World Journal of Advanced Research and Reviews
eISSN: 2581-9615 CODEN (USA): WJARAI
Cross Ref DOI: 10.30574/wjarr
Journal homepage: https://wjarr.com/

(Research Article)

Check for updates

Development of hybrid ridge–PCA estimators for addressing Multicollinearity in Gaussian linear regression models

Remilekun Enitan Alabi [1, *], Olatayo Olusegun Alabi [2] and Oluwadare O. Ojo [2]

[1] *Department of Statistics, School of Science and Computer Studies, The Federal Polytechnic Ado-Ekiti, Ekiti state, Nigeria.*
[2] *Department of Statistics, School of Physical Science, The Federal University of Technology, Akure, Ondostate, Nigeria.*

World Journal of Advanced Research and Reviews, 2025, 27(01), 942-957

Publication history: Received on 27 May 2025; revised on 05 July 2025; accepted on 07 July 2025

Article DOI: https://doi.org/10.30574/wjarr.2025.27.1.2559

Abstract

This study tackles the persistent issue of multicollinearity in Gaussian linear regression which undermines the efficiency of Ordinary Least Squares (OLS) estimators. While Ridge Regression and Principal Component Analysis (PCA) are common remedies, they have limitations in terms of bias control and interpretability. To address this, the research proposes hybrid Ridge – PCA estimators using four newly developed ridge parameters combined with PCA. A Monte Carlo simulation evaluated 21 estimators including OLS, Ridge, PCA, and Liu estimators under varying sample sizes, error variances and multicollinearity levels using Mean Squared Error (MSE) as the performance metric. Results show that a newly hybrid estimator consistently outperformed other proposed and existing estimators by achieving the lowest MSE. The study demonstrates the strength of integrating regularization with dimensionality reduction to improve regression under multicollinearity.

Keywords: Multicollinearity; Ridge Regression; Principal Component Estimator; Hybrid Estimators; Monte Carlo Simulation

1. Introduction

The linear regression model is a statistical method that analyzes the relationship between an effect or dependent variable and one or more independent variables helping to explain or predict outcomes (Fayose *et al.,* 2023b; Aladesuyi *et al.,* 2025). The model is simply defined as follows:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots\ldots + \beta_k x_{ik} + \varepsilon_i, i = 1,\ldots\ldots,n, \qquad \ldots\ldots\ldots\ldots \quad (1)$$

where y_i is the effect variable, x_{i1},..., x_{ik} are the concomitant variables, β_0, β_1, ..., β_k are the unknown parameters to be estimated, ε_i denotes the disturbance term and it is assumed to be normally distributed with mean zero and constant variance σ^2. The model is simply a simple regression model when there is one concomitant variable. The parameters in model (1) are mostly estimated by the Method of Least Squares (MLS). MLS is generally preferred and possesses some humbly properties when the assumptions of the linear regression models are intact, this makes the model to be classical (Owolabi *et al.,* 2022 and Dawoud *et al.,* 2022). These include linear relationship among the concomitant variables; the disturbance terms must come from a Gaussian distribution and has non scattered variance and others (Chatterjee and Hadi, 1977). In reality most of the aforementioned assumptions are normally violated. For instance, literature has shown that linear relationship often exists among concomitant variables which are termed multicollinearity (Fayose and Ayinde 2019; Shewa and Ugwuowo, 2022a). Multicollinearity is a phenomenon where

* Corresponding author: Remilekun Enitan Alabi

two or more concomitant variables are highly correlated in a Gaussian linear regression (Neter, *et al.* 1996; Fayose *et al,* 2023a and Aladesuyi *et al.,* 2025). There is tendency for perfect or strong or moderate linear dependency among the concomitant variables (Shewa and Ugwuowo, 2022a; Shewa and Ugwuowo, 2022b). The method of least squares is unbiased but inefficient when there is linear relationship among the concomitant variables (Gujarati *et al.,* 2012). It yields regression coefficients whose absolute values are too large and whose signs may actually reverse with negligible changes in the data (Buonaccorsi, 1996). If the multicollinearity is not perfect but high, the estimated coefficients can become unstable and highly sensitive to slight changes in the model leading to inflated standard errors and misleading inferences (Belseley *et al.,* 1980; Fayose and Ayinde, 2019). Consequently, reliable interpretation of the model parameters may be compromised undermining the credibility of the results derived from such analysis. The pursuit of effective techniques to mitigate the adverse effects of multicollinearity is of paramount importance in both theoretical and applied statistics.

Among the existing methods or approaches proposed to address or handle multicollinearity are Ridge Regression, Principal Component Analysis (PCA) among others. Both Ridge and PCA have emerged as prominent methodologies (Hoerl and Kennard, 1970; Jolliffe, 1986). Ridge regression offers a regularization technique that modifies the Ordinary least squares (OLSE) estimation process by introducing a penalty term (k) to the loss function thus allowing for the shrinkage of coefficient estimates towards zero (Hoerl and Kennard, 1970, Fayose *et al.,* 2023b). The ridge parameter (k) counteracts the inflation of variances associated with multicollinearity effectively enhancing the stability of the estimates and producing more reliable and responsive predictions.

In contrast, PCA serves as a dimensionality reduction technique that transforms the original correlated variables into a set of uncorrelated variables often called principal components (Jolliffe, 2002). By focusing on the principal components that explain the most variance in the data, PCA can help circumvent multicollinearity problem by ensuring that regression model utilizes orthogonal predictors. While both Ridge regression and PCA have demonstrated utility in addressing multicollinearity, each method presents notable limitations. Ridge regression while effective in controlling for multicollinearity does not completely eliminate correlation among predictors, it merely diminishes the variability of the coefficient estimates. Moreover, the choice of penalty parameter (k) can significantly influence the model's performance necessitating careful cross – validation (Tikhonov, 1963). Meanwhile, PCA while adept at reducing dimensionality and addressing multicollinearity transforms the original predictors into a new set of components that may lack interpretability in the context of the original variables posing challenges for practical application and insight derivation (Jolliffe, 1986).

To harness the strengths of both Ridge regression and PCA while mitigating their respective limitations, the concept of hybrid or combined estimators has been introduced by different authors in recent literature. These combined estimators integrate Ridge parameter (k) with PCA estimator to form or create a more robust framework or new hybrid estimator to tackle multicollinearity in Gaussian linear regression model. Wang *et al.,* (2013) proposed a hybrid estimator that combines Ridge parameter (k) with PCA to improve parameter estimation in high dimensional settings. Their approach demonstrated a significant reduction in mean squared error (MSE) compared to standard methods. Other authors that have utilized these combined approaches are Zou and Hastie (2005), Buhlmann and Van de Geer (2011); Chang and Yang (2012); Zhang and Li (2015); Huang and Wang (2018) and Lukman, *et al.,* (2020) among others. The paper intends to comprehensively propose new ridge parameter k's and combine them with PCA to form new hybrid estimators to resolve the problem of multicollinearity within the Gaussian linear regression framework through simulation studies. We compared the estimators' performance to that of some existing techniques. Section 2 contains the methodology. Section 3 presents the simulation design, while Section 4 presents simulation results and while Section 5 is the conclusion.

1.1. Existing and Proposed Estimators

The matrix form of a linear regression model is defined as:

$$y = X\beta + e \qquad \text{.................(2)}$$

Where y is the vector of response variables, X is $n \times p$ design matrix of concomitant variables or predictor β is $p \times 1$ true vector or regression coefficients $e \sim N\left(0, \sigma^2\right)$ is the disturbance which is normally distributed with mean 0 and variance σ^2.

1.2. The Ridge Estimator

The Ordinary Least Square estimator is defined as: $\hat{\beta}_{OLS} = S^{-1} X^{I} Y$ ………… (3)

where

$$S = (X^{I} X) \quad \text{…………} \quad (4)$$

The ridge regression estimator is the mostly used estimators in literature for handling multicollinearity problem. The Generalized Ridge Estimator is defined as:

$$\beta_{GRRE} = (S + kI)^{-1} X^{I} Y \quad \text{…………} \quad (5)$$

where S is a p x p product matrix of concomitant variables, $X^{1}Y$ is a p x 1 vector of the product of effect and concomitant variables, k = diagonal ($k_1, k_2, \ldots, k_p$), $k_i \geq 0$, i = 1, 2,..., p. k is a non – negative constant called biasing or ridge parameter. When k = 0, equation (5) returns to OLS estimator (Fayose and Ayinde, 2019; Kibria and Lukman, 2020; Fayose *et al.,* 2023a). In this paper, we considered these selected 'k' parameters: Hoerl and Kennard (1975), Fayose and Ayinde (2019), Kibria and Lukman (2020), Chand and Kibria (2024a), Chand and Kibria (2024b) and also proposed four new ridge parameters respectively.

2. Principal Component Estimator

The study considered PCA method to also handle multicollinearity in the model and also combined both ridge parameter and PCA method to form hybrid estimator to handle multicollinearity in the model.

PCA transforms the original predictors x into a new set of uncorrelated variables called principal components.

$$\beta_{PCR} = V(V'SV)^{-1} V'X^{1}Y \quad \text{…………} \quad (6)$$

'S' is defined in equation (4).

Let the covariance matrix of X be:

$C = X'X$ and the eigen value decomposition of C gives $C = VDV'$ ………… (7)

where V is $p \times p$ matrix of eigen vectors (Principal Components) and D = diag$\left(\lambda_1, \lambda_2 ., \ldots, \lambda_p\right)$ is the diagonal matrix of the eigen values.

The data matrix X is transformed into principal components $Z = XV$.

where Z is the new transformed data matrix of uncorrelated principal components

By regressing y on the principal components Z instead of X; is given as:

$$y = Z\gamma + e \quad \text{…………} \quad (8)$$

Where $\gamma = V'\beta$

The ordinary least square (OLS) estimator for γ in this model is given as;

$$\hat{\gamma} = \left(Z'Z\right)^{-1} Z'Y \quad \text{…………} \quad (9)$$

By substituting $Z = XV$, therefore the PCA estimator of β denoted as $\hat{\beta}_{PCA}$ is defined as:

$$\hat{\beta}_{PCA} = V\hat{\gamma} = V\left(Z'Z\right)^{-1}Z'y \qquad \text{............ (10)}$$

Where $Z'Z = V'X'XV = D$

Therefore the PCA estimator in equation (10) can further be defined as:

$$\hat{\beta}_{PCA} = V\left(D\right)^{-1}V'X'y \qquad \text{............ (11)}$$

2.1. Some Alternative Ridge Estimators to OLSE

The ridge estimator is defined as:

$$\hat{\beta}_{RIDGE} = (Z'Z + kI)^{-1}Z'y) \qquad \text{............ (12)}$$

where *k* is the non – negative tuning parameter. Different means of deriving *k* exists in the literature. These include:

Following (Hoerl and Kennard, 1975), *k* is given by:

$$KGRHK_{(median)} = \hat{k}_i^M(HK) = median(\frac{\sigma^2}{\hat{\alpha}_i^2}) \text{, i = 1, 2, 3, p} \qquad \text{............ (13)}$$

where $\hat{\sigma}^2 = \frac{\sum_{i=1}^{n} e_i^2}{n-p}$ is the mean square error from the MLS, α_i is the i[th] element of the vector, and is the regression coefficient from the MLS.

Following (Fayose and Ayinde, 2019), *k* is given by:

$$KGRFA = \hat{k}_i^{Min}(FA) = \frac{\hat{\sigma}^2}{\hat{\alpha}_i^2}\left\{\left[\left(\frac{\hat{\alpha}_i^4\lambda_{Min}^2}{4\hat{\sigma}^2}\right) + \left(\frac{6\hat{\alpha}_i^4\lambda_{Min}}{\hat{\sigma}^2}\right)\right]^{\frac{1}{2}} - \left(\frac{\hat{\alpha}_i^2\lambda_{Min}}{2\hat{\sigma}^2}\right)\right\} \qquad \text{............ (14)}$$

where $\lambda_{Min} = Min(\lambda_i) = 1,2,3,\ldots, p.$

Following (Kibria and Lukman, 2020), *k* is given by:

$$KGRKL = \hat{k}_i^{Min}(KL) = \min\left(\frac{\hat{\sigma}^2}{2\hat{\alpha}_i^2 + \left(\hat{\sigma}/\lambda_i\right)}\right) \qquad \text{............ (15)}$$

Following (Chand and Kibria, 2024), *k* is given by:

$$KGRCK_1 = \hat{k}_i(CK_1) = \hat{\sigma}p^{(1+p/n)} \qquad \text{............ (16)}$$

Following (Chand and Kibria, 2024), *k* is given by:

$$KGRCK_2 = \hat{k}_i(CK_2) = \hat{\sigma}\max(p^{(1+p/n)}, p^{(1+1/p)}) \qquad \text{...........} \qquad (17)$$

2.1.1. The Liu Estimator

The Liu estimator proposed by Liu (1993) combined the Stein estimator with Ordinary Ridge Regression estimator to handle multicollinearity. The Liu estimator of β is defined below as

$$\hat{\beta}_L = (X^I X + I)^{-I}(X^I X + dI)\hat{\beta}_{OLS} \qquad 0 < d < 1. \qquad \text{...........} \ (18)$$

$$\text{Where d} = \min\left[\frac{\hat{\alpha}^2}{\left(\hat{\sigma}^2/\lambda_i\right) + \hat{\alpha}_i^2}\right] \qquad \text{...........} \ (19)$$

Where d = ($diag(d_i)$ and is a diagonal matrix of the biasing parameter. The Liu estimator can return to OLS when d = 1.

2.2. Proposed Ridge Estimators

For the ridge parameter whose estimators are defined in (13), (14), (15), (16) and (17), the concept of different forms by Lukman and Ayinde (2017) and Fayose and Ayinde (2019) was introduced based on minimum (MI), maximum (MA) and Median (MD) of eigen values (λ_i) of X^IX of the design matrix of the regression model.

Consequently, in this paper, we proposed some new ridge parameters whose estimators are defined below:

RIDGE ESTIMATOR (PROPOSED 1)

$$\hat{k}_i(CK2) = \hat{\sigma}\min(p^{(1+p/n)}, p^{(1+1/p)}) \qquad \text{...........} \qquad (20)$$

i.e. The minimum version of Chand and Kibria (2024)

RIDGE ESTIMATOR (PROPOSED 2)

$$\hat{k}^{HM}_{KL_{(PROP)}} = p\sum_{i=1}^{p}\left[\frac{\sigma^2}{2\alpha_i^2 + \sigma/\lambda_i}\right] \qquad \text{...........} \qquad (21)$$

i.e. the Harmonic mean version of Kibria and Lukman (2020)

RIDGE ESTIMATOR (PROPOSED 3)

$$\hat{k}^{Median}_{KL_{(PROP)}} = median\left[\frac{\sigma^2}{2\alpha_i^2 + \sigma/\lambda_i}\right] \qquad \text{...........} \qquad (22)$$

i.e. the median version of Kibria and Lukman (2020)

RIDGE ESTIMATOR (PROPOSED 4

$$\hat{k}_{KL_{(PROP)}}^{FM} = \left[\frac{\sigma^2}{2\max(\alpha_i^2) + \sigma / \max(\lambda_i)} \right] \quad \ldots\ldots\ldots\ldots \quad (23)$$

Fixed maximum of Kibria and Lukman (2020)

2.3. Derivation of the Properties of PCA with Ridge Estimator

Ayinde *et al*. (2020) derived a new approach of Principal Component Analysis (PCA) estimator as an alternative to:

$$\hat{\beta}_{PCA} = VD^{-1}V'X'Y \quad \ldots\ldots\ldots\ldots \quad (24)$$

This is defined as

$$\hat{\beta}_{PCA} = (X'X)^{-1}X'\hat{y}_r \quad \ldots\ldots\ldots\ldots \quad (25)$$

Where $\hat{y}_r$ is the predicted variable by regressing y on the r-principal component defined as

$$\hat{y}_r = Z_r\left(Z_r'Z_r\right)^{-1} X'y \quad \ldots\ldots\ldots\ldots \quad (26)$$

Such that $Z_r = XT_r$

where T_r is the r – principal component and T is the orthogonal matrix. Therefore, combination of PCA with ridge estimator is defined as:

$$\hat{\boldsymbol{\beta}}_{R-PCA} = (\boldsymbol{X}'\boldsymbol{X} + \boldsymbol{kI})^{-1}\boldsymbol{X}'\boldsymbol{Z}_r(\boldsymbol{Z}_r'\boldsymbol{Z}_r)^{-1}\boldsymbol{X}'\boldsymbol{y} \quad \ldots\ldots\ldots\ldots \quad (27)$$

Where k is the biasing parameter for individual biasing parameter of Hoerl and Kennard (1975), Fayose and Ayinde (2019), Kibra and Lukman (2020), Chand and Kibra (2024a) and Chand and Kibra (2024b)

2.3.1. Properties of $\hat{\beta}_{R-PCA}$

Mean of the $\hat{\beta}_{R-PCA}$

To compute the mean of $\hat{\beta}_{R-PCA}$, the expected value of equation (3) is taking

$$E(\hat{\beta}_{R-PCA}) = E[(X'X + kI)^{-1}X'Z_r(Z_r'Z_r)^{-1}X'y] \quad \ldots\ldots\ldots\ldots \quad (28)$$

$y = X\beta + e$

$$= E(\hat{\beta}_{R-PCA}) = E\left((X'X + k\mathrm{I})^{-1} X'Z_r\left(Z_r'Z_r\right)^{-1} X'(X\beta + e) \right).$$

$$= E\left((X'X + k\mathrm{I})^{-1} X'Z_r\left(Z_r'Z_r\right)^{-1} X'X\beta + (X'X + k\mathrm{I})^{-1} X'Z_r\left(Z_r'Z_r\right)^{-1} X'e \right)$$

$$= (X'X + k\mathrm{I})^{-1} X'Z_r\left(Z_r'Z_r\right)^{-1} X'XE(\beta) + (X'X + k\mathrm{I})^{-1} X'Z_r\left(Z_r'Z_r\right)^{-1} X'E(e)$$

E(e) = 0 and $E(\beta)=\beta$

Therefore

$$(E(\hat{\beta}_{R-PCA}) = (X'X+kI)^{-1}X'Z_r(Z_r{}'Z_r)^{-1}X'.X\beta \quad \ldots\ldots\ldots\ldots \quad (29)$$

Variance of the $\hat{\beta}_{R-PCA}$

$$\text{Var}\,(\hat{\beta}_{R-PCA}) = E\left[\left(\hat{\beta}_{R-PCA}-E(\hat{\beta}_{R-PCA})\right)\left(\hat{\beta}_{R-PCA}-E(\hat{\beta}_{R-PCA})\right)'\right] \quad \ldots\ldots\ldots\ldots \quad (30)$$

$$=E\begin{bmatrix}[(X'X+kI)^{-1}X'Z_r(Z_r{}'Z_r)^{-1}X'y-(X'X+kI)^{-1}X'Z_r(Z_r{}'Z_r)^{-1}X'X\beta]\\ [(X'X+kI)^{-1}X'Z_r(Z_r{}'Z_r)^{-1}X'y-(X'X+kI)^{-1}X'Z_r(Z_r{}'Z_r)^{-1}X'X\beta]'.\end{bmatrix} \quad \ldots\ldots\ldots\ldots \quad (31)$$

Further expansion of (30) gives equation (31)

$$Var(\hat{\beta}_{R-PCA}) = \sigma^2 \; (X'X+k\mathrm{I})^{-2}(X'Z_r)^2\left(Z_r{}'Z_r\right)^{-2}X'.X \quad \ldots\ldots\ldots\ldots \quad (32)$$

Bias of the $\hat{\beta}_{R-PCA}$

$$\text{Bias}(\hat{\beta}_{R-PCA}) = E\left(\hat{\beta}_{R-PCA}\right)-\beta$$

$$=(X'X+k\mathrm{I})^{-1}X'Z_r\left(Z_r{}'Z_r\right)^{-1}X'X\beta-\beta.$$

$$=((X'X+kI)^{-1}X'Z_r(Z_r{}'Z_r)^{-1}X'X-I)\beta \quad \ldots\ldots\ldots\ldots \quad (33)$$

Mean Squared Error Matrix of $\hat{\beta}_{R-PCA}$

$$\text{MSEM}\,(\hat{\beta}_{R-PCA}) = Var\left(\hat{\beta}_{R-PCA}\right)+\text{Bias}^2(\hat{\beta}_{R-PCA}) \quad \ldots\ldots\ldots\ldots \quad (34)$$

$$\text{MSEM}(\hat{\beta}_{R-PCA}) = \sigma^2(X'X+kI)^{-2}(X'Z_r)^2(Z_r'Z_r)^{-1}X'X+[(X'X+kI)^{-1}X'Z_r(Z_r'Z_r)^{-1}X'X-I]^2\beta^2 \quad (\quad \ldots\ldots\ldots\ldots \quad (35)$$

Also equation (35) can still be written as:

$$MSEM(\hat{\beta}_{R-PCA}) = \sigma^2(X'X+kI)^{-2}(X'Z_r)'(X'Z_r)(Z_r'Z_r)^{-1}X'X+[(X'X+kI)^{-1}X'Z_r(Z_r'Z_r)^{-1}X'X-I]'\,[(X'X+kI)^{-1}X'Z_r(Z_r'Z_r)^{-1}X'X-I]\beta'\beta \quad \ldots\ldots\ldots\ldots \quad (36)$$

Recall the general form of linear regression model in matrix form as given in (2)

$$y = X\beta + e$$

Therefore the canonical form of equation (2) can be written as:

$$y = WX + e \quad \ldots\ldots\ldots\ldots \quad (37)$$

Where $W = XQ, \alpha = Q'\beta$ and Q is the orthogonal matrix whose columns indicate the eigen vectors of the design matrix $X'X$. Hence, $Q'X'XQ = \wedge = diag(\lambda_1, \lambda_2, ..., \lambda_p)$ where $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_p > 0$ are the ordered eigen values $X'X$. The OLS estimator in equation (2) can be defined as:

$$\hat{\alpha}_{ols} = \Lambda^{-1} X'y \qquad \text{...........} \quad (38)$$

Thus the canonical form of equation (26) is

$$\hat{\alpha}_{R-PCA} = (\wedge + k\text{I})^{-1} X'Z_r \left(Z_r{}' Z_r \right)^{-1} X'y \qquad \text{...........} \quad (39)$$

For the convenience of establishing the statistical properties of $\hat{\alpha}_{R-PCA}$, the following Lemmas will be useful.

Lemma I: Let F be a positive definite matrix such that F > 0 and let α be some vectors then;

$F - \alpha\alpha' \geq 0$ if and only if $\alpha' F^{-1} \alpha \leq 1$. (Trenkler and Tontenbury,1990)

Lemma II : Let $\hat{\alpha}_j = A_{ij}$ for j=1,2 be two competing estimators of α.. Also suppose that

$D = Cov(\hat{\alpha}_1) - Cov(\hat{\alpha}_2) > 0$ where $Cov(\hat{\alpha}_1) and Cov(\hat{\alpha}_2)$ are the covariance matrix of $\hat{\alpha}_1 and \hat{\alpha}_2$

Therefore $D = MSEM(\hat{\alpha}_1) - MSEM(\hat{\alpha}_2) \geq 0$ if and only if $a_2'\left[\sigma^2 D + a_1 a_1^2\right]^{-1} a_2 < 1$ where $MSEM(\hat{\alpha}_1)$ = $Cov(\hat{\alpha}_1)$ + $a_i a_i^{'}$

Such that a_i =Bias$(\hat{\alpha}_i) = (AiX - I)\alpha$

2.3.2. The Superiority of the Proposed Estimator in the Sense of MSEM Criterion

The proposed estimator is compared with some already existing estimators such as OLS, Ridge estimators in the sense of MSEM.

Comparison between $\hat{\alpha}_{ols}$ and $\hat{\alpha}_{R-PCA}$

Recall the MSEM of the OLS estimator as

$\hat{\alpha}_{ols} = \Lambda^{-1} X'y$ as:

$$\text{MSEM}(\hat{\alpha}_{ols}) = \sigma^2 \Lambda^{-1} \qquad \text{............} \quad (40)$$

Equation (35) can be written as:

$$\text{MSEM}(\hat{\alpha}_{R-PCA}) = \sigma^2 N^{-2} N_r' N_r \Lambda_r^{-1} \Lambda + \left[N^{-1} N_r \Lambda_r^{-1} \Lambda - I\right]\left[N^{-1} N_r \Lambda_r^{-1} \Lambda - I\right] \alpha^1 \alpha \quad (41)$$

where N= $(\Lambda + kI)$, N_r' = $X'Z_r$, $\Lambda_r = Z_r^{'} Z_r$, $\Lambda = X'X$

Therefore the difference between (39) and (40) is given as

MSEM($\hat{\alpha}_{ols}$)-MSEM($\hat{\alpha}_{R-PCA}$) = $D_{\alpha_{ols}}^{\alpha_{R-PCA}}$

= $\sigma^2\Lambda^{-1}$ - $\sigma^2 N^{-2}N_r'N_r\Lambda_r^{-1}\Lambda + \left[N^{-1}N_r\Lambda_r^{-1}\Lambda - I\right]^1\left[N^{-1}N_r\Lambda_r^{-1}\Lambda - I\right]\alpha^1\alpha$

= $\sigma^2\Lambda^{-1}$ - $\sigma^2 N^{-2}N_r'N_r\Lambda_r^{-1}\Lambda + \left[N^{-1}N_r\Lambda_r^{-1}\Lambda - I\right]\alpha'\alpha\left[N^{-1}N_r\Lambda_r^{-1}\Lambda - I\right]$

$$= \sigma^2\left[\Lambda^{-1} - N^{-2}N_r'N_r\Lambda_r^{-1}\Lambda\right] + \left[N^{-1}N_r\Lambda_r^{-1}\Lambda - I\right]^1\alpha'\alpha\left[N^{-1}N_r\Lambda_r^{-1}\Lambda - I\right] \quad \text{............} \quad (42)$$

Where $k>0$ is the individual biasing parameter of Hoerl and Kennard (1970), Fayose and Ayinde (2019), Kibria and Lukman (2020) and Chand and Kibra (2024).

The proposed estimator $\hat{\alpha}_{R-PCA}$ is superior to $\hat{\alpha}_{ols}$ if and only if $\alpha'\left[N^{-1}N_r\Lambda_r^{-1}\Lambda - I\right]\sigma^2\left[\Lambda^{-1} - N^{-2}N_r^1N_r\Lambda_r^{-1}\Lambda\right]^{-1}\left[N^{-1}N_r\Lambda_r^{-1}\Lambda - I\right]\alpha < 1$

Proof: By considering the dispersion matrix difference

$$D_{\alpha_{ols}}^{\alpha_{R-PCA}} = \sigma^2\left[\Lambda^{-1} - N^{-2}N_r^1N_r\Lambda_r^{-1}\Lambda\right] \quad \text{............} \quad (43)$$

=trace ($D_{\alpha_{ols}}^{\alpha_{R-PCA}}$)

$= \sum_{i=1}^{p} diag\,(D_{\alpha OLS}^{\alpha R-PCA})$

$= \sigma^2 \sum_{i=1}^{p} diag\,\left[\Lambda^{-1} - N^{-2}N_r^1N_r\Lambda_r^{-1}\Lambda\right]$

$$= \sigma^2 \sum_{i=1}^{p} diag\left[\frac{1}{\lambda_i} - \frac{n_r^2\lambda_i}{(\lambda_i + k)^2\lambda_{ir}}\right]_{i=1}^{p} \quad \text{............} \quad (44)$$

Where λ_i is the diag (X'X), n_r =diag $(X'Z_r)$ *and* λ_{ir} =diag $(Z'Z_r)$

The difference will be positive definite if and only if $(\lambda_{i+}k)^2\lambda_{ir}$ - $n_{ir}^2\lambda_i^{\,2}$ > 0. It can be observed that

$(\lambda_i + k)^2\lambda_{ir}$ - $n_{ir}^2\lambda_i^{\,2}$ will be greater than zero if $k > 0$. Hence by Lemma II the proof is completed

Comparison between $\hat{\alpha}_{RE}$ *and* $\hat{\alpha}_{R-PCA}$

The bias vector covariance matrix and MSEM of $\hat{\alpha}_{RE}$ estimator defined as

$$\hat{\alpha}_{RE} = (\Lambda + kI)^{-1}W'y \quad \text{............} \quad (45)$$

$$E(\hat{\alpha}_{RE}) = (\Lambda + kI)^{-1}\Lambda\alpha \quad \text{............ (46)}$$

$$\text{Var}(\hat{\alpha}_{RE}) = \sigma^2 (\Lambda + kI)^{-1}\Lambda\alpha \quad \text{............ (47)}$$

$$\text{MSEM}(\hat{\alpha}_{RE}) = \sigma^2(\Lambda + kI)^{-1}\Lambda(+kI)^{-1} + k^2(\Lambda + kI)^{-1}\alpha\alpha'(\Lambda + kI)^{-1} \quad \text{............ (48)}$$

$$\text{MSEM}(\hat{\alpha}_{RE}) = \sigma^2 N^{-1}\Lambda N^{-1} + k^2 N^{-1}\alpha\alpha^1 N^{-1} \quad \text{............ (49)}$$

Therefore the difference between (48) and (49)

MSEM ($\hat{\alpha}_{RE}$) – MSEM ($\hat{\alpha}_{R-PCA}$)

$$\sigma^2 N^{-1}\Lambda N^{-1} - \sigma^2 N^{-2}N_r^1 N_r \Lambda_r^{-1}\Lambda + k^2 N^{-1}\alpha\alpha^1 N^{-1} - [N^{-1}N_r\Lambda_r^{-1}\Lambda - I]^1[N^{-1}N_r\Lambda_r^{-1}\Lambda - I]\alpha'\alpha$$
(............ (50)

Let k>0, the estimator $\hat{\alpha}_{R-PCA}$ is superior to $\hat{\alpha}_{RE}$ if and only if{ MSEM ($\hat{\alpha}_{RE}$) – MSEM ($\hat{\alpha}_{R-PCA}$)}>0,if and only if =
$\sigma^2(N^{-2}\Lambda - N^{-2}N_r^1 N_r\Lambda_r^{-1}\Lambda) + k^2 N^{-1}\alpha\alpha^1 N^{-1} - [N^{-1}N_r\Lambda_r^{-1}\Lambda - I]^1[N^{-1}N_r\Lambda_r^{-1}\Lambda - I]\alpha'\alpha < 1$

Proof

Considering the dispersion matrix difference between $\hat{\alpha}_{RE}$ and $\hat{\alpha}_{R-PCA}$

$$D_{RE}^{\alpha_{R-PCA}} = \sigma^2(N^{-2}\Lambda - N^{-2}N_r^1 N_r\Lambda_r^{-1}\Lambda) \quad \text{............ (51)}$$

$$= \sigma^2(\Lambda + kI)^{-2}\Lambda - \sigma^2(\Lambda + kI)^{-2}(X'Z_r)^2\left(Z_r'Z_r\right)^{-1}\Lambda$$

$$= \sigma^2(\Lambda + kI)^{-2}\left[\Lambda - (X'Z_r)^2\left(Z_r'Z_r\right)^{-1}\Lambda\right]$$

$$= \sigma^2 \text{ diag } \left[\frac{\lambda_i}{(\lambda_i + k)^2} - \frac{n_r^2\lambda_i}{(\lambda_i + k)^2\lambda_{ir}}\right]_{i=1}^{p} \quad \text{............ (52)}$$

$D_{RE}^{\alpha_{R-PCA}}$ will be pdf if and only if $\lambda_i\lambda_{ir} - n_{ir}^2 > 0$ for k > 0. Hence it can be observed that $\lambda_i\lambda_{ir} - n_{ir}^2 > 0$, therefore by Lemma II the proof is completed

2.3.3. Determination of Biasing Parameter 'k'.

Finding an appropriate ridge shrinkage or biasing parameter has been the bone of contention in the study of ridge regression. This is because the parameter may either be non – stochastic or may depend on the observed or real – life data set. Therefore, the shrinkage parameter k employed in this study are stated in equation (13), (14), (15), (16) and (17) as well as the proposed ones (20) to (23) respectively.

For practical purpose the MSE of $\hat{\alpha}_{R-PCA}$ can be written as

$$\text{MSE}(\hat{\alpha}_{R-PCA}) = \text{trace}\{\text{MSEM}(\hat{\alpha}_{R-PCA})\} \quad \ldots\ldots\ldots\ldots (53)$$

$$= \sum_{I=1}^{P} diag\ (\text{MSEM}(\hat{\alpha}_{R-PCA})) \quad \ldots\ldots\ldots\ldots (54)$$

$$= \sum_{i=1}^{P} diag\left[\sigma^2 N^{-2} N_1 N_r \Lambda_r^{-1}\Lambda + \left[N^{-2}N_r^2\Lambda_r^{-2}\Lambda^2 - 2N^{-1}N_r\Lambda_r^{-1}\Lambda + I\right]\hat{\alpha}_i^2\right]$$

$$\text{MSE}(\hat{\alpha}_{R-PCA}) = \sigma^2 \sum_{i=1}^{p}\left[\frac{n_{ir}{}^2\lambda_i}{(\lambda_i+k)^2\lambda_{ir}}\right] + \sum_{i=1}^{p}\left[\frac{n_{ir}{}^2\lambda^2{}_i}{(\lambda_i+k)^2\lambda^2{}_{ir}} - \frac{n_{ir}\lambda_i}{(\lambda_i+k)\lambda_{ir}} + 1\right]\hat{\alpha}_i^2$$

$$\text{MSE}(\hat{\alpha}_{R-PCA}) = \sigma^2 \sum_{i=1}^{p}\left[\frac{n_{ir}{}^2\lambda_i}{(\lambda_i+k)^2\lambda_{ir}}\right] + \sum_{i=1}^{p}\left[\frac{n_{ir}{}^2\lambda^2{}_i - (\lambda_i+k)\lambda_{ir}(n_{ir}\lambda_i + (\lambda_i+k)\lambda_{ir}}{(\lambda_i+k)^2\lambda^2{}_{ir}}\right]\hat{\alpha}_i^2$$

(………… (55)

Where n_{ir} is the eigen value of $N_r = X\ Z_{r,}\lambda_{ir}$ is the eigen value of matrix $Z'_r Z_{r,}\lambda_r$ is the eigen value of the design matrix X^1X, p is the number of the concomitant variables and k is the generalized biasing parameter which is the individual biasing parameter which were earlier defined in (13) to (17), (19) to (23) respectively.

3. Simulation Procedure and Design

A Monte Carlo simulation study is performed in the study to show the performance of the proposed estimator over some existing estimators in literature.

Consider the linear regression of the form:

$$y_t = \beta_0 + \beta_1 X_{t1} + \beta_2 X_{t2} + \ldots + \beta_p X_{tp} + U_t \quad \ldots\ldots\ldots\ldots (56)$$

where t = 1, 2, …, n; p = 3, 6, $U_t \approx N(0, \sigma^2)$, $X_{ti,}$ t = 1,2,…,n; i = 1,2,...,p are fixed concomitant variables. The concomitant variables are generated using the following procedure (Lukman and Ayinde 2017; Fayose and Ayinde, 2019; Fayose *et al.,* 2023a):

$$X_{ti} = (1-\rho^2)^{\frac{1}{2}} Z_{ti} + \rho Z_{tp} \quad \ldots\ldots\ldots\ldots (57)$$

where Z_{ti} is independent standard normal distribution with mean zero and constant variance, ρ is the correlation between any two concomitant variables and p is the number of concomitant variables. The error terms U_t were generated to be normally distributed with mean zero and variance σ^2. U_t is the error term. The study used Monte Carlo simulation to conduct the experiment with varying parameters such as sample sizes (n = 10, 20, 30, 50, 100 and 250); level of multicollinearity (ρ = 0, 0.8, 0.9, 0.95, 0.99, 0.999). In the study, σ^2 values were 1, 25 and 100. $E(y_i)$ = expected value of the regression under consideration. $y_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + u_i$ for p = 3 and $y_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 + \beta_6 x_6 + u_i$ for p = 6. When p = 3; β_0 = 0.1550494, β_1 = 0.6162552, β_2 = 0.5311015, β_3 = 0.5604644. When p = 6; β_0 = 0.09867566, β_1 = 0.47436618, β_2 = 0.35373615, β_3 = 0.41280825, β_4 = 0.32653190,

β_5 = 0.40968387, β_6 = 0.44185515. The experiment was repeated 1000 times (number of replication). The performances of the estimators were compared using the Mean Square Error criterion. For any estimator $\hat{\beta}$, MSE is defined as follows:

$$MSE(\hat{\beta}) = \frac{1}{1000}\sum_{i=1}^{p}\sum_{j=1}^{1000}\left(\hat{\beta}_{ij} - \beta_i\right)^2 \qquad \text{........... (58)}$$

where $\hat{\beta}_{ij}$ is i^{th} element of the estimator β in the j^{th} replication which gives the estimate of β_i. β_i are the true value of the parameter previously mentioned. Estimator with the minimum MSE was considered best. The statistical package R Studio was used to write the program that accommodated Twenty – One (21) estimators (OLS, PCA estimator, Generalized Ridge estimators, Liu estimator, Ridge – PCA estimators). Out of the Twenty – One (21) estimators, thirteen (13) are proposed estimators, eight (8) are existing estimators. At a particular level of error variance, multicollinearity and sample size, R studio package gave MSE values. These were recorded 180 times (Multicollinearity levels x Error Variance x Sample Sizes x Number of Regressors types = 5 x 3 x 6 x 2) accordingly. Statistical Package for the Social Sciences (SPSS 25.0) was further used to rank the estimators on the basis of their MSE values. Estimators with high MSE were sorted and removed using SPSS software. The MSE obtained by each estimator was ranked for each degree of multicollinearity and error variance. The degrees of multicollinearity and error variance were tallied to determine the number of times each estimator had the smallest MSE (rank 1 and 2). An estimator is optimal or most efficient if it has the most counts; the mode.

4. Results and Discussion

We represented the outcomes visually and in Tables.

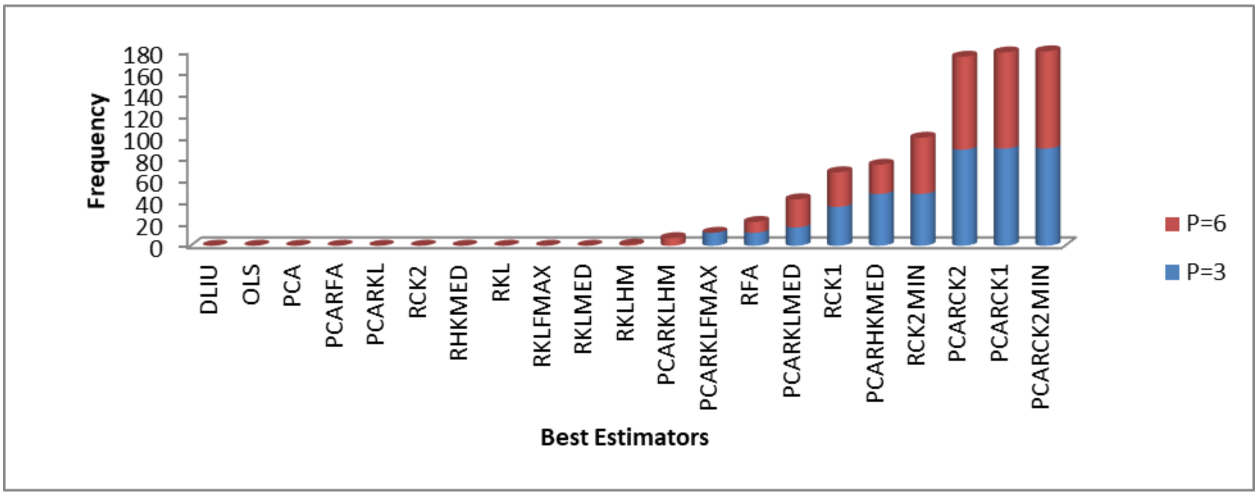

Figure 1 Component Bar Chart showing frequency of counts at which MSE is minimum at p = 3 and p = 6 for OLS and PCA estimation methods

Figure 1 shows that PCARCK2MIN is the best (most efficient) estimator when dealing with the model's multicollinearity problem. i.e. proposed one – parameter ridge estimator (minimum version of Chand and Kibria 2 (2024) with Principal Component estimator) followed by combined estimator nicknamed (PCARCK1) i.e. proposed one – parameter ridge estimator with PCA (Chand and Kibria 1 (2024) with Principal Component estimator) and combined estimator nicknamed (PCARCK2) i.e. proposed one – parameter ridge estimator with PCA (Chand and Kibria 2 (2024) with Principal Component estimator) respectively. Meanwhile the best estimator without Principal Component estimator is the proposed one – parameter ridge estimator i.e. minimum version of Chand and Kibria 2 (2024) followed by existing one – parameter ridge estimator of Chand and Kibria 1 (2024) and existing one – parameter ridge estimator of Fayose and Ayinde (2019) respectively.

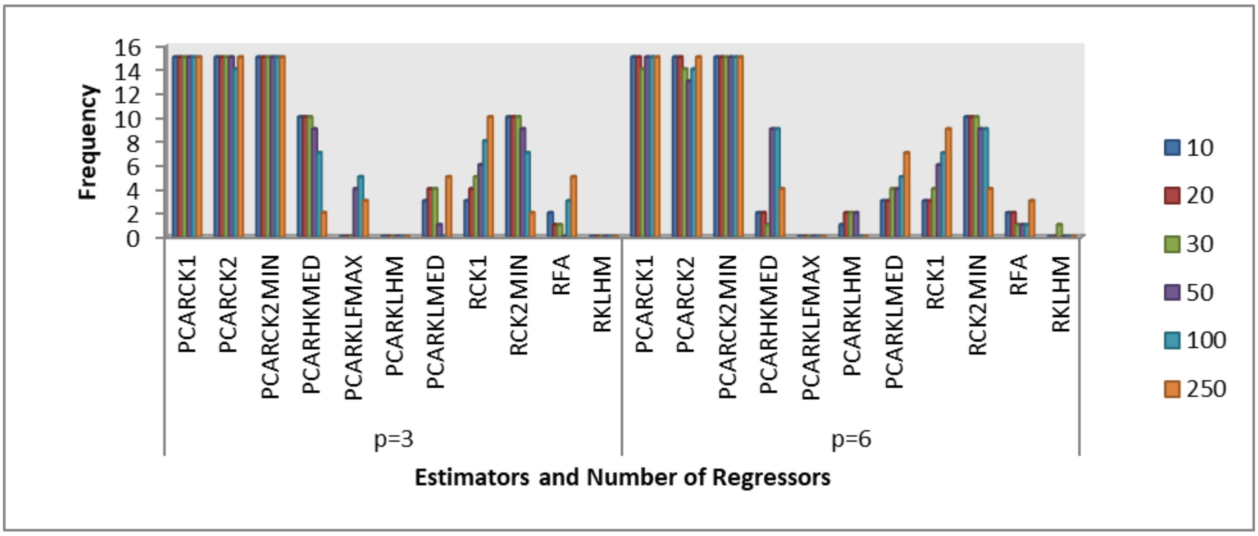

Figure 2 Multiple Bar chart showing performance of the estimators at different sample sizes when p = 3 and p = 6.

From Figure 2, proposed (PCARCK2MIN) i.e. minimum version of Chand and Kibra 2 (2024) with Principal Component estimator performed efficiently across all sample sizes at p = 3 and at p = 6 respectively. Similarly, proposed (PCARCK1) i.e. Chand and Kibra 1 (2024) with Principal Component estimator performed efficiently across all sample sizes at p = 3 but when p = 6, it performed across all sample sizes expect at sample size 30.

The simulation results are available on request but for ease of comparison the results are summarized in Table 1.

Table 1 Number of Times Each Estimator Produced Minimum MSE when counted over levels of Multicollinearity and Error Variance

p	Estimators	Sample Size (n)							
		10	20	30	50	100	250	TOTAL	RANK
3	PCARCK1	15	15	15	15	15	15	90	1st
	PCARCK2	15	15	15	15	14	15	89	2nd
	PCARCK2MIN	15	15	15	15	15	15	90	1st
	PCARHKMED	10	10	10	9	7	2	48	3rd
	PCARKLFMAX	0	0	0	4	5	3	12	6th
	PCARKLHM	0	0	0	0	0	0	0	7th
	PCARKLMED	3	4	4	1	0	5	17	5th
	RCK1	3	4	5	6	8	10	36	4th
	RCK2MIN	10	10	10	9	7	2	48	3rd
	RFA	2	1	1	0	3	5	12	6th
	RKLHM	0	0	0	0	0	0	0	7th
6	PCARCK1	15	15	14	15	15	15	89	2nd
	PCARCK2	15	15	14	13	14	15	86	3rd
	PCARCK2MIN	15	15	15	15	15	15	90	1st
	PCARHKMED	2	2	1	9	9	4	27	6th
	PCARKLFMAX	0	0	0	0	0	0	0	11th

	PCARKLHM	1	2	2	2	0	0	7	9^{th}
	PCARKLMED	3	3	4	4	5	7	26	7^{th}
	RCK1	3	3	4	6	7	9	32	5^{th}
	RCK2MIN	10	10	10	9	9	4	52	4^{th}
	RFA	2	2	1	1	1	3	10	8^{th}
	RKLHM	0	0	1	0	0	0	1	10^{th}

NOTE: Estimator with highest frequency at each sample size is bolded at p = 3 and p = 6.

From Table 1, when p = 3, it is observed that the best or most efficient estimator are PCARCK1 i.e. proposed one – parameter ridge estimator with PCA (Chand and Kibria 1 (2024) with Principal Component estimator) and PCACK2MIN i.e. proposed one – parameter ridge estimator with PCA (minimum version of Chand and Kibria 2 (2024) with Principal Component estimator) followed by PCARCK2 i.e. proposed one – parameter ridge estimator with PCA (Chand and Kibria 2 (2024) with Principal Component estimator) respectively. But when p = 6, the most efficient estimator is PCACK2MIN i.e. proposed one – parameter ridge estimator with PCA (minimum version of Chand and Kibria 2 (2024) with Principal Component estimator) followed by PCARCK1 i.e. proposed one – parameter ridge estimator with PCA (Chand and Kibria 1 (2024) with Principal Component estimator) respectively. The top best or most efficient estimators are the proposed ones.

5. Conclusion

This study addressed the multicollinearity problem in Gaussian linear regression by proposing robust hybrid estimators that combine Ridge Regression with Principal Component Analysis (PCA). Four novel ridge parameters were developed and integrated with PCA to form advanced Ridge – PCA estimators. Through extensive Monte Carlo simulations, the proposed estimators especially PCARCK2MIN consistently outperformed traditional methods like OLS, Ridge, PCA, and Liu estimators by achieving the lowest Mean Squared Error (MSE). These results highlight the efficiency and reliability of the hybrid approach. The study demonstrates that combining regularization with dimensionality reduction enhances model performance and offers a strong solution for multicollinear regression analysis.

Compliance with ethical standards

Acknowledgments

The authors wish to express their gratitude to TETFUND for their support in this research.

Disclosure of conflict of interest

The authors declared that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this journal. Finally, the authors agreed that all information supplied here are real and original.

References

[1] Aladesuyi, A., Ayinde, K. and Fayose, T. S. (2025): Assessing the Role of Significant Roots in Parameter Estimation of Linear Regression Models under Multicollinearity. Tech – Sphere Journal of Pure and Applied Sciences (TSJPAS), 2(1), 1 – 16. https://doi:/10.5281/zenodo.15470100.

[2] Belsley, D. A., Kuh, E. and Welsch, R. E. (1980). Regression Diagnostics: Identifying Influential Data and Sources of Collinearity. New York: John Wiley and Sons.

[3] Buhlmann, P. and Van de Geer, S. (2011). Statistics for High – Dimensional Data: Methods, Theory and Applications. Berlin: Springer – Verlag.

[4] Buonaccorsi, J. P. (1996). A Modified Estimating Equation Approach to Correcting for Measurement Error in Regression, Biometrika 83, 433 – 440.

[5] Chang, Y. and Yang, Y. (2012). Analysis of Multicollinearity in Regression Models with an Application to a Real Data Example. Journal of Statistical Theory and Practice, 6(3), 455 – 470.

[6] Chatterjee, S. and Hadi, A. S. (1977). Regression Analysis by Example, Wiley, Hoboken, NJ, USA.

[7] Chatterjee, S., Hadi, A. S. and Price, B. (2000). Regression Analysis by Example. 3rd Edition, A Wiley – Interscience Publication. John Wiley and Sons.

[8] Dawoud, I., Lukman, A. F. and Haadi, A. (2022). A New Biased Regression Estimator: Theory, Simulation and Application. Scientific African, 15, e01100. https://doi.org/10.1016/j.sciaf.2022.e01100.

[9] Fayose, T. S. and Ayinde, K. (2019). Different Forms Biasing Parameter for Generalized Ridge

[10] Regression Estimator. International Journal of Computer Applications, 181, 21 – 29.

[11] Fayose, T. S.; Ayinde K. and Alabi, O. O. (2023a). M Robust Weighted Ridge Estimator in Linear Regression Model. African Scientific Reports, 2(123), 1 – 28.

[12] Fayose, T. S.; Ayinde K.; Alabi, O. O. and Bello, A. H. (2023b). Robust Weighted Ridge Regression Based on S – Estimator. African Scientific Reports, 2(126), 1 – 28.

[13] Gujarati, N. D., Porter, C. D. and Gunasekar, S. (2012). Basic Econometrics. 5th Edition. New Delhi: Tata McGraw – Hill.

[14] Hoerl, A. E. and Kennard, R. W. (1970). Ridge Regression: Biased Estimation for Non – Orthogonal Problems. Technometrics, 12(1), 55 – 67.

[15] Hoerl, A. E. and Kennard, R.W. (1975). Ridge Regression: Iterative Estimation of the Biasing Parameter. Communications in Statistics - Theory and Methods, 4(1), 77 – 88.

[16] Huang, J. and Wang, H. (2018). A Regularized Estimator for High – Dimensional Generalized Linear Models with PCA. Journal of Statistical Computation and Simulation, 88(4), 675 – 686.

[17] Jolliffe, I. T. (1986). Principal Component Analysis. Springer Series in Statistics. New York: Springer – Verlag.

[18] Jolliffe, I. T. (2002). Principal Component Analysis. New York: Springer – Verlag.

[19] Kibria, G. B. M. and Lukman, A. F. (2020). A New Ridge – Type Estimator for the Linear Regression Model: Simulations and Applications. Scientifica. 10. https://doi.org/10.1155/2020/9758378.

[20] Liu, K. (1993). A New Class of Biased Estimate in Linear Regression. Communications in Statistics – Theory and Methods. 22, 393 – 402, MR1212418.

[21] Lukman, A. F. Lukman and Ayinde, K. (2017). Review and Classifications of the Ridge Parameter Estimation Techniques. Hacettepe Journal of Mathematics and Statistics, 46, 953 – 967.

[22] Lukman, A. F., Kibria, G. B. M., Ayinde, K. and Jegede, S. L. (2020). Modified One – Parameter Liu Estimator for the Linear Regression Model. Modelling and Simulation in Engineering, 10, 9574304. https://doi.org/10.1155/2020/9574304.

[23] Lukman, A. F., Ayinde, K., Kibria, G. B. M. and Jegede, S. L. (2020). Two – Parameter Modified Ridge – Type M – Estimator for Linear Regression Model. The Scientific World Journal. 10, https://doi.org/10.1155/2020/3192852.

[24] Lukman, A. F., Dawoud, I. and Haadi, A. R. (2022). A New Biased Regression Estimator: Theory, Simulation and Application. Scientific African. 15, 2468 – 2276. Doi: 10.1016/j.sciaf.2022.e01100.

[25] Neter, J., Wasserman, W. and Kutner, M. H. (1996). Applied Linear Statistical Models. Chicago: Irwin.

[26] Owolabi, A. T., Ayinde, K., Idowu, J. I., Oladapo, O. J. and Lukman, A. F. (2022). New Two – Parameter Estimator in the Linear Regression Model with Correlated Concomitant Variables. J. Stat. Appl. Pro., 11, 499 – 512.

[27] Shewa, G. A. and Ugwuowo, F. I. (2022a). Kibria – Lukman type Estimator for Gamma Regression Model. Concurrency Computat Pract Exper. 10, e7741. https://doi.org/10.1002/cpe.7441

[28] Shewa, G. A. and Ugwuowo, F. I. (2022b). Combating the Multicollinearity in Bell Regression Model: Simulation and Application. Journal of the Nigerian Society of Physical Sciences, 4(3), 713. https://doi.org/10.46481/jnsps.2022.713.

[29] Tikhonov, A. N. (1963). Solution of III – Posed Problems and the Regularization Method. Soviet Mathematics Doklady, 4, 1624 – 1627.

[30] Trenkler, G. and Toutenburg, H. (1990). Mean Squared Error Matrix Comparisons between Biased Estimators – An Overview of Recent Results. Stat Papers. 31(1), 165 – 179.

[31] Wang, Y., Li, R. and Zhang, H. H. (2013). A hybrid Ridge – PCA Estimator in High – Dimensonal Linear Models. Computational Statistics and Data Analysis, 57(1), 36 – 46.

[32] Zhang, L. and Li, Y. (2015). Nonlinear Modeling with Principal Component Analysis and Penalized Regression Techniques. Journal of Computational and Graphical Statistics, 24(2), 372 – 392.

[33] Zou, H. and Hastie, T. (2005). Regularization and Variable Selection via the Elastic Net. Journal of the Royal Statistical Society: Series B (Statistical Methodology). 67(2), 301 – 320.