An unbiased estimator with prior information

Adewale F. Lukman\textsuperscript{a,c}, Kayode Ayinde\textsuperscript{b}, Benedicta Aladeitan\textsuperscript{a} and Rasak Bamidele\textsuperscript{a}

\textsuperscript{a}Department of Physical Sciences and Social Sciences, Landmark University, Omu-Aran, Kwara State, Nigeria; \textsuperscript{b}Department of Statistics, Federal University of Technology, Akure, Ondo State, Nigeria; \textsuperscript{c}Department of Mathematics, Centre Emile Borel, Institut Henri Poincare, Paris, France

\textbf{ABSTRACT} 

The ordinary least square (OLS) estimator suffers a breakdown in the presence of multicollinearity. The estimator is still unbiased but possesses a significant variance. In this study, we proposed an unbiased modified ridge-type estimator as an alternative to the OLS estimator and the biased estimators for handling multicollinearity in linear regression models. The properties of this new estimator were derived. The estimator is also unbiased with minimum variance. A real-life application to the higher heating value of poultry waste from proximate analysis and simulation study generally supported the findings.

\section{1. Introduction}

Consider the linear regression model

\[ y = X\beta + e, \quad e \sim N(0, \sigma^2 I) \quad (1) \]

where \( y \) is a \( n \times 1 \) vector of the dependent variable, \( X \) is a known \( n \times p \) full rank matrix of explanatory variables, \( \beta \) is a \( p \times 1 \) vector of regression coefficients and \( I \) is an \( n \times n \) identity matrix. The ordinary least squares estimator (OLS) of \( \beta \) in model (1) is defined as:

\[ \hat{\beta}_{OLS} = (S)^{-1}X'y \quad (2) \]

where \( S = X'X \).

This estimator is the most widely used method to estimate the parameters in a linear regression model. It performs best when certain assumptions are satisfied. One of them is that the independent variables are not associated. However, in practice, there often exist strong or perfect linear relationships among the independent variables. This situation is called multicollinearity. The OLS estimator suffers a breakdown in the presence of multicollinearity. The estimator is still unbiased but possesses a significant variance (Ayinde, Lukman, Samuel, & Attah, 2018). Different approaches are available in the literature to handle this problem. These include Hoerl and Kennard (1970), Swindel (1976), Farebrother (1976), Liu (1993), Sakalloglu and Akdeniz (2003), Ozkale and Kaciranlar (2007), Yang and Chang (2010), Li and Yang (2012), Wu and Yang (2013), Wu (2014) and recently, Arumairajan and Wijekoon (2017), Ayinde et al. (2018), Lukman, Ayinde, Binuomote, and Onate (2019). The estimators by these authors are biased. Crouse, Jin, and Hanumara (1995) and Sakalloglu and Akdeniz (2003) proposed the unbiased version of the ridge estimator and Liu estimator, respectively, with the addition of prior information. These methods effectively handle the problem of multicollinearity and eliminate bias.

In this article, we proposed an unbiased modified ridge-type estimator (UMRT) with prior information and derived its properties. Furthermore, we discuss the performance of the proposed estimator over the OLS estimator, the Ridge estimator (RE) and the modified ridge-type estimator (MRT) using the mean square error matrix (MSEM) criteria.

The remaining part of this article is as follows. In Section 2, we proposed the unbiased modified ridge-type estimator and compared its performance with some existing estimators using the mean square error matrix (MSEM) criterion in Section 3. We estimate the biasing parameter \( k \) and \( d \) in Section 4. We conducted a simulation study and a real-life data application in Section 5. Finally, we provide some concluding remarks in Section 6.

\section{2. Unbiased modified ridge-type estimator with prior information}

Hoerl and Kennard (1970) defined the ridge estimator of \( \beta \) as:

\[ \hat{\beta}_{RE}(k) = (S + kI)^{-1}X'y, \quad k > 0 \quad (3) \]

where \( k \) is the biasing parameter.

\textbf{CONTACT} Adewale F. Lukman \textsuperscript{a} adewale.folaranmi@lmu.edu.ng Department of Physical Sciences and Social Sciences, Landmark University, Omu-Aran, Kwara State, Nigeria.
Swindel (1976) defined the ridge estimator with prior information \( b \)

\[
\hat{\beta}_{\text{MRE}}(k, b) = (S + k I)^{-1}(X'y + kb)
\]

(4)

Crouse et al. (1995) introduced the unbiased ridge estimator based on the ridge estimator and prior information \( J \). This is defined as

\[
\hat{\beta}_{\text{UMRE}} = (S + k I)^{-1}(X'y + kJ)
\]

(5)

where \( J \) and \( \hat{\beta}_{\text{OLS}} \) are uncorrelated and \( J \sim N(\beta, V) \) such that \( V = \left( \frac{1}{k} \right) I_p \) and \( I_p \) is \( p \times p \) identity matrix. \( J \) is estimated by \( J = \frac{1}{p} \sum_{i=1}^{p} \hat{\beta}_i \).

Lukman et al. (2019) proposed the modified ridge-type estimator which is defined as follows:

\[
\hat{\beta}_{\text{MRT}}(k, d) = \left[ S + k(1 + d) \right]^{-1} S \hat{\beta}_{\text{OLS}} = F_{kd} \hat{\beta}_{\text{OLS}}
\]

(6)

where \( F_{kd} = \left[ S + k(1 + d) \right]^{-1} S \).

Following the convex estimator

\[
\hat{\beta}(C, J) = C \hat{\beta}_{\text{OLS}} + (I - C)J
\]

(7)

where \( C \) is a \( p \times p \) matrix and \( I \) is a \( p \times p \) identity matrix. Consequently, the mean square error of \( \hat{\beta}(C, J) \) is

\[
\text{MSE}(\hat{\beta}(C, J)) = \sigma^2 C S^{-1} C' + (I - C)V(I - C)'
\]

(8)

Then,

\[
\frac{\partial \text{MSE}(\hat{\beta}(C, J))}{\partial C} = 2C(\sigma^2 S^{-1} + V) - 2V = 0
\]

(9)

From (9), \( C \) is obtained to be

\[
C = V(\sigma^2 S^{-1} + V)^{-1}.
\]

Accordingly, \( V = \sigma^2 (I - C) S^{-1} (I - C)' \). The convex estimator \( \hat{\beta}(C, J) \) has minimum MSE for optimal value of \( C \) and it’s an unbiased estimator of \( \beta \). Therefore, the new estimator in this study is defined as

\[
\hat{\beta}_{\text{UMRT}}(F_{kd}, J) = F_{kd} \hat{\beta}_{\text{OLS}} + (I - F_{kd})J
\]

(10)

where \( F_{kd} = \left[ S + k(1 + d) \right]^{-1} S \), then, the value of \( V = \frac{\sigma^2}{k(1+d)} \). Consequently, \( J \sim N(\beta, \frac{\sigma^2}{k(1+d)}) \) for \( k > 0, 0 < d < 1 \).

It is easy to show that \( \hat{\beta}_{\text{UMRT}}(F_{kd}, J) \) is an unbiased estimator of \( \beta \). The expectation vector, bias vector, dispersion matrix and mean square error matrix of the proposed estimator are:

\[
\mathbb{E}(\hat{\beta}_{\text{UMRT}}(F_{kd}, J)) = \frac{\mathbb{E}(\hat{\beta}_{\text{MRT}} + (I - F_{kd})J)}{\mathbb{E}(\hat{\beta}_{\text{OLS}})}
\]

(11)

\[
\text{Bias}(\hat{\beta}_{\text{UMRT}}(F_{kd}, J)) = \mathbb{E}(\hat{\beta}_{\text{UMRT}}(F_{kd}, J)) - \beta = \beta - \beta = 0
\]

(12)

\[
\text{D}(\hat{\beta}_{\text{UMRT}}(F_{kd}, J)) = \text{D}(\beta_{\text{OLS}} + (I - F_{kd})J)
\]

(13)

Since \( \text{Bias} = 0 \), then

\[
\text{MSEM}(\hat{\beta}_{\text{UMRT}}(F_{kd}, J)) = \mathbb{E}(\hat{\beta}_{\text{MRT}} + (I - F_{kd})J)
\]

(14)

Consequently, the estimator \( \hat{\beta}_{\text{UMRT}}(F_{kd}, J) \) is an unbiased estimator of \( \beta \).

Suppose there exist an orthogonal matrix \( Q \) such that \( Q'XQ = \Lambda = \text{diag}(\lambda_1, \lambda_2, ..., \lambda_p) \) where \( \lambda_i \) is the \( i \)th eigenvalue of \( X'X \). \( \Lambda \) and \( Q \) are the matrices of eigenvalues and eigenvectors of \( X'X \), respectively. Model (1) can be written in canonical form as:

\[
y = Zx + \epsilon
\]

(15)

where \( Z = XQ \), \( x = Q'\beta \) and \( Z'Z = \Lambda \). For model (15), we get the following representations:

\[
\hat{\beta}_{\text{OLS}} = \Lambda^{-1} Z'y
\]

(16)

\[
\hat{\beta}_{\text{MRT}} = (\Lambda + k)\Lambda^{-1} Z'y
\]

(17)

\[
\hat{\beta}_{\text{UMRT}}(k, d) = (\Lambda + k(1 + d))^{-1} \Lambda \hat{\beta}_{\text{OLS}}
\]

(18)

\[
\hat{\beta}_{\text{UMRT}}(F_{kd}, J) = \hat{\beta}_{\text{MRT}}(k, d) + (I - F_{kd})J
\]

(19)

**Lemma 2.1.** Let \( M \) be an \( n \times n \) positive definite matrix, that is \( M > 0 \), and \( x \) be some vector, then \( Mx'x \geq 0 \) if and only if \( x'M^{-1}x \leq 1 \) (Farebrother, 1976).

**Lemma 2.2.** Let \( \hat{\beta}_1 = Ay_i = 1, 2 \) be two linear estimators of \( \beta \). Suppose that \( D = \text{Cov}(\hat{\beta}_1) - \text{Cov}(\hat{\beta}_2) > 0 \), where \( \text{Cov}(\hat{\beta}_i), i = 1, 2 \) denotes the covariance matrix of \( \hat{\beta}_i \) and \( b_1 = \text{Bias}(\hat{\beta}_i) = (AX - I)\beta, i = 1, 2 \). Consequently,

\[
\Delta(\hat{\beta}_1 - \hat{\beta}_2) = \text{MSEM}(\hat{\beta}_1) - \text{MSEM}(\hat{\beta}_2)
\]

\[
= \sigma^2 D + b_1 b_1' - b_2 b_2' > 0
\]

if and only if \( b_2'\sigma^2 D + b_1 b_1' \) is positive definite for all \( k > 0 \) and \( 0 < d < 1 \). (Trenkler & Toutenburg, 1990).

3. Theoretical Comparisons

3.1. Comparison of the OLS estimator and the unbiased modified ridge-type estimator

**Theorem 3.1.** The unbiased modified ridge-type estimator \( \hat{\beta}_{\text{UMRT}}(F_{kd}, J) \) is superior to the OLS estimator in the mean square error sense for \( k > 0 \) and \( 0 < d < 1 \).

**Proof.** By Definition,

\[
\text{MSEM}(\hat{\beta}_{\text{OLS}}) = \sigma^2 \Lambda^{-1}
\]

(21)

The MSEM difference between Eqs. (14) and (21)

\[
\text{MSEM}(\hat{\beta}_{\text{OLS}}) - \text{MSEM}(\hat{\beta}_{\text{UMRT}}(F_{kd}, J))
\]

\[
= \sigma^2 \Lambda^{-1} - \sigma^2 (\Lambda + k(1+d))^{-1}
\]

\[
= \sigma^2 (\Lambda^{-1} - (\Lambda + k(1+d))^{-1})
\]

\[
= \sigma^2 \text{diag} \left[ \frac{1}{\lambda_i} - \frac{1}{(\lambda_i + k(1+d))} \right]_{i=1}^p
\]

It was observed that \( \Lambda^{-1} - (\Lambda + k(1+d))^{-1} \) will be positive definite if and only if \( \lambda_i \) is such that \( \lambda_i + k(1+d) > \lambda_i > 0 \). However, for \( k > 0 \) and \( 0 < d < 1 \), \( \lambda_i + k(1+d) > \lambda_i \)
The difference between Eqs. (14) and (23)

\[ \text{MSEM}(\hat{\beta}_{RE}(k)) = \sigma^2 \text{diag} \left( \frac{\hat{\lambda}_i}{(\hat{\lambda}_i+k(1+d))^2} \right) \]

where \( \hat{\lambda}_i > 0 \) for any \( i \) in \( \{1, \ldots, p\} \). We observed that \( (\Lambda + k(1+d))^{-1} \Lambda(\Lambda + k(1+d))^{-1} \) will be positive definite if and only if \( \hat{\lambda}_i(\hat{\lambda}_i + k(1+d)) - (\hat{\lambda}_i + k(1+d))^2 \geq 0 \) or \( \hat{\lambda}_i(d+1) > k \). where \( k > 0 \) and \( 0 < d < 1 \).

Theorem 3.3. The unbiased modified ridge type estimator always dominates the modified ridge type estimator in the MSEM sense for \( k > 0 \) and \( 0 < d < 1 \).

Proof: The difference between Eqs. (14) and (28)

\[ \text{MSEM}(\hat{\beta}_{MRT}(k,d)) - \text{MSEM}(\hat{\beta}_{MRT}(F_{kd,J})) = \sigma^2 k(1+d)[\Lambda + (k(1+d))^{-1}][\Lambda + k(1+d)]^{-1} \]

(29)

Therefore, MSEM(\( \hat{\beta}_{MRT}(k,d) \)) is a non-negative matrix for \( k > 0 \) and \( 0 < d < 1 \). The proof of Theorem 3.3 is completed.

4. Estimation of the biasing parameters \( k \) and \( d \)

In this section, we discuss the estimation of the biasing parameter \( k \) and \( d \).

4.1. The estimation of parameter \( d \)

In the definition of the new estimator, \( J \) and \( \tilde{z}_{OLS} \) are uncorrelated. Therefore, \( (\tilde{z}_{OLS} - J) \sim N(0, \frac{\sigma^2}{(\Lambda^1k(1+d) + 1)}) \) and

\[ E[(\tilde{z}_{OLS} - J)(\tilde{z}_{OLS} - J)^t] = \frac{\sigma^2}{k(1+d)}[p + k(1+d)tr(\Lambda^{-1})] \]

(30)

From (30), if \( \sigma^2 \) is known for a fixed \( k \), we can get an unbiased estimator of \( d \) as follows:

\[ \hat{d} = \frac{p\sigma^2}{k[(\hat{\beta}_{OLS} - J)(\hat{\beta}_{OLS} - J)^t] \sigma^2 tr(\Lambda^{-1}) - 1} \]

(31)

When \( \sigma^2 \) is unknown, \( s^2 \) is used as an estimate of \( \sigma^2 \):

\[ s^2 = \frac{(Y - X\hat{\beta}_{OLS})'(Y - X\hat{\beta}_{OLS})}{n-p} \]

(32)

Consequently,

\[ \hat{d} = \frac{p s^2}{k[(\hat{\beta}_{OLS} - J)(\hat{\beta}_{OLS} - J)^t] - s^2 tr(\Lambda^{-1}) - 1} \]

(33)

where \( tr(\Lambda^{-1}) = \sum_{i=1}^p \frac{1}{\lambda_i} \) and \( \lambda_i \) is the eigen-value of \( X'X \). It was observed that the estimator of \( d \) in (33) can return a negative value. To eliminate the negative value, Wu (2014) suggests replacing \( \hat{d} \) with one (1) when its estimate is negative. Here, in this study, when \( d \) in Eq. (33) is negative, we adopt the estimator of \( \hat{d} \) suggested by Ozkale and Kaciranlar (2007) as follows:

\[ \hat{d}^* = \min \left[ \frac{s^2}{\tilde{\lambda}_i^2 + s^2} \right] \]

(34)

4.2. Estimating the biasing parameter \( k \)

From Eq. (30), if \( \sigma^2 \) is known and \( d \) is assumed to be fixed, an unbiased estimate of \( k \) is defined as follows:
Table 1. Correlation matrix of the variables.

|          | HHV   | FC     | VM     | ASH    |
|----------|-------|--------|--------|--------|
| HHV      | 1     | -0.55  | -0.59  | -0.53  |
| FC       | 1     | 1      | 1      | 1      |
| VM       | 1     | 1      | 1      | 1      |
| ASH      | 1     | 1      | 1      | 1      |

Table 2. Regression coefficients and MSE.

| Coeff. | $\hat{\beta}_{OLS}$ | $\hat{d}_{R}$ | $\hat{d}_{MRT}$ | $\hat{\delta}_{MSE}$ |
|--------|----------------------|---------------|-----------------|---------------------|
| $\alpha_0$ | 167.7189            | 102.0991      | 8.0166          | 167.7189            |
| $\alpha_1$ | -1.2704             | -0.6143       | -0.6141         | -1.2704             |
| $\alpha_2$ | -1.5311             | -0.8763       | -0.8760         | -1.5311             |
| $\alpha_3$ | -1.6840             | -1.0267       | -1.0262         | -1.6840             |
| Bias    | 0                   | 0.6589        | 0.6591          | 0                   |
| MSE     | 4521.101            | 1675.788      | 1674.967        | 1674.533            |
| $AMSE_{cv}$ | 349.9209           | 5.6562        | 5.6528          | 3.9485              |

The theoretical results are illustrated with real-life data which was analyzed in the study of Qian, Lee, Soto, and Chen (2018). A total of 48 samples of poultry waste were collected from different published open literature reviews to form a database for derivation, evaluation and validation of proximate-based higher heating value (HHV) models. Six samples (#43, 44, 45, 46, 47 and 48) were deleted due to incomplete information. The linear regression model is:

$$HHV = \beta_0 + \beta_1 FC + \beta_2 VM + \beta_3 A + \varepsilon$$  \hspace{1cm} (37)

where HHV denotes Higher Heating Value, FC denotes Fixed Carbon, VM denotes Volatile Matter, A denotes ASH and $\varepsilon$ is the random error term that is expected to be normally distributed. The relationship between the variables were obtained by the correlation matrix as follows.

From Table 1, there is a strong positive relationship between higher heating value and Fixed Carbon while a negative relationship exists between HHV and VM; HHV and Ash. To identify the distribution of the error term, we used the Jarque-Bera (JB) test. The test statistic and the corresponding $p$ value are $JB = 0.6409$ and $p$ value $= 0.7258$, respectively. Since this $p$ value is larger than any reasonable alpha value used in the literature, we conclude that the error term follows the normal distribution. We diagnosed the model for a possible presence of multicollinearity. The variance inflation factor (VIF) values are $VIF_{FC} = 997.819$, $VIF_{VM} = 2163.504$, $VIF_{ASH} = 1533.782$. Literature shows that a model suffers from multicollinearity when $VIF_i > 10$. Since the values of the VIF in the above model is higher than 10, we conclude that the model suffers from severe multicollinearity. Alternatively, we can use the condition number (CN) to examine if the explanatory variables are related where $CN = \text{maximum(eigenvalue)} / \text{minimum(eigenvalue)}$. If CN is between 100 and 1000 there is moderate to strong multicollinearity and if it exceeds 1000 there is severe multicollinearity (Arumairajan & Wijekoon, 2017; Gujarati, 1995). The condition number is 581291.39 which indicates the presence of severe multicollinearity. Therefore, it will be appropriate to predict higher heating value with an alternative unbiased estimator possessing minimum variance. We adopt K fold crossvalidation to validate the performances of the estimators. The data is partitioned into K equal size folds ($K = 10$ in this study). In these $K$ folds, onefold will be treated as the test set and use the remaining $K - 1$ (9) folds as the training set. The MSE is computed on the observations in the held-out fold. The process is repeated ten times, taking out a different part each time. The validation test error is obtained by computing the average $K$ estimates of the test error, and we get an estimated validation (test) error rate for new observations. The estimator with the lowest validation MSE is the best. The average MSE of the validation error in this study is defined as:

$$AMSE_{cv} = \frac{\sum_{k=1}^{10} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{10}$$  \hspace{1cm} (38)

where $n_k$ is the number of subsample in each fold, $\hat{y}_i$ is the fitted value for observation $i$, obtained from the data with fold $k$ removed. The result is presented in Table 2.

The result in Table 2 shows that the unbiased modified ridge-type estimator (UMRT) produced the same estimates with the OLS estimator. Also, the technique was able to circumvent the problem of large variance which is peculiar to the OLS estimator. The proposed estimator has the smallest mean square error and prediction error, respectively.

5.2. Monte–Carlo simulation

We carried out a Monte–Carlo simulation to investigate the performances of these estimators. The explanatory variables were generated in line with the study of McDonald and Galanrneau (1975), Liu...
Table 3. Estimated MSE for OLS, Ridge, MRT and UMRT when \( n = 30 \), Sig = 1 and \( p = 3 \).

|       | \( d = 0.4 \) | \( d = 0.7 \) | \( d = 0.85 \) |
|-------|--------------|--------------|---------------|
| \( y = 0.85 \) | \( \gamma = 0.5 \) \( k = 0.8 \) \( k = 0.95 \) | \( \gamma = 0.5 \) \( k = 0.8 \) \( k = 0.95 \) | \( \gamma = 0.5 \) \( k = 0.8 \) \( k = 0.95 \) |
| OLS   | 0.7149       | 0.7149       | 0.7149        |
| RIDGE | 0.4992       | 0.4280       | 0.4280        |
| MRT   | 0.4853       | 0.3970       | 0.3664        |
| UMRT  | 0.4486       | 0.3767       | 0.3520        |

|       | \( d = 0.4 \) | \( d = 0.7 \) | \( d = 0.85 \) |
|-------|--------------|--------------|---------------|
| \( y = 0.95 \) | \( \gamma = 0.5 \) \( k = 0.8 \) \( k = 0.95 \) | \( \gamma = 0.5 \) \( k = 0.8 \) \( k = 0.95 \) | \( \gamma = 0.5 \) \( k = 0.8 \) \( k = 0.95 \) |
| OLS   | 2.0294       | 2.0294       | 2.0294        |
| RIDGE | 0.8523       | 0.6412       | 0.5758        |
| MRT   | 0.7056       | 0.6173       | 0.5132        |
| UMRT  | 0.6968       | 0.5200       | 0.4689        |

|       | \( d = 0.4 \) | \( d = 0.7 \) | \( d = 0.85 \) |
|-------|--------------|--------------|---------------|
| \( y = 0.99 \) | \( \gamma = 0.5 \) \( k = 0.8 \) \( k = 0.95 \) | \( \gamma = 0.5 \) \( k = 0.8 \) \( k = 0.95 \) | \( \gamma = 0.5 \) \( k = 0.8 \) \( k = 0.95 \) |
| OLS   | 9.5571       | 9.5571       | 9.5571        |
| RIDGE | 1.0061       | 0.6503       | 0.5616        |
| MRT   | 0.7978       | 0.6407       | 0.5298        |
| UMRT  | 0.7333       | 0.4932       | 0.4365        |

Table 4. Estimated MSE for OLS, RE, D and MRT when \( n = 30 \), Sig = 5 and \( p = 3 \).

|       | \( d = 0.4 \) | \( d = 0.7 \) | \( d = 0.85 \) |
|-------|--------------|--------------|---------------|
| \( y = 0.85 \) | \( \gamma = 0.5 \) \( k = 0.8 \) \( k = 0.95 \) | \( \gamma = 0.5 \) \( k = 0.8 \) \( k = 0.95 \) | \( \gamma = 0.5 \) \( k = 0.8 \) \( k = 0.95 \) |
| OLS   | 17.872       | 17.890       | 17.890        |
| RIDGE | 10.123       | 10.119       | 9.303         |
| MRT   | 9.972        | 9.057        | 8.825         |
| UMRT  | 8.517        | 8.510        | 7.684         |

|       | \( d = 0.4 \) | \( d = 0.7 \) | \( d = 0.85 \) |
|-------|--------------|--------------|---------------|
| \( y = 0.95 \) | \( \gamma = 0.5 \) \( k = 0.8 \) \( k = 0.95 \) | \( \gamma = 0.5 \) \( k = 0.8 \) \( k = 0.95 \) | \( \gamma = 0.5 \) \( k = 0.8 \) \( k = 0.95 \) |
| OLS   | 50.749       | 50.749       | 50.749        |
| RIDGE | 20.139       | 14.149       | 12.214        |
| MRT   | 19.573       | 13.922       | 11.414        |
| UMRT  | 15.758       | 10.520       | 8.025         |

|       | \( d = 0.4 \) | \( d = 0.7 \) | \( d = 0.85 \) |
|-------|--------------|--------------|---------------|
| \( y = 0.99 \) | \( \gamma = 0.5 \) \( k = 0.8 \) \( k = 0.95 \) | \( \gamma = 0.5 \) \( k = 0.8 \) \( k = 0.95 \) | \( \gamma = 0.5 \) \( k = 0.8 \) \( k = 0.95 \) |
| OLS   | 238.93       | 238.93       | 238.93        |
| RIDGE | 21.40        | 11.79        | 9.35          |
| MRT   | 16.38        | 11.50        | 8.74          |
| UMRT  | 14.05        | 7.44         | 5.84          |

Table 5. Estimated MSE for OLS, RE, D and MRT when \( n = 50 \), Sig = 1 and \( p = 3 \).

|       | \( d = 0.4 \) | \( d = 0.7 \) | \( d = 0.85 \) |
|-------|--------------|--------------|---------------|
| \( y = 0.85 \) | \( \gamma = 0.5 \) \( k = 0.8 \) \( k = 0.95 \) | \( \gamma = 0.5 \) \( k = 0.8 \) \( k = 0.95 \) | \( \gamma = 0.5 \) \( k = 0.8 \) \( k = 0.95 \) |
| OLS   | 0.3124       | 0.3124       | 0.3124        |
| RIDGE | 0.2617       | 0.2395       | 0.2301        |
| MRT   | 0.2519       | 0.2243       | 0.2299        |
| UMRT  | 0.2463       | 0.2206       | 0.2104        |

|       | \( d = 0.4 \) | \( d = 0.7 \) | \( d = 0.85 \) |
|-------|--------------|--------------|---------------|
| \( y = 0.95 \) | \( \gamma = 0.5 \) \( k = 0.8 \) \( k = 0.95 \) | \( \gamma = 0.5 \) \( k = 0.8 \) \( k = 0.95 \) | \( \gamma = 0.5 \) \( k = 0.8 \) \( k = 0.95 \) |
| OLS   | 0.822        | 0.822        | 0.822         |
| RIDGE | 0.533        | 0.443        | 0.411         |
| MRT   | 0.519        | 0.409        | 0.389         |
| UMRT  | 0.469        | 0.381        | 0.352         |

|       | \( d = 0.4 \) | \( d = 0.7 \) | \( d = 0.85 \) |
|-------|--------------|--------------|---------------|
| \( y = 0.99 \) | \( \gamma = 0.5 \) \( k = 0.8 \) \( k = 0.95 \) | \( \gamma = 0.5 \) \( k = 0.8 \) \( k = 0.95 \) | \( \gamma = 0.5 \) \( k = 0.8 \) \( k = 0.95 \) |
| OLS   | 3.775        | 3.775        | 3.775         |
| RIDGE | 0.973        | 0.567        | 0.593         |
| MRT   | 0.792        | 0.526        | 0.467         |
| UMRT  | 0.750        | 0.524        | 0.465         |
Table 6. Estimated MSE for OLS, RE, D and MRT when \( n = 50 \), sig = 5 and \( p = 3 \).

| Method | \( d = 0.4 \) | \( d = 0.7 \) | \( d = 0.85 \) |
|--------|--------------|--------------|--------------|
| \( y = 0.85 \) | \( k = 0.5 \) | \( k = 0.8 \) | \( k = 0.95 \) |
| OLS    | 7.217        | 7.217        | 7.217        |
| RIDGE  | 6.138        | 5.619        | 5.388        |
| MRT    | 6.112        | 5.610        | 5.268        |
| UMRT   | 5.783        | 5.146        | 4.873        |
| \( y = 0.95 \) | \( k = 0.5 \) | \( k = 0.8 \) | \( k = 0.95 \) |
| OLS    | 19.912       | 19.912       | 19.912       |
| RIDGE  | 12.910       | 10.505       | 9.580        |
| MRT    | 12.327       | 9.613        | 8.359        |
| UMRT   | 11.213       | 8.693        | 7.781        |
| \( y = 0.99 \) | \( k = 0.5 \) | \( k = 0.8 \) | \( k = 0.95 \) |
| OLS    | 93.640       | 93.640       | 93.640       |
| RIDGE  | 22.567       | 14.290       | 11.879       |
| MRT    | 18.875       | 11.675       | 9.043        |
| UMRT   | 16.390       | 9.869        | 8.065        |

Table 7. Estimated MSE for OLS, RE, D and MRT when \( n = 100 \), sig = 1 and \( p = 3 \).

| Method | \( d = 0.4 \) | \( d = 0.7 \) | \( d = 0.85 \) |
|--------|--------------|--------------|--------------|
| \( y = 0.85 \) | \( k = 0.5 \) | \( k = 0.8 \) | \( k = 0.95 \) |
| OLS    | 0.1285       | 0.1285       | 0.1285       |
| RIDGE  | 0.1199       | 0.1151       | 0.1132       |
| MRT    | 0.1253       | 0.1245       | 0.1243       |
| UMRT   | 0.1165       | 0.1112       | 0.1090       |
| \( y = 0.95 \) | \( k = 0.5 \) | \( k = 0.8 \) | \( k = 0.95 \) |
| OLS    | 0.370        | 0.36964      | 0.370        |
| RIDGE  | 0.300        | 0.27206      | 0.261        |
| MRT    | 0.350        | 0.34732      | 0.348        |
| UMRT   | 0.281        | 0.24961      | 0.238        |
| \( y = 0.99 \) | \( k = 0.5 \) | \( k = 0.8 \) | \( k = 0.95 \) |
| OLS    | 1.7988       | 1.7988       | 1.7988       |
| RIDGE  | 0.8014       | 0.6084       | 0.5472       |
| MRT    | 1.5095       | 1.4892       | 1.4873       |
| UMRT   | 0.6598       | 0.4944       | 0.4457       |

Table 8. Estimated MSE for OLS, Ridge, MRT and UMRT when \( n = 100 \), Sig = 5 and \( p = 3 \).

| Method | \( d = 0.4 \) | \( d = 0.7 \) | \( d = 0.85 \) |
|--------|--------------|--------------|--------------|
| \( y = 0.85 \) | \( k = 0.5 \) | \( k = 0.8 \) | \( k = 0.95 \) |
| OLS    | 3.2125       | 3.2125       | 3.2125       |
| RIDGE  | 2.9762       | 2.8490       | 2.7691       |
| MRT    | 3.1210       | 3.0748       | 3.0537       |
| UMRT   | 2.8903       | 2.7239       | 2.6471       |
| \( y = 0.95 \) | \( k = 0.5 \) | \( k = 0.8 \) | \( k = 0.95 \) |
| OLS    | 9.2410       | 9.2410       | 9.2410       |
| RIDGE  | 7.4076       | 6.5890       | 6.2371       |
| MRT    | 8.5726       | 8.3126       | 8.2076       |
| UMRT   | 6.8435       | 5.8766       | 5.4793       |
| \( y = 0.99 \) | \( k = 0.5 \) | \( k = 0.8 \) | \( k = 0.95 \) |
| OLS    | 44.971       | 44.971       | 44.971       |
| RIDGE  | 18.971       | 13.474       | 11.659       |
| MRT    | 36.254       | 34.771       | 34.290       |
| UMRT   | 14.969       | 10.053       | 8.526        |
Table 9. Estimated MSE for OLS, Ridge, MRT and UMRT when \( n = 200 \), Sig = 1 and \( p = 3 \).

| \( \gamma \) | \( d \) | OLS | Ridge | MRT | UMRT |
|---|---|---|---|---|---|
| \( 0.85 \) | 0.4 | 0.0676 | 0.0643 | 0.0668 | 0.0647 |
| | 0.7 | 0.0667 | 0.0638 | 0.0667 | 0.0633 |
| | 0.95 | 0.0676 | 0.0633 | 0.0667 | 0.0639 |
| \( 0.95 \) | 0.4 | 0.1998 | 0.1929 | 0.1927 | 0.1928 |
| | 0.7 | 0.1998 | 0.1927 | 0.1927 | 0.1933 |
| | 0.95 | 0.1998 | 0.1929 | 0.1927 | 0.1933 |
| \( 0.99 \) | 0.4 | 0.9936 | 0.8783 | 0.9369 | 0.9528 |
| | 0.7 | 0.9936 | 0.8773 | 0.9369 | 0.9528 |
| | 0.95 | 0.9936 | 0.8773 | 0.9369 | 0.9528 |

Table 10. Estimated MSE for OLS, Ridge, MRT and UMRT when \( n = 200 \), Sig = 5 and \( p = 3 \).

| \( \gamma \) | \( d \) | OLS | Ridge | MRT | UMRT |
|---|---|---|---|---|---|
| \( 0.85 \) | 0.4 | 1.6888 | 1.5980 | 1.6510 | 1.5642 |
| | 0.7 | 1.6888 | 1.5980 | 1.6510 | 1.5642 |
| | 0.95 | 1.6888 | 1.5980 | 1.6510 | 1.5642 |
| \( 0.95 \) | 0.4 | 4.9937 | 4.2205 | 4.7109 | 3.9668 |
| | 0.7 | 4.9937 | 4.2205 | 4.7109 | 3.9668 |
| | 0.95 | 4.9937 | 4.2205 | 4.7109 | 3.9668 |
| \( 0.99 \) | 0.4 | 24.8400 | 20.5760 | 20.2440 | 19.5890 |
| | 0.7 | 24.8400 | 20.5760 | 20.2440 | 19.5890 |
| | 0.95 | 24.8400 | 20.5760 | 20.2440 | 19.5890 |

Table 11. Estimated MSE for OLS, Ridge, MRT and UMRT when \( n = 30 \), Sig = 1 and \( p = 6 \).

| \( \gamma \) | \( d \) | OLS | Ridge | MRT | UMRT |
|---|---|---|---|---|---|
| \( 0.85 \) | 0.4 | 1.4789 | 0.8110 | 0.7210 | 0.5270 |
| | 0.7 | 1.4789 | 0.8110 | 0.7210 | 0.5270 |
| | 0.95 | 1.4789 | 0.8110 | 0.7210 | 0.5270 |
| \( 0.95 \) | 0.4 | 2.3910 | 1.1470 | 1.0990 | 1.0210 |
| | 0.7 | 2.3910 | 1.1470 | 1.0990 | 1.0210 |
| | 0.95 | 2.3910 | 1.1470 | 1.0990 | 1.0210 |
| \( 0.99 \) | 0.4 | 23.7790 | 1.9090 | 1.0970 | 0.9520 |
| | 0.7 | 23.7790 | 1.9090 | 1.0970 | 0.9520 |
| | 0.95 | 23.7790 | 1.9090 | 1.0970 | 0.9520 |
## Table 12. Estimated MSE for OLS, RE, D and MRT when $n = 30$, sig = 5 and $p = 6$.  

| $d$ = 0.4 | $d$ = 0.7 | $d$ = 0.85 |
|---|---|---|
| $y = 0.85$ | $y = 0.85$ | $y = 0.85$ |
| $k = 0.5$ | $k = 0.5$ | $k = 0.5$ |
| $k = 0.8$ | $k = 0.8$ | $k = 0.8$ |
| $k = 0.95$ | $k = 0.95$ | $k = 0.95$ |
| **OLS** | **OLS** | **OLS** |
| 36.944 | 36.944 | 36.944 |
| 36.944 | 36.944 | 36.944 |
| 36.944 | 36.944 | 36.944 |
| **RIDGE** | **RIDGE** | **RIDGE** |
| 25.322 | 25.322 | 25.322 |
| 25.322 | 25.322 | 25.322 |
| 25.322 | 25.322 | 25.322 |
| **MRT** | **MRT** | **MRT** |
| 23.855 | 23.855 | 23.855 |
| 23.855 | 23.855 | 23.855 |
| 23.855 | 23.855 | 23.855 |
| **UMRT** | **UMRT** | **UMRT** |
| 22.464 | 22.464 | 22.464 |
| 22.464 | 22.464 | 22.464 |
| 22.464 | 22.464 | 22.464 |

## Table 13. Estimated MSE for OLS, RE, D and MRT when $n = 50$, sig = 1 and $p = 6$.  

| $d$ = 0.4 | $d$ = 0.7 | $d$ = 0.85 |
|---|---|---|
| $y = 0.85$ | $y = 0.85$ | $y = 0.85$ |
| $k = 0.5$ | $k = 0.5$ | $k = 0.5$ |
| $k = 0.8$ | $k = 0.8$ | $k = 0.8$ |
| $k = 0.95$ | $k = 0.95$ | $k = 0.95$ |
| **OLS** | **OLS** | **OLS** |
| 0.816 | 0.816 | 0.816 |
| 0.816 | 0.816 | 0.816 |
| 0.816 | 0.816 | 0.816 |
| **RIDGE** | **RIDGE** | **RIDGE** |
| 0.631 | 0.631 | 0.631 |
| 0.631 | 0.631 | 0.631 |
| 0.631 | 0.631 | 0.631 |
| **MRT** | **MRT** | **MRT** |
| 0.624 | 0.624 | 0.624 |
| 0.624 | 0.624 | 0.624 |
| 0.624 | 0.624 | 0.624 |
| **UMRT** | **UMRT** | **UMRT** |
| 0.581 | 0.581 | 0.581 |
| 0.581 | 0.581 | 0.581 |
| 0.581 | 0.581 | 0.581 |

## Table 14. Estimated MSE for OLS, RE, D and MRT when $n = 50$, sig = 5 and $p = 6$.  

| $d$ = 0.4 | $d$ = 0.7 | $d$ = 0.85 |
|---|---|---|
| $y = 0.85$ | $y = 0.85$ | $y = 0.85$ |
| $k = 0.5$ | $k = 0.5$ | $k = 0.5$ |
| $k = 0.8$ | $k = 0.8$ | $k = 0.8$ |
| $k = 0.95$ | $k = 0.95$ | $k = 0.95$ |
| **OLS** | **OLS** | **OLS** |
| 20.393 | 20.393 | 20.393 |
| 20.393 | 20.393 | 20.393 |
| 20.393 | 20.393 | 20.393 |
| **RIDGE** | **RIDGE** | **RIDGE** |
| 15.748 | 15.748 | 15.748 |
| 15.748 | 15.748 | 15.748 |
| 15.748 | 15.748 | 15.748 |
| **MRT** | **MRT** | **MRT** |
| 15.284 | 15.284 | 15.284 |
| 15.284 | 15.284 | 15.284 |
| 15.284 | 15.284 | 15.284 |
| **UMRT** | **UMRT** | **UMRT** |
| 14.466 | 14.466 | 14.466 |
| 14.466 | 14.466 | 14.466 |
| 14.466 | 14.466 | 14.466 |
Table 15. Estimated MSE for OLS, RE, D and MRT when \( n = 100, \text{sig} = 1 \) and \( p = 6 \).

| Method  | \( k = 0.5 \) | \( k = 0.8 \) | \( k = 0.95 \) |
|---------|---------------|---------------|---------------|
| OLS     | 0.27131       | 0.27131       | 0.27131       |
| RIDGE   | 0.25222       | 0.25222       | 0.25222       |
| MRT     | 0.26867       | 0.26867       | 0.26867       |
| UMRT    | 0.2457        | 0.2336        | 0.2283        |

| Method  | \( k = 0.5 \) | \( k = 0.8 \) | \( k = 0.95 \) |
|---------|---------------|---------------|---------------|
| OLS     | 0.27131       | 0.27131       | 0.27131       |
| RIDGE   | 0.25222       | 0.25222       | 0.25222       |
| MRT     | 0.26867       | 0.26867       | 0.26867       |
| UMRT    | 0.2457        | 0.2336        | 0.2283        |

Table 16. Estimated MSE for OLS, Ridge, MRT and UMRT when \( n = 100, \text{Sig} = 5 \) and \( p = 6 \).

| Method  | \( k = 0.5 \) | \( k = 0.8 \) | \( k = 0.95 \) |
|---------|---------------|---------------|---------------|
| OLS     | 6.7828        | 6.7828        | 6.7828        |
| RIDGE   | 6.2695        | 5.9981        | 5.8712        |
| MRT     | 6.6616        | 6.6019        | 6.5750        |
| UMRT    | 6.0858        | 5.7340        | 5.5733        |

| Method  | \( k = 0.5 \) | \( k = 0.8 \) | \( k = 0.95 \) |
|---------|---------------|---------------|---------------|
| OLS     | 6.7828        | 6.7828        | 6.7828        |
| RIDGE   | 6.2695        | 5.9981        | 5.8712        |
| MRT     | 6.6616        | 6.6019        | 6.5750        |
| UMRT    | 6.0858        | 5.7340        | 5.5733        |

Table 17. Estimated MSE for OLS, Ridge, MRT and UMRT when \( n = 200, \text{sig} = 1 \) and \( p = 6 \).

| Method  | \( k = 0.5 \) | \( k = 0.8 \) | \( k = 0.95 \) |
|---------|---------------|---------------|---------------|
| OLS     | 0.1583        | 0.1583        | 0.1583        |
| RIDGE   | 0.1500        | 0.1457        | 0.1437        |
| MRT     | 0.1564        | 0.1554        | 0.1550        |
| UMRT    | 0.1471        | 0.1415        | 0.1390        |

| Method  | \( k = 0.5 \) | \( k = 0.8 \) | \( k = 0.95 \) |
|---------|---------------|---------------|---------------|
| OLS     | 0.1583        | 0.1583        | 0.1583        |
| RIDGE   | 0.1500        | 0.1457        | 0.1437        |
| MRT     | 0.1564        | 0.1554        | 0.1550        |
| UMRT    | 0.1471        | 0.1415        | 0.1390        |
Table 18. Estimated MSE for OLS, Ridge, MRT and UMRT when \( n = 200, \) Sig = 5 and \( p = 6. \)

| \( \gamma \) | \( d = 0.4 \) | \( d = 0.7 \) | \( d = 0.85 \) |
|----------|-----------|------------|-----------|
| \( \kappa = 0.85 \) | \( \kappa = 0.8 \) | \( \kappa = 0.95 \) | \( \kappa = 0.8 \) | \( \kappa = 0.95 \) | \( \kappa = 0.8 \) | \( \kappa = 0.95 \) |
| OLS | 3.9583 | 3.9583 | 3.9583 | 3.9583 | 3.9583 | 3.9583 | 3.9583 |
| RIDGE | 3.7453 | 3.6289 | 3.5736 | 3.7453 | 3.6289 | 3.5736 | 3.7453 | 3.6289 |
| MRT | 3.9083 | 3.8822 | 3.8701 | 3.8987 | 3.8684 | 3.8545 | 3.8940 | 3.8618 |
| UMRT | 3.6668 | 3.5130 | 3.4412 | 3.6102 | 3.4312 | 3.3487 | 3.5827 | 3.3918 |

When \( \gamma = 0.95 \) and \( \kappa = 0.8 \), \( \kappa = 0.85 \) and \( \kappa = 0.95 \), of the estimators.

4. As sample sizes remain constant, increasing the value of \( \sigma \) increases the mean square errors of each of the estimators.

5. As the number of explanatory variables increases, the mean squared error of all the estimators increases.

The response variable is defined as:

\[ y_i = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \epsilon_i \quad (40) \]

where \( \epsilon_i \sim (0, \sigma^2) \). The values of \( \beta \) were chosen such that \( \beta' \beta = 1 \) (Newhouse & Oman, 1971). The sample size used are 30 and 50. Two different values of \( \sigma: 1 \) and 5. The experiment is repeated 1000 times. The estimated MSE is calculated as

\[ \text{MSE}(\hat{\beta}) = \frac{1}{1000} \sum_{j=1}^{1000} (\hat{y}_j - \beta_1)^2 (\hat{y}_j - \beta_1) \quad (41) \]

where \( \hat{y}_j \) denotes the estimate of the \( j \)th parameter in \( j \)th replication and \( \beta_1 \) is the true parameter values. The estimated MSES of the estimators for different values of \( n, k, d, \sigma \) and \( \gamma \) are shown in Tables 3–18. The following observations were made:

1. The unbiased estimator is superior to OLS in all the cases. OLS estimator has the least performance when there is multicollinearity.
2. Also, the unbiased estimator consistently outperforms the ridge and modified ridge estimators. Even though, ridge and modified ridge estimators dominate OLS in all cases.
3. When the sample size increase, the MSE decreases even when the correlation between the explanatory variables increases.

6. Conclusion

The OLS estimator suffers a breakdown in the presence of multicollinearity. The estimator is unbiased but possesses a significant variance. An alternative estimator called unbiased modified ridge-type estimator with prior information was proposed in this study. This estimator was proved to be unbiased and possess minimum variance theoretically. Also, a simulation study and real-life application were conducted to establish the superiority of this estimator over the existing estimators in terms of the MSEM criterion and crossvalidation prediction error. The performance of this new estimator is better than the OLS estimator and ridge estimator for all degree of multicollinearity. This estimator was able to circumvent the problem of inflated variance that faces the OLS estimator. Finally, this estimator should be adopted as a replacement to the OLS estimator and the biased estimators when there is multicollinearity in a linear model.

Acknowledgements

The authors are grateful to the anonymous reviewers for their valuable comments and suggestions, which certainly improved the quality and presentation of this article.
Disclosure statement
No potential conflict of interest was reported by the authors.

References

Arumairajan, S., & Wijekoon, P. (2017). Modified almost unbiased Liu estimator in linear regression model. *Communications in Mathematics and Statistics*, 5, 261–276. doi:10.1007/s40304-017-0111-z

Ayinde, K., Lukman, A. F., Samuel, O. O., & Attah, O. M. (2018). Some new adjusted ridge estimators of linear regression model. *International Journal of Civil Engineering and Technology*, 11, 2838–2852.

Crouse, R. H., Jin, C., & Hanumara, R. C. (1995). Unbiased ridge estimation with prior information and ridge trace. *Communications in Statistics—Theory and Methods*, 24, 2341–2354. doi:10.1080/03610929508831620

Farebrother, R. W. (1976). Further results on the mean square error of ridge regression. *Journal of the Royal Statistical Society: Series B (Methodological)*, 38, 248–250. doi:10.1111/j.2517-6161.1976.tb01588.x

Gujarati, D. N. (1995). *Basic econometrics*. New York, NY: McGraw-Hill.

Hoerl, A. E., & Kennard, R. W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, 12(1), 55–67. doi:10.1080/00401706.1970.10488634

Li, Y., & Yang, H. (2012). A new Liu-type estimator in linear regression model. *Statistical Papers*, 53, 427–437. doi:10.1007/s00362-010-0349-y

Liu, K. (1993). A new class of biased estimator in linear regression. *Communications in Statistics - Theory and Methods*, 22, 393–402.

Lukman, A. F., & Ayinde, K. (2017). Review and classifications of the ridge parameter estimation techniques. *Hacettepe Journal of Mathematics and Statistics*, 46, 953–967. doi:10.15672/HJMS.201815671

Lukman, A. F., Ayinde, K., Binuomote, S., & Onate, A. C. (2019). Modified ridge-type estimator to combat multicollinearity: Application to chemical data. *Journal of Chemometrics*, e3125. 10.1002/cem.3125

McDonald, M. C., & Galanou, D. I. (1975). A Monte Carlo evaluation of some ridge-type estimators. *Journal of the American Statistical Association*, 70, 407–416. doi:10.2307/2285832

Newhouse, J. P., & Oman, S. D. (1971). An evaluation of ridge estimators. *Rand Report*, 1–28. R-716-PR.

Ozkale, M. R., & Kaciranlar, S. (2007). The restricted and unrestricted two-parameter estimators. *Communications in Statistics - Theory and Methods*, 36, 2707–2725.

Qian, X., Lee, S., Soto, A., & Chen, G. (2018). Regression model to predict the higher heating value of poultry waste from proximate analysis. *Resources*, 7, 39. doi:10.3390/resources7030039

Sakallinoğlu, S., & Akdeniz, F. (2003). Unbiased Liu estimation with prior information. *International Journal of Mathematical Sciences*, 2(1), 205–217.

Swindel, F. F. (1976). Good ridge estimators based on prior information. *Communications in Statistics - Theory and Methods*, 11, 1065–1075. doi:10.1080/03610927608827423

Trenkler, G., & Toutenburg, H. (1990). Mean squared error matrix comparisons between biased estimators an overview of recent results. *Statistical Papers*, 31(1), 165–179. doi:10.1007/BF02292467

Wu, J. (2014). An unbiased two-parameter estimation with prior information in linear regression model. *The Scientific World Journal*, 2014, 1–8. doi:10.1155/2014/206943

Wu, J., & Yang, H. (2013). Efficiency of an almost unbiased two-parameter estimator in linear regression model. *Statistics*, 47, 535–545. doi:10.1080/02331882.2011.605891

Yang, H., & Chang, X. (2010). A new two-parameter estimator in linear regression. *Communications in Statistics - Theory and Methods*, 39, 923–934. doi:10.1080/03610920902807911