AL-Bahir
Journal for Engineering and Pure Sciences
BJEPS

Volume 2 | Issue 1 Article 1

Biasing Estimator to Mitigate Multicollinearity in Linear Regression Model

Abdulrasheed Bello Badawaire
Department of Mathematics and Statistics, Federal University Wukari, Wukari, Nigeria

Issam Dawoud
Department of Mathematics, Al-Aqsa University, Gaza, Palestine

Adewale Folaranmi Lukman
Department of Mathematics, University of Medical Scieces, Ondo, Nigeria

Victoria Laoye
Department of Statistics, University of Ibadan, Ibadan, Nigeria

Arowolo Olatunji
Department of Mathematical Sciences, Lagos State University of Science and Technology, Lagos, Nigeria.

ORIGINAL STUDY

Biasing Estimator to Mitigate Multicollinearity in Linear Regression Model

Abdulrasheed Bello Badawaire [a,*], Issam Dawoud [b], Adewale F. Lukman [c], Victoria Laoye [d], Arowolo Olatunji [e]

[a] Department of Mathematics and Statistics, Federal University Wukari, Wukari, Nigeria
[b] Department of Mathematics, Al-Aqsa University, Gaza, Palestine
[c] Department of Mathematics, University of Medical Scieces, Ondo, Nigeria
[d] Department of Statistics, University of Ibadan, Ibadan, Nigeria
[e] Department of Mathematical Sciences, Lagos State University of Science and Technology, Lagos, Nigeria

Abstract

A new two-parameter estimator was developed to combat the threat of multicollinearity for the linear regression model. Some necessary and sufficient conditions for the dominance of the proposed estimator over ordinary least squares (OLS) estimator, ridge regression estimator, Liu estimator, KL estimator, and some two-parameter estimators are obtained in the matrix mean square error sense. Theory and simulation results show that, under some conditions, the proposed two-parameter estimator consistently dominates other estimators considered in this study. The real-life application result follows suit.

Mathematics Subject Classification: 62J05, 62J07

Keywords: Liu estimator, MSE, Multicollinearity, Ridge regression, Simulation study

1. Introduction

We consider the following linear regression model:

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

where $y_{n\times 1}$ vector of the response variable, $X_{n\times p}$ matrix of explanatory variables, $\beta_{p\times 1}$ vector of regression parameters, $\varepsilon_{n\times 1}$ vector of errors ($E(\varepsilon) = 0$, $V(\varepsilon) = \sigma^2 I_n$, and I_n identity matrix). The ordinary least squares estimator (OLSE) of β is

$$\widehat{\beta} = S^{-1}X'y \tag{2}$$

where $S = X'X$.

The OLSE performs poorly when the explanatory variables in the linear regression model (LRM) are collinear. Collinearity arises when there exists a linear relationship among the explanatory variables. In that case, it becomes difficult to make a valid inference about the regression parameters of the model. Several estimators have been suggested as alternatives to the OLSE to handle the threat of multicollinearity as [1,2,11,12,19,20,22,27–32,34,36], (B. M. [15], (B. M [15]., [8], (S [6,38].

In some situations, a linear relationship between predictors and other assumption violations can jointly occur. Researchers such as [20,22,33] proposed various Ridge-type M-Estimators that jointly combat the problem of multicollinearity and outliers in Y-direction. Also [3,7], suggested some method of parameter estimation in the linear regression model when multicollinearity and outliers are jointly present [24]. Proposed a Ridge-type estimator for the Gamma regression to handle the problem of multicollinearity when the response variable is right-skewed [39]. Proposed a two-stage K-L estimator that handled both multicollinearity and autocorrelation. In some real-life scenarios, predicting the

Received 9 November 2022; accepted 1 December 2022.
Available online 30 January 2023

* Corresponding author.
E-mail addresses: abdulrasheed@fuwukari.edu.ng (A.B. Badawaire), fadewale@unimed.edu.ng (A.F. Lukman), laoyevictoria@gmail.com (V. Laoye), arowolo.ot@lasustech.edu.ng (A. Olatunji).

https://doi.org/10.55810/2313-0083.1011

current values of a dependent variable based on the current values of a predictor when the explanatory variables are correlated with their lagged values yield unreliable results. To overcome this challenge [25], proposed the Almon-KL estimator (A-KLE).

We developed a new two-parameter estimator for parameter estimation in the LRM in this study and compared its performance with some existing ones. We will give very brief descriptions of these existing estimators in the following subsections.

1.1. Some alternative biased estimators

One of the most popular estimators dealing with multicollinearity is the ordinary ridge regression (ORR) estimator which was proposed by [11] and is defined as

$$\widehat{\beta}_k = W\,\widehat{\beta},\ \ k \geq 0 \tag{3}$$

where $W = (I_p + k\,S^{-1})^{-1}$ and k is the biasing parameter.

The other alternative biased estimator for dealing with multicollinearity is the Liu estimator defined by [12] as

$$\widehat{\beta}_d = F_d\,\widehat{\beta},\ \ 0 < d < 1 \tag{4}$$

where $F_d = (S + I_p)^{-1}(S + dI_p)$ and d is the biasing parameter.

Ozkale and Kaciranlar (2007) proposed the two-parameter estimator and is defined as

$$\widehat{\beta}_{kd} = R_{kd}\,\widehat{\beta},\ \ k \geq 0\,, 0 < d < 1 \tag{5}$$

where $R_{kd} = (S + k\,I_p)^{-1}(S + k\,d\,I_p)$.

Then [36], proposed another two-parameter estimator based on the Liu estimator and the ORR estimator, and is defined as

$$\widehat{\beta}(k,d) = F_d\,W\,\widehat{\beta},\ \ k \geq 0\,, 0 < d < 1 \tag{6}$$

And recently, (B. M. [15]), proposed a one-parameter estimator called Kibria-Lukman estimator and defined as,

$$\widehat{\beta}_{KL} = W\,M\ \widehat{\beta}, k \geq 0 \tag{7}$$

where $M = (I_p - k\,S^{-1})$.

1.2. The new two-parameter estimator

The new estimator is obtained by replacing $\widehat{\beta}$ with $\widehat{\beta}_{KL}$ in the Liu estimator and becomes as follows:

$$\widehat{\beta}_{proposed} = F_d\,W\,M\,\widehat{\beta},\ \ k \geq 0\,, 0 < d < 1 \tag{8}$$

The properties of the estimator are as follows:

$$E\left(\widehat{\beta}_{proposed}\right) = F_d\,W\,M\,E(\widehat{\beta}) = F_d\,W\,M\,\beta\,. \tag{9}$$

$$B\left(\widehat{\beta}_{proposed}\right) = \left[\,F_d\,W\,M - I_p\right]\beta\,. \tag{10}$$

$$D\left(\widehat{\beta}_{proposed}\right) = \sigma^2\,F_d\,W\,M\,S^{-1}\,M'\ \ W'\ \ F_d' \tag{11}$$

and the matrix mean squared error (MSEM) is

$$\begin{aligned} MSEM\left(\widehat{\beta}_{proposed}\right) &= \sigma^2\,F_d\,W\,M\,S^{-1}\,M'\ \ W'\ \ F_d' \\ &+ \left[\,F_d\,W\,M - I_p\right]\beta\,\beta'\left[\,F_d\,W\,M - I_p\right]' \end{aligned} \tag{12}$$

We rewrite (1) in the canonical form as follows:

$$y = Z\alpha + \varepsilon, \tag{13}$$

where $Z = X\,\tilde{Q}$ and $\alpha = \tilde{Q}'\,\beta$. $\tilde{Q}$ is an orthogonal matrix such that $Z'\,Z = \tilde{Q}\,X'\,X\,\tilde{Q} = \Lambda = diag\,(e_1, e_2, \ldots, e_p)$.

Then the OLS estimator of α in terms of eigenvalues is obtained as follows:

$$\widehat{\alpha} = \Lambda^{-1} Z'\,y, \tag{14}$$

with mean square error matrix (MSEM),

$$MSEM(\widehat{\alpha}) = \sigma^2 \Lambda^{-1}. \tag{15}$$

The ridge regression estimator of α is

$$\widehat{\alpha}(k) = W\,\widehat{\alpha}, \tag{16}$$

where $W = [I_p + k\,\Lambda^{-1}]^{-1}$.
and its MSEM is

$$MSEM(\widehat{\alpha}(k)) = \sigma^2\,W\,\Lambda^{-1} W' + \left(W - I_p\right)\alpha\,\alpha'\left(W - I_p\right)' \tag{17}$$

where $(W - I_p) = -\,k\,(\Lambda + k\,I_p)^{-1}$.
The Liu estimator of α is

$$\widehat{\alpha}(d) = F_d\,\widehat{\alpha}, \tag{18}$$

where $F_d = [\Lambda + I_p]^{-1}[\Lambda + d\,I_p]$.
and its MSEM is

$$\begin{aligned} MSEM(\widehat{\alpha}(d)) &= \sigma^2\,F_d\,\Lambda^{-1} F_d' \\ &+ (1-d)^2\left(\Lambda + I_p\right)^{-1}\alpha\alpha'\left(\Lambda + I_p\right)^{-1} \end{aligned} \tag{19}$$

The Ozkale and Kaciranlar estimator of α is

$$\widehat{\alpha}_{kd} = R_{kd}\,\widehat{\alpha}, \tag{20}$$

where $R_{kd} = [\Lambda + k\,I_p]^{-1}[\Lambda + k\,d\,I_p]$.

$$\begin{aligned} MSEM(\widehat{\alpha}_{kd}) &= \sigma^2\,R_{kd}\,\Lambda^{-1}\,R_{kd}' \\ &+ \left[R_{kd} - I_p\right]\ \alpha\,\alpha'\left[R_{kd} - I_p\right]' \end{aligned} \tag{21}$$

The Yang and Chang estimator of α is

$$\widehat{\alpha}(k,d)=F_d\,W\,\widehat{\alpha} \quad (22)$$

$$MSEM(\widehat{\alpha}(k,d))=\sigma^2\,F_d\,W\,\Lambda^{-1}\,W'\,F_d' + \left[F_d\,W-I_p\right]\,\alpha\,\alpha'\,\left[F_d\,W-I_p\right]' \quad (23)$$

The KL estimator of α is

$$\widehat{\alpha}_{KL}=W\,M\,\widehat{\alpha}, \quad (24)$$

where $M=[I_p-k\,\Lambda^{-1}]$ and its MSE is

$$MSEM(\widehat{\beta}_{KL})=\sigma^2\,W\,M\,\Lambda^{-1}\,M'\quad W' + \left[W\,M-I_p\right]\,\alpha\,\alpha'\,\left[W\,M-I_p\right]' \quad (25)$$

The Proposed estimator of α is

$$\widehat{\alpha}_{proposed}=F_d\,W\,M\,\widehat{\alpha}, \quad (26)$$

and its MSEM is,

$$MSEM\left(\widehat{\beta}_{proposed}\right)=\sigma^2\,F_d\,W\,M\,\Lambda^{-1}\,M'\,W'\,F_d' + \left[F_d\,W\,M-I_p\right]\,\alpha\,\alpha'\,\left[F_d\,W\,M-I_p\right]' \quad (27)$$

The following lemma is used to make some theoretical comparison among estimators in the section follow.

Lemma 1: Let $\widehat{\tilde{\beta}}_1$ and $\widehat{\tilde{\beta}}_2$ be two linear estimators of $\tilde{\beta}$. The difference in the covariance of the two estimators D is positive definite (pd) if and only if $bias(\widehat{\tilde{\beta}}_2)'[\sigma^2\tilde{D}+bias(\widehat{\tilde{\beta}}_1)bias(\widehat{\tilde{\beta}}_1)']^{-1}bias(\widehat{\tilde{\beta}}_2)<1$. Consequently, $MSEM(\widehat{\tilde{\beta}}_1)-MSEM(\widehat{\tilde{\beta}}_2)=\sigma^2 D+bias(\widehat{\tilde{\beta}}_1)bias(\widehat{\tilde{\beta}}_1)'$ $-bias(\widehat{\tilde{\beta}}_2)\;bias(\widehat{\tilde{\beta}}_1)'>0$ [35].

2. Comparison among the estimators

2.1. $\widehat{\alpha}$ and $\widehat{\alpha}_{proposed}$

Theorem 2.1: $\widehat{\alpha}_{proposed}$ dominates $\widehat{\alpha}$ if and only if

$$\alpha'\left[F_d\,W\,M-I_p\right]'\;\left[\sigma^2\left(\Lambda^{-1}-F_d\,W\,M\,\Lambda^{-1}\,M'\,W'\,F_d'\right)\right] \times\left[F_d\,W\,M-I_p\right]\;\alpha<1 \quad (28)$$

Proof:

$$D(\widehat{\alpha})-D\left(\widehat{\alpha}_{proposed}\right)=\sigma^2\left(\Lambda^{-1}-F_d\,W\,M\,\Lambda^{-1}\,M'\,W'\,F_d'\right) =\sigma^2\,diag\left\{\frac{1}{e_i}-\frac{(e_i+d)^2\,(e_i-k)^2}{e_i\,(e_i+1)^2(e_i+k)^2}\right\}_{i=1}^{p} \quad (29)$$

where $\Lambda^{-1}-\sigma^2\,F_d\,W\,M\,\Lambda^{-1}\,M'\,W'\,F_d'$ is pd since $(e_i+1)(e_i+k)-(e_i+d)(e_i-k)>0$.

Consequently, $\Lambda^{-1}-F_d\,W\,M\,\Lambda^{-1}\,M'\,W'\,F_d'$ is pd.

2.1. $\widehat{\alpha}(k)$ and $\widehat{\alpha}_{proposed}$

Theorem 2.2: $\widehat{\alpha}_{proposed}$ dominates $\widehat{\alpha}(k)$ if and only if

$$\alpha'\left[F_d\,W\,M-I_p\right]'\;\left[V_1+\left[W-I_p\right]\;\alpha\,\alpha'\;\left[W-I_p\right]'\right] \times\left[F_d\,W\,M-I_p\right]\;\alpha<1 \quad (30)$$

where $V_1=\sigma^2\left(W\,\Lambda^{-1}W'-F_d\,W\,M\,\Lambda^{-1}\,M'\,W'\,F_d'\right)$

Proof:

$$V_1=\sigma^2\left(W\,\Lambda^{-1}W'-F_d\,W\,M\,\Lambda^{-1}\,M'\,W'\,F_d'\right) =\sigma^2\,diag\left\{\frac{e_i}{(e_i+k)^2}-\frac{(e_i+d)^2\,(e_i-k)^2}{e_i\,(e_i+1)^2(e_i+k)^2}\right\}_{i=1}^{p} \quad (31)$$

where $W\,\Lambda^{-1}W'-F_d\,W\,M\,\Lambda^{-1}\,M'\,W'\,F_d'$ is pd since $e_i\,(k+1-d)+kd>0$.

2.3. $\widehat{\alpha}(d)$ and $\widehat{\alpha}_{proposed}$

Theorem 2.3: $\widehat{\alpha}_{proposed}$ dominates $\widehat{\alpha}(d)$ if and only if

$$\alpha'\left[F_d\,W\,M-I_p\right]'\left[V_2+\left[F_d-I_p\right]\right. \left.\alpha\,\alpha'\left[F_d-I_p\right]'\right]\left[F_d\,W\,M-I_p\right]\;\alpha<1 \quad (32)$$

where $V_2=\sigma^2\left(\mathrm{F_d}\,\Lambda^{-1}\mathrm{F_d'}-\mathrm{F_d}\,\mathrm{W}\,\mathrm{M}\,\Lambda^{-1}\,\mathrm{M'}\,\mathrm{W'}\,\mathrm{F_d'}\right)$

Proof: The difference between the dispersion matrices is

$$V_2=\sigma^2\left(F_d\,\Lambda^{-1}F_d'-F_d\,W\,M\,\Lambda^{-1}\,M'\,W'\,F_d'\right) =\sigma^2\,diag\left\{\frac{(e_i+d)^2}{e_i\,(e_i+1)^2}-\frac{(e_i+d)^2\,(e_i-k)^2}{e_i\,(e_i+1)^2(e_i+k)^2}\right\}_{i=1}^{p} \quad (33)$$

where $F_d\,\Lambda^{-1}F_d'-F_d\,W\,M\,\Lambda^{-1}\,M'\,W'\,F'$ *is pd since* $(e_i+k)-(e_i-k)=2\,k>0$.

2.4. $\widehat{\alpha}_{KL}$ and $\widehat{\alpha}_{proposed}$

Theorem 2.4: $\widehat{\alpha}_{proposed}$ dominates $\widehat{\alpha}_{KL}$ if and only if

$$\alpha'\left[F_d\,W\,M-I_p\right]'\;\left[V_3+\left[W\,M-I_p\right]\;\alpha\,\alpha'\left[W\,M-I_p\right]'\right] \times\left[F_d\,W\,M-I_p\right]\;\alpha<1 \quad (34)$$

where $V_3=\sigma^2(WM\Lambda^{-1}M'W'-F_dWM\Lambda^{-1}M'W'F_d')$

Proof:

$$V_3=\sigma^2\left(W\,M\,\Lambda^{-1}M'\,W'-F_d\,W\,M\,\Lambda^{-1}\,M'\,W'\,F_d'\right) =\sigma^2\,diag\left\{\frac{(e_i-k)^2}{e_i\,(e_i+k)^2}-\frac{(e_i+d)^2\,(e_i-k)^2}{e_i\,(e_i+1)^2(e_i+k)^2}\right\}_{i=1}^{p} \quad (35)$$

where $W\,M\,\Lambda^{-1}M'\,W'-F_d\,W\,M\,\Lambda^{-1}\,M'\,W'\,F_d'$ is pd since $(e_i+1)-(e_i+d)=1-d>0$.

2.5. $\widehat{\alpha}_{kd}$ and $\widehat{\alpha}_{proposed}$

Theorem 2.5: $\widehat{\alpha}_{proposed}$ dominates $\widehat{\alpha}_{kd}$ if and only if

$$\alpha' \left[F_d\, W\, M - I_p\right]' \left[V_4 + \left[R_{kd} - I_p\right]\ \alpha\, \alpha' \left[R_{kd} - I_p\right]'\right] \left[F_d\, W\, M - I_p\right]\ \alpha < 1 \tag{36}$$

where $V_4 = \sigma^2\,(R_{kd}\,\Lambda^{-1} R'_{kd} - F_d\, W\, M\, \Lambda^{-1}\, M'\, W'\, F'_d\,)$
Proof:

$$V_4 = \sigma^2\left(R_{kd}\,\Lambda^{-1} R'_{kd} - F_d\, W\, M\, \Lambda^{-1}\, M'\, W'\, F'_d\,\right) = \sigma^2\, diag\left\{\frac{(e_i + kd)^2}{e_i\,(e_i + k\,)^2} - \frac{(e_i + d)^2\,(e_i - k\,)^2}{e_i\,(e_i + 1\,)^2 (e_i + k\,)^2}\right\}_{i=1}^{p} \tag{37}$$

where $R_{kd}\,\Lambda^{-1} R'_{kd} - F_d\, W\, M\, \Lambda^{-1}\, M'\, W'\, F'_d$ is pd since $e_i\,(kd + k + 1 - d) + 2kd > 0$.

$$\frac{\partial m(k,d)}{\partial k} = -2\sigma^2 \sum_{i=1}^{p} \frac{(e_i + d)^2 (e_i - k)}{e_i (e_i + 1)^2 (e_i + k)^2} - 2\sigma^2 \sum_{i=1}^{p} \frac{(e_i + d)^2 (e_i - k)^2}{e_i (e_i + 1)^2 (e_i + k)^3}$$
$$+2(2e_i + d + 1) \sum_{i=1}^{p} \frac{(e_i(2k - d + 1) + k(d + 1))\,\alpha_i^2}{(e_i + 1)^2 (e_i + k)^2} - 2 \sum_{i=1}^{p} \frac{(e_i(2k - d + 1) + k(d + 1))^2 \alpha_i^2}{(e_i + 1)^2 (e_i + k)^3} \tag{41}$$

2.6. $\widehat{\alpha}(k,d)$ and $\widehat{\alpha}_{proposed}$

Theorem 2.6: $\widehat{\alpha}_{proposed}$ dominates $\widehat{\alpha}(k,d)$ if and only if

$$\alpha' \left[F_d\, W\, M - I_p\right]' \left[V_5 + \left[F_d\, W - I_p\right] \alpha\, \alpha' \left[F_d\, W - I_p\right]'\right] \times \left[F_d\, W\, M - I_p\right]\ \alpha < 1 \tag{38}$$

where $V_5 = \sigma^2\,(F_d\, W \Lambda^{-1} W'\, F'_d - F_d\, W\, M\, \Lambda^{-1}\, M'\, W'\, F'_d\,)$
Proof:

$$V_5 = \sigma^2\left(F_d\, W \Lambda^{-1} W'\, F'_d - F_d\, W\, M\, \Lambda^{-1}\, M'\, W'\, F'_d\,\right) = \sigma^2\, diag\left\{\frac{e_i\,(e_i + d)^2}{(e_i + 1)^2 (e_i + k\,)^2} - \frac{(e_i + d)^2\,(e_i - k\,)^2}{e_i\,(e_i + 1\,)^2 (e_i + k\,)^2}\right\}_{i=1}^{p} \tag{39}$$

where $F_d\, W \Lambda^{-1} W'\, F'_d - F_d\, W\, M\, \Lambda^{-1}\, M'\, W'\, F'_d$ is pd since $e_i - (e_i - k) = k > 0$.

2.7. Determination of the parameters k and d

In determining the optimal value of k, first we assume that d is fixed [10], (B. M. G. [14], (B. M. [13], and [17]. Then the optimal value of k can be obtained by choosing k that minimize

$$MSEM(\widehat{\alpha}_{proposed}) = E\left((\widehat{\alpha}_{proposed} - \alpha)'(\widehat{\alpha}_{proposed} - \alpha)\right)$$

$$m\,(k,d) = tr\left(MSEM(\widehat{\alpha}_{proposed})\right)$$

$$m\,(k,d) = \sigma^2 \sum_{i=1}^{p} \frac{(e_i + d)^2 (e_i - k)^2}{e_i (e_i + 1)^2 (e_i + k)^2} + \sum_{i=1}^{p} \frac{(e_i(2k - d + 1) + k(d + 1))^2 \alpha_i^2}{(e_i + 1)^2 (e_i + k)^2} \tag{40}$$

Differentiate equation (40) with respect to k

$\frac{\partial m(k,d)}{\partial k} = 0$, then

$$\widehat{k} = \frac{\sigma^2 e_i (e_i + d) + e_i^2 (d - 1)\alpha_i^2}{\sigma^2 (e_i + d) + e_i \alpha_i^2 (2e_i + d + 1)} \tag{42}$$

Equation (42) returns the biasing parameter for the KL estimator when $d = 1$, which is defined as follows:

$$\widehat{k} = \frac{\sigma^2}{2\alpha_i^2 + \frac{\sigma^2}{e_i}} \tag{43}$$

Differentiate equation (40) with respect to d gives

$$\frac{\partial m(k,d)}{\partial d} = 2\sigma^2 \sum_{i=1}^{p} \frac{(e_i + d)(e_i - k)^2}{e_i (e_i + 1)^2 (e_i + k)^2} - 2(e_i - k) \sum_{i=1}^{p} \frac{(e_i(2k - d + 1) + k(d + 1))\alpha_i^2}{(e_i + 1)^2 (e_i + k)^2} \tag{45}$$

$\frac{\partial m(k,d)}{\partial d} = 0$, then

$$\widehat{d} = \frac{e_i(e_i(2k + 1) + k)\ \alpha_i^2 - \sigma^2 e_i (e_i - k)}{\sigma^2 (e_i - k) + e_i (k - e_i)\ \alpha_i^2} \tag{46}$$

Substituting k to be zero in equation (46), equation (46) returns the biasing parameter for the Liu estimator in (46)

$$\hat{d}_{liu} = \frac{\alpha_i^2 - \sigma^2}{\alpha_i^2 + \frac{\sigma^2}{e_i}} \tag{47}$$

For the purpose of simplifying the selection of the biasing parameter k and d we carry out the following:

From equation (46),

$$\hat{d} = \frac{e_i(e_i(2k+1)+k)\,\alpha_i^2 - \sigma^2 e_i(e_i-k)}{\sigma^2(e_i-k)+e_i(k-e_i)\,\alpha_i^2} > 0$$

It implies that $e_i(e_i(2k+1)+k)\,\alpha_i^2 - \sigma^2 e_i(e_i-k) > 0$; parameter k is obtained as follows:

$$k > \frac{e_i(\sigma^2 - \alpha_i^2)}{2e_i\alpha_i^2 + \alpha_i^2 + \sigma^2} \tag{48}$$

In this study we take the absolute value and the minimum value of equation (48). The absolute value is taken to ensure the estimate return a positive value.

$$k_p > \left| \min\left(\frac{e_i(\sigma^2 - \alpha_i^2)}{2e_i\alpha_i^2 + \alpha_i^2 + \sigma^2} \right) \right| \tag{49}$$

The steps for the practical selection of the biasing parameters are as follows.

1. Obtain $\hat{k}_p$ by replacing σ^2 and α_i^2 with their unbiased estimates.
2. Substitute $\hat{k}_p$ into equation (46) and obtain the absolute value and the minimum value.

$$\hat{d}_p = \left| \min\left(\frac{e_i(e_i(2\hat{k}_p+1)+\hat{k}_p)\,\hat{\alpha}_i^2 - \hat{\sigma}^2 e_i(e_i-\hat{k}_p)}{\hat{\sigma}^2(e_i-\hat{k}_p)+\lambda_i(\hat{k}_p-e_i)\,\alpha_i^2} \right) \right| \tag{50}$$

The biasing parameters for the KL estimator is obtained using the following biasing parameter:

$$\hat{k}_{\min} = \min\left[\frac{\hat{\sigma}^2}{2\hat{\alpha}_i^2 + \frac{\hat{\sigma}^2}{e_i}} \right] \tag{51}$$

The biasing parameter by Ozkale and Kaciranlar (2007) is adopted for the Liu estimator in this study.

$$\hat{d} = \min\left[\frac{\hat{\sigma}^2}{\hat{\alpha}_i^2 + \frac{\hat{\sigma}^2}{e_i}} \right] \tag{52}$$

The biasing parameter for the ridge regression estimator in this study is given by

$$\hat{k} = \frac{p\hat{\sigma}^2}{\sum_{i=1}^{p} \hat{\alpha}_i^2} \tag{53}$$

3. Simulation study

In this section, we compare the performances of the following estimators: the OLSE, the ridge regression estimator, Liu estimator, KL estimator, Ozkale and Kaciranlar estimator, Yang and Chang estimator and the proposed estimator through a simulation study. The simulation study was conducted using the RStudio programming language. It contains two parts (i) Simulation technique and (ii) Simulation results discussion.

The simulation design depends on factors such as sample size n, degree of collinearity ρ and the number of explanatory variable p. Following [9], the degree of collinearity among the explanatory variables are generated using the following equation:

$$x_{ij} = (1-\rho^2)^{\frac{1}{2}} z_{ij} + \rho z_{i,p+1},\ i = 1,2,\ldots,n,\ j = 1,2,3,\ldots.,p, \tag{54}$$

where z_{ij} are independent standard normal pseudo-random numbers and ρ represents correlation between any two explanatory variables. We consider p = 4, and 8 explanatory variables in the simulation and the variables are standardized. The response variable y is obtained as follows:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \ldots + \beta_p x_{ip} + e_i \tag{55}$$

where i=1, 2, …,30,..,50, …,100, …,200, $e_i \sim N(0,\sigma^2)$. β is chosen such that $\beta'\beta = 1$ [20] and Newhouse and Oman (1971)). The MSE is computed as follows:

$$MSE(\alpha^*) = \frac{1}{5000}\sum_{i=1}^{5000} (\alpha^* - \alpha)'(\alpha^* - \alpha) \tag{56}$$

where α^* would be any of the estimators in this study. The estimator with the least MSE dominates others. The simulation results are presented in Tables 1 and 2. For a graphical representation, we also displayed the MSE values vs different values of the parameters and presented them in Fig. 1.

3.1. Simulation results discussion

We examine the effects of the different factors on the values of the MSE. It appears from Tables 1–2 and from Fig. 1, that the MSE increases as the variance and the number of explanatory variables increases, respectively when other variables are kept constant. For instance, from Table 1, when ρ = 0.99 and n = 50, the MSE increases from 9.603 to

Table 1. MSE of the estimators when $p = 4$.

Sigma	n	ρ	Estimators							
			$\widehat{\alpha}$	$\widehat{\alpha}(k)$	$\widehat{\alpha}(d)$	$\widehat{\alpha}_{KL}$	$\widehat{\alpha}_{kd}$	$\widehat{\alpha}(k,d)$	$\widehat{\alpha}_{proposed}$	$\widehat{\alpha}_{proposed}$
				$\widehat{k}$	$\widehat{d}$	$\widehat{k}_{\min}$	$(\widehat{k},\widehat{d})$	$(\widehat{k},\widehat{d})$	$(\widehat{k}_{\min},\widehat{d})$	$(\widehat{k}_p,\widehat{d}_p)$
1	50	0.9	3.956	0.318	0.447	0.140	0.243	0.362	0.111	**0.059**
		0.95	4.638	0.521	0.710	0.170	0.323	0.624	0.116	**0.111**
		0.99	9.603	1.907	1.384	0.615	0.519	2.480	**0.179**	0.399
		0.999	64.640	17.122	5.111	6.995	1.943	22.948	**0.644**	4.748
	100	0.9	1.231	0.137	0.174	0.092	0.124	0.146	0.084	**0.017**
		0.95	3.554	0.220	0.310	0.110	0.184	0.244	0.087	**0.030**
		0.99	5.019	0.664	0.882	0.157	0.337	0.826	**0.093**	0.129
		0.999	21.338	5.064	2.088	1.714	0.741	6.768	**0.219**	0.645
	200	0.9	1.152	0.072	0.083	0.059	0.069	0.075	0.056	**0.011**
		0.95	3.279	0.125	0.156	0.086	0.114	0.132	0.079	**0.013**
		0.99	4.016	0.385	0.567	0.119	0.261	0.452	0.084	**0.062**
		0.999	11.761	2.531	1.465	0.759	0.509	3.318	**0.146**	0.372
5	50	0.9	18.035	4.464	10.867	1.672	3.448	5.861	**1.293**	1.990
		0.95	33.220	8.756	17.392	3.401	5.468	11.643	**2.083**	5.011
		0.99	154.852	43.220	34.338	17.640	11.708	57.896	4.261	**3.971**
		0.999	1526.1	432.06	129.425	179.103	49.237	579.185	16.229	**12.350**
	100	0.9	5.720	1.566	4.236	0.480	1.434	1.984	0.445	**0.290**
		0.95	12.338	2.848	7.641	0.893	2.409	3.753	0.760	**0.652**
		0.99	48.560	13.136	22.146	4.590	6.807	17.566	**2.300**	8.945
		0.999	456.84	129.02	54.607	46.804	19.762	173.126	**6.115**	8.832
	200	0.9	3.200	0.792	2.053	0.227	0.759	0.978	0.219	**0.157**
		0.95	7.373	1.380	3.868	0.385	1.268	1.763	0.356	**0.167**
		0.99	24.405	6.109	14.124	2.075	4.190	8.082	**1.396**	1.451
		0.999	216.810	59.662	36.827	21.999	12.201	79.690	**2.807**	12.230

Table 2. MSE of the estimators when $p = 8$.

Sigma	n	ρ	Estimators							
			$\widehat{\alpha}$	$\widehat{\alpha}(k)$	$\widehat{\alpha}(d)$	$\widehat{\alpha}_{KL}$	$\widehat{\alpha}_{kd}$	$\widehat{\alpha}(k,d)$	$\widehat{\alpha}_{proposed}$	$\widehat{\alpha}_{proposed}$
				$\widehat{k}$	$\widehat{d}$	$\widehat{k}_{\min}$	$(\widehat{k},\widehat{d})$	$(\widehat{k},\widehat{d})$	$(\widehat{k}_{\min},\widehat{d})$	$(\widehat{k}_p,\widehat{d}_p)$
1	50	0.9	6.273	0.525	0.754	0.228	0.403	0.546	0.185	**0.111**
		0.95	7.289	0.829	1.166	0.265	0.518	0.875	0.185	**0.151**
		0.99	15.381	2.962	1.916	0.931	0.700	3.210	**0.228**	0.419
		0.999	105.843	26.494	6.981	10.442	2.713	29.040	**0.741**	1.539
	100	0.9	6.174	0.252	0.395	0.117	0.224	0.257	0.104	**0.021**
		0.95	6.615	0.359	0.690	0.187	0.286	0.371	0.172	**0.036**
		0.99	10.087	0.871	1.672	0.274	0.362	0.932	0.200	**0.146**
		0.999	49.033	6.100	2.265	6.084	0.827	6.989	**0.243**	0.734
	200	0.9	5.796	0.164	0.188	0.133	0.155	0.166	0.126	**0.009**
		0.95	6.000	0.279	0.348	0.190	0.251	0.283	0.141	**0.010**
		0.99	7.518	0.836	1.184	0.236	0.530	0.875	0.155	**0.036**
		0.999	24.541	5.314	1.787	1.075	0.546	5.753	0.170	**0.149**
5	50	0.9	30.508	7.103	18.697	2.363	5.521	7.743	**1.822**	2.417
		0.95	55.239	13.570	29.094	4.861	8.583	14.841	**2.961**	5.729
		0.99	254.712	65.714	47.887	25.761	15.621	72.013	**5.441**	12.747
		0.999	2509.207	654.713	149.564	265.615	57.679	718.474	**17.014**	21.284
	100	0.9	16.670	3.294	9.711	0.715	2.936	3.556	0.642	**0.493**
		0.95	27.383	6.055	17.144	1.440	4.859	6.586	1.154	**0.830**
		0.99	113.671	28.203	42.737	7.752	11.821	30.821	**3.115**	10.793
		0.999	1088.421	278.121	63.763	79.847	22.570	304.251	**3.243**	18.840
	200	0.9	10.553	1.700	4.663	0.331	1.612	1.818	0.316	**0.254**
		0.95	15.283	2.950	8.649	0.556	2.662	3.171	0.503	**0.257**
		0.99	53.054	12.852	29.559	2.918	8.185	13.970	1.828	**1.665**
		0.999	477.882	124.180	44.763	30.492	12.814	135.340	**2.898**	25.439

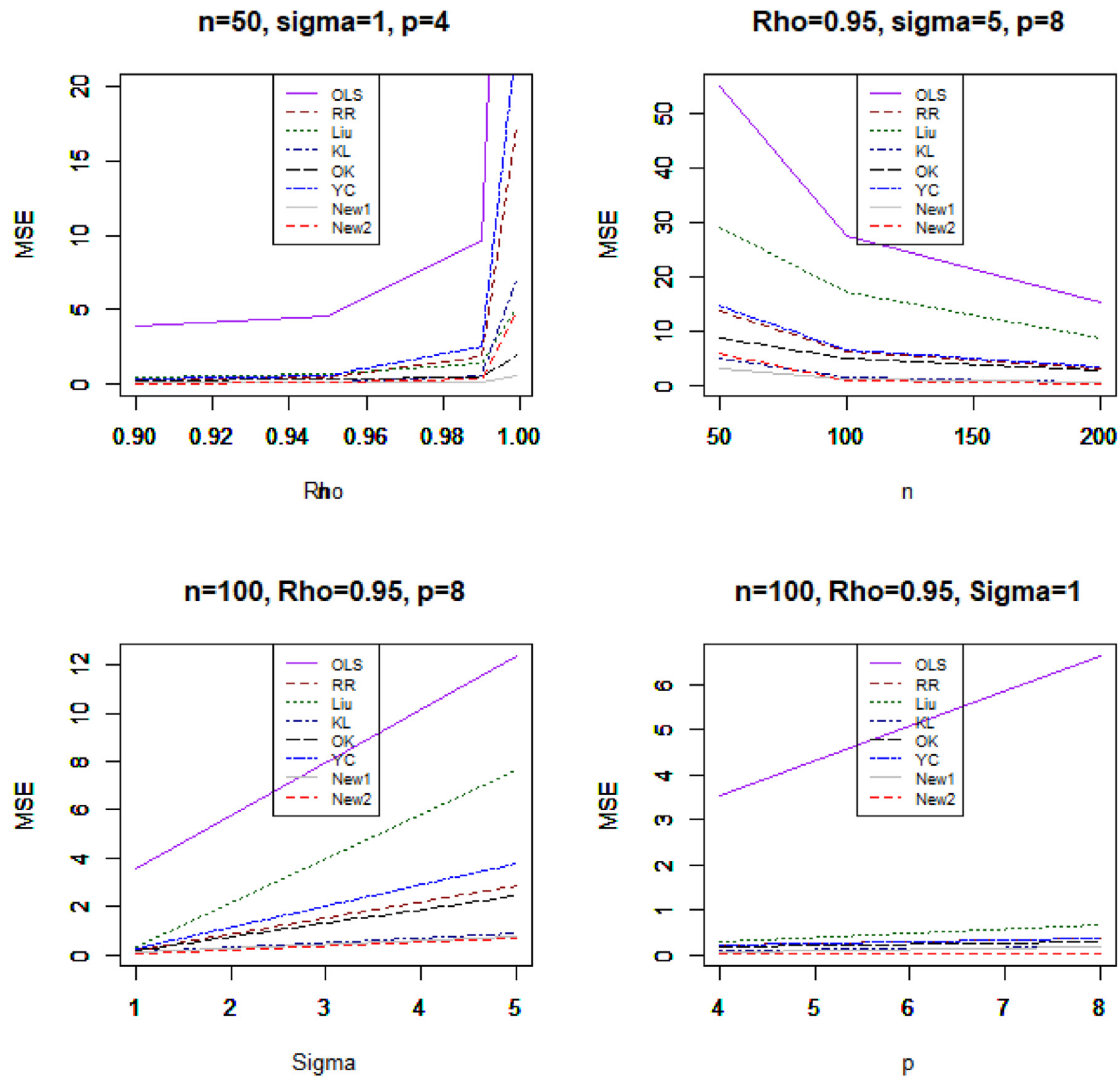

Fig. 1. : Estimated MSE values of the estimators (OLS, RR, Liu, KL, OK (Ozkale and Kaçiranlar), YC (Yang and Chang), New 1 and New 2) for different values of ρ.

154.852 as σ increases from 1 to 5. Keeping other variables constant, the MSE values of each of the estimators decrease as the sample size increases. Every other estimator considered dominate the OLSE as expected. Recall that the biased estimators considered in this study as an alternative to the OLSE can be classified into two types. Some of the estimators possess a single biasing parameter e.g. Ridge estimator, Liu estimator and the KL estimator. These estimators are called single parameter estimator while the proposed estimator, Yang and Chang estimator and Ozkale and Kaciranlar estimator exhibit two biasing parameter and are called Two-parameter estimators. From the result in Tables 1 and 2 and Fig. 1, we observed that the KL estimator outperforms the other single parameter estimator with few exceptions when $\rho = 0.999$. Among the two-parameter estimators, $\widehat{\alpha}_{proposed}$ consistent dominates. The overall performance

Table 3. Correlation coefficient among the variables.

	y	x1	x2	x3	x4	x5	x6
y	1	0.971[a]	0.984[a]	0.960[a]	0.502[b]	0.457	0.971[a]
x_1	0.971[a]	1	0.992[a]	0.979[a]	0.621[b]	0.465	0.991[a]
x_2	0.984[a]	0.992[a]	1	0.991[a]	0.604[b]	0.446	0.995[a]
x_3	0.960[a]	0.979[a]	0.991[a]	1	0.687[a]	0.364	0.994[a]
x_4	0.502[b]	0.621[b]	0.604[b]	0.687[a]	1	−0.177	0.668[a]
x_5	0.457	0.465	0.446	0.364	−0.177	1	0.417
x_6	0.971[a]	0.991[a]	0.995[a]	0.994[a]	0.668[a]	0.417	1

[a] Correlation is significant at the 0.01 level (2-tailed).
[b] Correlation is significant at the 0.05 level (2-tailed).

Table 4. Parameter estimation output and their MSE.

	Estimators							
	$\widehat{\alpha}$	$\widehat{\alpha}(k)$	$\widehat{\alpha}(d)$	$\widehat{\alpha}_{KL}$	$\widehat{\alpha}_{kd}$	$\widehat{\alpha}(k,d)$	$\widehat{\alpha}_{proposed}$	$\widehat{\alpha}_{proposed}$
		$\widehat{k}$	$\widehat{d}$	$\widehat{k}_{\min}$	$(\widehat{k},\widehat{d})$	$(\widehat{k},\widehat{d})$	$(\widehat{k}_{\min},\widehat{d})$	$(\widehat{k}_p,\widehat{d}_p)$
β_1	−52.994	1.093	−49.641	−5.041	1.253	−7.794	−4.492	−5.042
β_2	0.071	0.053	0.070	0.061	0.053	s	0.061	0.061
β_3	−0.414	−0.206	−0.408	−0.327	−0.206	−0.240	−0.325	−0.326
β_4	−0.423	−0.646	−0.432	−0.543	−0.646	−0.609	−0.544	−0.543
β_5	−0.573	−0.561	−0.574	−0.598	−0.561	−0.563	−0.599	−0.598
β_6	48.418	37.120	48.046	42.921	37.085	38.976	42.843	42.908
K		262.88		9.5524	262.88	262.88	9.3125	9.5524
D			0.1643		0.1643	0.1643	0.1643	5.06e-08
MSE	17095.15	3190.611	15182.57	2915.835	3204.145	2945.292	2916.601	**2914.771**

revealed that the proposed estimator outperformed all other estimators considered in this study. Furthermore, we observed that the performance of the proposed estimator is a function of its biasing parameter. The biasing parameter $(\widehat{k}_p, \widehat{d}_p)$ demonstrates great performance when there is severe multicollinearity when $\rho = 0.999$.

4. Example: longley data

This data was originally adopted by [16] and has been adopted by other authors in the study that involve multicollinearity (A. [37], and [18]. The model is as follows:

$$y = \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_5 x_5 + \beta_6 x_6 + e \tag{62}$$

The variable descriptions are available in the references. According to Walker and Birch (1988), the condition number is 43,275 which revealed the model suffers a severe multicollinearity. The correlation coefficient is illustrated in Table 3. The result in Table 3 shows that a significant and strong positive relationship exist among the following explanatory variables: x_1 and x_2; x_1 and x_3; x_1 and x_6; x_2 and x_3; x_2 and x_6; x_3 and x_6; The variance inflation factor (VIF) of the variables are as follows: 7266.945, 1096.619, 101193.162, 42.538, 39.890 and 2365.095. Following [20], there is multicollinearity when VIF exceeds 10. The result is also evident that there is severe multicollinearity. The result in Table 4 shows that the OLSE performed least when compared with other estimators. The estimator has a very significant MSE and the results agree with the simulation study. The proposed estimator with biasing parameter $(\widehat{k}_{\min}, \widehat{d})$ and $(\widehat{k}_p, \widehat{d}_p)$, and the K-L estimator performed very well by possessing the lowest MSE. The proposed estimator with biasing parameter $(\widehat{k}_p, \widehat{d}_p)$ dominated other estimators at some particular k and d while the K-L estimator competes favorably. The performance of the proposed agreed with the simulation result.

5. Some concluding remarks

In this paper, a new two-parameter biased estimator is proposed, and its biasing parameter were developed to circumvent the problem of multicollinearity in the linear regression set up. The properties of the new estimator are obtained, and we compared the estimator with some existing estimators theoretically. The theoretical comparison shows the proposed estimator dominate other existing estimators at some particular k and d. However, this is not sufficient and thus, it is necessary to conduct a simulation and a real-life study to further support the theoretical results. It is also evident from both results that the proposed estimator outperformed other existing estimators considered in this study. Consequently, we recommended this estimator for researchers in modelling a linear regression model with the challenge of multicollinearity. This method can also be extended to the generalized linear models.

Disclosure of any funding to the study

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

[1] Akdeniz F, Kaçiranlar S. On the almost unbiased generalized Liu estimator and unbiased estimation of the bias and MSE. Commun Stat Theor Methods 1995;24(7):1789–97.

[2] Akdeniz F, Roozbeh M. Generalized difference-based weighted mixed almost unbiased ridge estimator in partially linear models. Stat Pap 2019;60(5):1717–39.

[3] Awwad FA, Dawoud I, Abonazel MR. Development of robust Özkale–Kaçiranlar and Yang–Chang estimators for regression models in the presence of multicollinearity and outliers. Concurrency Comput Pract Ex 2021:e6779.

[6] Dawoud I. A new improved estimator for reducing the multicollinearity effects. Commun Stat Simulat Comput 2021:1–12.

[7] Dawoud I, Abonazel MR. Robust Dawoud–Kibria estimator for handling multicollinearity and outliers in the linear

regression model. J Stat Comput Simulat 2021;91(17): 3678–92.
[8] Dawoud I, Kibria BM. A new biased estimator to combat the multicollinearity of the Gaussian linear regression model. Stats 2020;3(4):526–41.
[9] Gibbons DG. A simulation study of some ridge estimators. J Am Stat Assoc 1981;76(373):131–9.
[10] Hoerl AE, Kannard RW, Baldwin KF. Ridge regression: some simulations. Commun Stat Theor Methods 1975;4(2): 105–23.
[11] Hoerl AE, Kennard RW. Ridge regression: biased estimation for nonorthogonal problems. Technometrics 1970;12(1): 55–67.
[12] Kejian L. A new class of blased estimate in linear regression. Commun Stat Theor Methods 1993;22(2):393–402.
[13] Kibria BM, Banik S. Some ridge regression estimators and their performances. 2020.
[14] Kibria BMG. Performance of some new ridge regression estimators. Commun Stat Simulat Comput 2003;32(2):419–35.
[15] Kibria BM, Lukman AF. A new ridge-type estimator for the linear regression model: simulations and applications. Scientifica; 2020.
[16] Longley JW. An appraisal of least squares programs for the electronic computer from the point of view of the user. J Am Stat Assoc 1967;62(319):819–41.
[17] Lukman AF, Ayinde K. Review and classi cations of the ridge parameter estimation techniques. Hacettepe Journal of Mathematics and Statistics 2017;46(5):953–67.
[18] Lukman AF, Ayinde K. Detecting influential observations in two-parameter Liu-ridge estimator. J Data Sci 2018;16(2): 2017–218.
[19] Lukman AF, Ayinde K, Binuomote S, Clement OA. Modified ridge-type estimator to combat multicollinearity: application to chemical data. J Chemometr 2019a;33(5):e3125.
[20] Lukman AF, Ayinde K, Kibria BMG, Adewuyi ET. Modified ridge-type estimator for the gamma regression model. Commun Stat Simulat Comput 2020:1–15.
[22] Lukman AF, Ayinde K, Oludoun O, Onate CA. Combining modified ridge-type and principal component regression estimators. Scientific African 2020c;9:e00536.
[24] Lukman AF, Dawoud I, Kibria BM, Algamal ZY, Aladeitan B. A new Ridge-type estimator for the Gamma regression model. Scientifica; 2021.
[25] Lukman AF, Kibria GBM. Almon-KL estimator for the distributed lag model. Arab Journal of Basic and Applied Sciences 2021;28(1):406–12.
[27] Massy WF. Principal components regression in exploratory statistical research. J Am Stat Assoc 1965;60(309):234–56.
[28] Mayer LS, Willke TA. On biased estimation in linear models. Technometrics 1973;15(3):497–508.
[29] Özkale MR, Kaciranlar S. The restricted and unrestricted two-parameter estimators. Commun Stat Theor Methods 2007;36(15):2707–25.
[30] Roozbeh M. Optimal QR-based estimation in partially linear regression models with correlated errors using GCV criterion. Comput Stat Data Anal 2018;117:45–61.
[31] Sakallõoğlu S, Kaçõranlar S. A new biased estimator based on ridge estimation. Stat Pap 2008;49(4):669–89.
[32] Stein C. Inadmissibility of the usual estimator for the mean of a multivariate normal distribution. Relatório técnico. Stanford University Stanford United States; 1956.
[33] Suhail M, Chand S, Babar I. ON some new ridge M-ESTIMATORS for linear regression models under various error distributions. Pak. J. Statist 2021;37(4):369–91. Pak J. Statist. 2021 Vol. 37 (4), 369-391.
[34] Swindel BF. Good ridge estimators based on prior information. Commun Stat Theor Methods 1976;5(11):1065–75.
[35] Trenkler G, Toutenburg H. Mean squared error matrix comparisons between biased estimators—an overview of recent results. Stat Pap 1990;31(1):165–79.
[36] Yang H, Chang X. A new two-parameter estimator in linear regression. Commun Stat Theor Methods 2010;39(6):923–34.
[37] Yasin A, Erisoglu M. Influence diagnostics in two-parameter ridge regression. J Data Sci 2016;14(1):33–51.
[38] Yasin S, Kamal S, Suhail M. Performance of some new ridge parameters in two-parameter ridge regression model. Iran J Sci Technol Trans A-Science 2021;45(1):327–41.
[39] Zubair MA, Adenomon MO. Comparison of estimators efficiency for linear regressions with joint presence of autocorrelation and multicollinearity. Sci World J 2021;16(2):103–9.