	1.8579	2.4492	2.6313	3.0779	3.3611

Table 2. Ratio of AMSE of OLS over various ridge estimators ($p = 4$, $\beta = (2, 1, 6, 2)'$ and $\rho = 0.99$)

Estimator	n = 20			50			100		
	$\hat{\sigma}^2 = 1$	5	10	1	5	10	1	5	10
$\hat{\alpha}_{ORR}$	1.2365	1.4166	1.9591	2.2037	2.7241	2.8145	2.2735	2.7244	2.9822
$\hat{\alpha}_{JRR}$	1.1365	1.1309	1.4754	1.2585	2.4739	1.7537	2.2602	2.3833	1.6211
$\hat{\alpha}_{AUL}$	1.0873	1.0061	1.052	1.068	1.0458	1.02	1.0022	1.0142	1.1438
$\hat{\alpha}_{URR}$	1.2691	1.4654	0.9704	1.7314	2.281	2.8455	1.5181	1.742	2.257
$\hat{\alpha}_{TP}$	1.0435	0.9784	0.8843	0.9211	1.2129	0.8857	1.0332	1.1264	0.8024
$\hat{\alpha}_{NTP}$	1.1173	1.1683	1.5396	0.516	2.587	2.9986	2.5124	2.1497	2.0449
$\hat{\alpha}_{AUTP}$	1.0054	1.0006	1.0176	1.0036	1.0409	1.0119	1.0016	1.0153	1.0336
$\hat{\alpha}_{MTP}$	1.2300	1.4168	1.8960	2.2039	2.7245	2.8150	2.2740	2.7250	2.9820
$\hat{\alpha}_{ETP}$	1.34	1.5982	2.5048	2.6136	3.0088	4.0831	2.2873	3.1294	4.2129

Table 3. Ratio of AMSE of OLS over various ridge estimators ($p = 3$, $\beta = (1, 1, 3)'$ and $\rho = 0.95$)

Estimator	n = 20			50			100		
	$\hat{\sigma}^2 = 1$	5	10	1	5	10	1	5	10
$\hat{\alpha}_{ORR}$	1.5559	2.0183	2.7696	1.8432	1.923	2.3064	2.0479	2.0930	3.0024
$\hat{\alpha}_{JRR}$	1.5030	1.3665	1.4734	1.7442	1.5766	1.7138	1.3825	1.3013	1.9361
$\hat{\alpha}_{AUL}$	1.0061	1.0577	1.0181	1.0044	1.0591	1.0157	1.0690	1.0469	1.0214
$\hat{\alpha}_{URR}$	1.1864	1.7946	2.5463	1.8317	1.8465	0.7616	1.7940	1.8067	2.5083
$\hat{\alpha}_{TP}$	0.9606	0.8749	0.9631	1.0025	0.8328	0.8223	0.8684	0.8958	0.9171
$\hat{\alpha}_{NTP}$	1.2280	1.7430	2.1870	1.8514	1.5227	1.6014	1.6203	1.8242	2.7112
$\hat{\alpha}_{AUTP}$	1.0023	1.0176	1.0044	1.003	1.0301	1.027	1.0118	1.0090	1.0063
$\hat{\alpha}_{MTP}$	1.5550	2.0189	2.7755	1.8440	1.933	2.3164	2.0586	2.0980	3.0124
$\hat{\alpha}_{ETP}$	1.6037	2.3405	3.5764	1.9354	2.2614	2.8938	2.3319	2.4589	4.0607

Table 4. Ratio of AMSE of OLS over various ridge estimators (p = 3, $\beta = (1, 1, 3)'$ and ρ = 0.99)

Estimator	n = 20			50			100		
	$\hat{\sigma}^2 = 1$	5	10	1	5	10	1	5	10
$\hat{\alpha}_{ORR}$	1.5622	2.8334	3.0927	1.9421	2.2077	2.4474	2.5685	2.6181	3.5616
$\hat{\alpha}_{JRR}$	1.2507	1.8883	2.2305	1.2062	1.7607	1.4847	1.7945	1.4024	2.1656
$\hat{\alpha}_{AUL}$	1.0227	1.0308	1.0204	1.0817	1.0179	1.1277	1.0273	1.1743	1.0506
$\hat{\alpha}_{URR}$	1.4735	1.8497	2.1001	1.5143	1.4785	2.1487	2.3512	2.0673	2.2603
$\hat{\alpha}_{TP}$	0.8635	0.8921	0.8776	0.8935	0.915	0.8246	0.9067	0.8155	0.8338
$\hat{\alpha}_{NTP}$	1.4160	2.0758	2.5563	1.3947	1.9748	1.7623	0.5709	1.6083	3.438
$\hat{\alpha}_{AUTP}$	1.0032	1.0115	1.0058	1.0076	1.0096	1.0288	1.0111	1.0246	1.0251
$\hat{\alpha}_{MTP}$	1.5620	2.8634	3.2927	2.1011	2.3077	2.8474	3.1685	2.9182	3.7521
$\hat{\alpha}_{ETP}$	1.8216	3.9784	3.6171	2.1426	2.7185	3.2002	3.3286	3.1803	4.4847

6.2. Numerical example

To validate the theoretical results, the numerical example used by Gruber (1988) was adopted. It was established that this data set suffers 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 four other independent variables X_1, X_2, X_3, and X_4.

The estimated MSE values for $\hat{\alpha}_{OLS}, \hat{\alpha}_{ORR}, \hat{\alpha}_{NTP}, \hat{\alpha}_{AUTP}, \hat{\alpha}_{MTP}$ and $\hat{\alpha}_{ETP}$ estimators, are obtained and reported in Table 5.

Table 5. Values of MSE

Estimator	$\hat{\alpha}_{OLS}$	$\hat{\alpha}_{ORR}$	$\hat{\alpha}_{NTP}$	$\hat{\alpha}_{AUTP}$	$\hat{\alpha}_{MTP}$	$\hat{\alpha}_{ETP}$
MSE	0.2833	0.1256	0.1255	0.2832	0.1257	0.1251

From Table 5, it was observed that the estimated MSE value of the efficient two-parameter estimator ($\hat{\alpha}_{ETP}$**)** is always smaller than those of the $\hat{\alpha}_{OLS}$, $\hat{\alpha}_{ORR}$, $\hat{\alpha}_{NTP}$, $\hat{\alpha}_{AUTP}$ and $\hat{\alpha}_{MTP}$ estimators. The results agree with the theoretical findings in Section 4. Finally, $\hat{\alpha}_{ETP}$ is meaningful in practice.

7. Conclusion

A biased efficient two-parameter estimator has been proposed for estimating the parameter of the linear regression model with multicollinearity. The proposed estimator is examined against OLS and ORR estimator in terms of scalar MSE criterion. Finally, from the simulation study and numerical example, the performance of the proposed estimator is satisfactory in the presence of multicollinearity over other estimators reviewed in this article.

Acknowledgements

The author would like to thank the referee for his valuable comments, which resulted in the present version of this article. In the part Research Award Scheme, the present studies were supported by UGC, India Project No. F. 30-1/2014/RA-2014-16-GE-MAH-5958 (SA-II).

REFERENCE

AKDENIZ, F., KACIRANLAR, S., (1995). On the almost unbiased generalized Liu estimator and unbiased estimation of the bias and MSE, Commun. Statist. Theor. Meth., 24, pp. 1789–1797.

CROUSE, R. H., JIN, C., HANUMARA, R. C., (1995). Unbiased ridge estimation with prior information and ridge trace, Commun. Statist. Theor. Meth., 24, pp. 2341–2354.

DORUGADE, A. V., (2014). Modified Two Parameter Estimator in Linear regression, Statistics in Transition - new series, 15 (1), pp. 23–36.

GRUBER, M. H. J., (1998). Improving efficiency by shrinkage the James–Stein and Ridge regression estimators. Marcell Dekker, NewYork.

HOERL, A. E., (1962). Application of ridge analysis to regression problems, Chemical Engineering Progress., 68, pp. 54–59.

HOERL, A. E., KENNARD, R. W., (1968). On regression analysis and biased estimation, Technometrics, 10, pp. 422–423.

HOERL, A. E., KENNARD, R. W., (1970a). Ridge regression: Biased estimation for non orthogonal problems, Technometrics, 12, pp. 55–67.

HOERL, A. E., KENNARD, R. W., (1970b). Ridge regression: Applications to Nonorthogonal problems, Technometrics, 12, pp. 69–82.

HOERL, A. E., KENNARD, R. W., BALDWIN, K. F., (1975). Ridge regression: Some Simulations, Commun. Statist., 4, pp. 105–123.

KACIRANLAR, S., SAKALLIOGLU, S., AKDENIZ, F., STYAN, G. P. H., WERNER, H. J., (1999). A new biased estimator in linear regression and a detailed analysis of the widely-analysed dataset on Portland Cement, Sankhya Ind. J. Statist., 61, pp. 443–459.

LIU, K., (1993). A new class of biased estimate in linear regression, Commun. Statist. Theor. Meth., 22, pp. 393–402.

LIU, K., (2003). Using Liu-type estimator to combat Collinearity. Commun, Statist. Theor. Meth., 32, pp. 1009–1020.

MCDONALD G. C., GALARNEAU, D. I., (1975). A Monte Carlo evaluation of some ridge-type estimators, J Am Stat Assoc., 20, pp. 407–416.

MONTGOMERY, D. C., PECK, E. A., VINING, G. G., (2006). Introduction to linear regression analysis. John Wiley and Sons, New York.

OZKALE, M. R., KACIRANLAR , S., (2007). The restricted and unrestricted two-parameter estimators, Commun. Statist. Theor. Meth., 36, pp. 2707–2725.

SINGH, B., CHAUBEY, Y. P., (1987). On some improved ridge estimators, Stat Papers, 28, pp. 53–67.

YANG, H., CHANG, X., (2010). A New Two-Parameter Estimator in Linear Regression, Commun. Statist. Theor. Meth., 39, pp. 923–934.

WU, J., YANG, H., (2011). Efficiency of an almost unbiased two-parameter estimator in linear regression model, Statistics., 47 (3), pp. 535–545.