IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received April 17, 2021, accepted April 30, 2021, date of publication May 4, 2021, date of current version May 13, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3077385

On the Performance of Jackknife Based Estimators for Ridge Regression

ISMAIL SHAH[1], FAIZA SAJID[1], SAJID ALI[1], AMJAD REHMAN[2], (Senior Member, IEEE), SAEED ALI BAHAJ[3,4], AND SULIMAN MOHAMED FATI[5], (Senior Member, IEEE)

[1]Department of Statistics, Quaid-i-Azam University, Islamabad 45320, Pakistan
[2]Artificial Intelligence and Data Analytics Laboratory, CCIS, Prince Sultan University, Riyadh 11586, Saudi Arabia
[3]MIS Department, College of Business Administration, Prince Sattam Bin Abdulaziz University, Alkharj 11942, Saudi Arabia
[4]Department of Computer Engineering, Hadhramout University, Al Mukalla, Yemen
[5]College of Computer and Information Sciences, Prince Sultan University, Riyadh 11586, Saudi Arabia

Corresponding author: Sajid Ali (sajidali.qau@hotmail.com)

This work was supported in part by the Artificial Intelligence and Data Analytics Laboratory, CCIS, Prince Sultan University.

ABSTRACT Regression techniques are generally used to predict a response variable using one or more predictor variables. In many fields of study, the regressors can be highly intercorrelated, which leads to the problem of multicollinearity. Consequently, the ordinary least squares estimates become inconsistent and lead to wrong inferences. To handle the problem, machine learning techniques particularly, the ridge regression approach, are commonly used. In this paper, we revisit the problem of estimating the ridge parameter "k" by proposing some new estimators using the Jackknife method and compare them with some existing estimators. The performance of the proposed estimators compared to the existing ones is evaluated using extensive Monte Carlo simulations as well as two real data sets. The results suggested that the proposed estimators outperform the existing estimators.

INDEX TERMS Ridge regression, multicollinearity, Monte Carlo simulations, mean squared error, Jackknife technique, machine learning.

I. INTRODUCTION

The primary goal of the regression analysis is to predict the response variable with the help of one or more predictor variables. In many fields of study, including medical sciences, engineering, economics, and social sciences, the predictor variables are highly intercorrelated and in such situations, the ordinary least squares (OLS) estimators are inconsistent. Consequently, the OLS estimators have very large standard errors and thus, lead to wrong inferences. To cope with this problem, machine learning techniques are widely used. Within these techniques, ridge regression is a well-known regression method that can handle multicollinearity and issues due to the high dimensionality of the data [1]. In ridge regression, the main idea is to introduce biased estimators in order to decrease the overall variance. The ridge parameter "k", also called the tuning parameter, is used to control the trade-off between bias and variance. To define ridge regression, consider the following multiple linear regression model.

$$Y = X\boldsymbol{\beta} + \boldsymbol{\epsilon} \tag{1}$$

where $Y = (y_1, y_2, \ldots, y_n)'$ is an $n \times 1$ vector of response variable, $X = \begin{bmatrix} x_{11} & x_{12} & \ldots & x_{1p} \\ x_{21} & x_{22} & \ldots & x_{2p} \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \ldots & x_{np} \end{bmatrix}$ is an $n \times p$ matrix of the observed regressors, $\boldsymbol{\beta} = (\boldsymbol{\beta_1}, \boldsymbol{\beta_2}, \ldots, \boldsymbol{\beta_p})'$ is an $p \times 1$ vector of unknown regression parameters, and $\boldsymbol{\epsilon} = (\epsilon_1, \epsilon_2, \ldots, \epsilon_n)'$ is an $n \times 1$ vector of random errors which are normally distributed with mean vector 0 and covariance matrix $\sigma^2 \boldsymbol{I_n}$ with $\boldsymbol{I_n}$ is an identity matrix of order n, where n is the number of rows and p is the number of columns of the design matrix. The OLS estimator of the regression coefficients $\boldsymbol{\beta}$ is obtained as

$$\hat{\boldsymbol{\beta}} = (X'X)^{-1}XY$$

and

$$\text{Cov}(\hat{\boldsymbol{\beta}}) = \sigma^2 (X'X)^{-1}$$

The associate editor coordinating the review of this manuscript and approving it for publication was Wen-Sheng Zhao.

Note that the OLS estimator and its covariance matrix heavily depend on the characteristics of the matrix $X'X$. Since the OLS estimators are the best linear unbiased estimator, that is, have the smallest mean squared error among the set of unbiased estimators, these are generally preferred. However, if the matrix $X'X$ is ill-conditioned, it indicates that there exists a multicollinearity problem. In such cases, the OLS estimators are inconsistent and have large variances. In addition, some of the regression coefficients may be statistically insignificant or have the wrong sign and thus can be misleading [2].

To overcome the problem of multicollinearity, different researchers suggested different methods, and one of the most useful methods is the ridge regression method. The literature indicates that an abundance of works has been done in this direction [3]. Ridge regression was first proposed by Hoerl and Kennard [1] to solve the multicollinearity problem. The main idea was to add a small positive number to the diagonal elements of the $X'X$ matrix. The resulting estimator of $\boldsymbol{\beta}$ is obtained as

$$\hat{\boldsymbol{\beta}} = (X'X + kI_p)^{-1}X'Y, \quad k \geq 0 \tag{2}$$

where I_p represents an identity matrix of order p, and k is the ridge parameter. The above estimator is known as a ridge regression estimator where k plays an important role in the tradeoff between the consistency and bias of the estimator. When $k \rightarrow \infty$, $\hat{\boldsymbol{\beta}} \rightarrow 0$, i.e., we obtain a stable but biased estimator of $\boldsymbol{\beta}$. On the other hand, when $k \rightarrow 0$, $\hat{\boldsymbol{\beta}} \rightarrow OLS$ and we obtain an unbiased but unstable estimator of $\boldsymbol{\beta}$. Furthermore, for a positive value of k, this estimator provides a smaller mean squared error (MSE) compared to the OLS [2]. As the ridge estimator is heavily dependent on the unknown value of k, the optimum value for k that can produce the best results to some extent is still an open problem in the literature. The estimator $\hat{\boldsymbol{\beta}}_{\boldsymbol{k}}$ is a complicated function of k and several authors presented their proposals for the estimation of k [3]–[25].

Due to the importance of the problem, this article aims to revisit the estimation problem of ridge parameter k and to propose some new and efficient estimators. The proposed methods will be compared with the existing methods using a Monte Carlo simulation study and real data sets. The remaining article unfolds as follows. Section 2 contains the review of some well known existing ridge estimators as well as our proposed estimators for estimating the ridge parameter k. A Monte Carlo simulation study is given in Section 3. Section 4 provides real data applications to assess the proposals' performance, whereas some concluding remarks are discussed in Section 5.

II. SOME EXISTING AND NEW ESTIMATORS

This section reviews some existing ridge estimators and proposes some new estimators. To understand how the previous estimators are build, suppose that there exists an orthogonal matrix Ψ such that $\Psi'\Omega\Psi = \Theta$, where $\Theta = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_p)$ contain the eigenvalues of the matrix $\Omega = X'X$. Consequently, the modified form of equation 1 can be written as

$$Y = X^{*}\boldsymbol{\alpha}\ \mathcal{C}\ \boldsymbol{\epsilon} \tag{3}$$

where

$$X^{*} = X\Psi \quad \textbf{and } \boldsymbol{\alpha} = \Psi'\boldsymbol{\beta}$$

In the case of strong multicollinearity, some of the eigenvalues of the matrix Ω tend to zero. To overcome this issue, a small quantity is added to the diagonal of the matrix $X'X$, i.e., $X'X + kI_p$ $(k > 0)$, which is the same as replacing the λ_i by $\lambda_i + k$, and it accommodates the estimator for the strength of the linear multicollinearity. Consequently, the generalized ridge regression estimator can be written as

$$\hat{\boldsymbol{\alpha}}(k) = (X^{*}X + KI_p)^{-1}X^{*}Y = (I_p + K(X^{*}X))^{-1}\hat{\boldsymbol{\alpha}} \tag{4}$$

where $K = \text{diag}(k_1, k_2, \ldots, k_p)\, k_i \geq 0$, and $\hat{\boldsymbol{\alpha}} = \Omega^{-1}X^{*}Y$ is the OLS estimator of $\boldsymbol{\alpha}$. [1] stated the value of k_i that minimizes the mean squared error (MSE) can be obtained by

$$k_i = \frac{\sigma^2}{\alpha_i} \tag{5}$$

where σ^2 denotes the variance of the regression residuals and α_i represents the ith element of the vector $\boldsymbol{\alpha}$. Different authors suggested different estimation techniques to estimate the optimal value of $\hat{k}_i$ and some of the widely used estimators are described below.

A. HOERL AND KENNARD ESTIMATOR

The pioneering work of Hoerl and Kennard [1] suggested to replace the variance σ^2 of the OLS estimator and regression coefficients α_i^2 by their corresponding unbiased estimators $\hat{\sigma}^2$ and $\hat{\alpha}_i^2$, respectively, i.e.,

$$\hat{k}_i = \frac{\hat{\sigma}^2}{\hat{\alpha}_i^2}$$

where the residual mean square error $\hat{\sigma}^2 = \frac{\sum \hat{\epsilon}_i^2}{n-p} = \frac{(y-y^2)'(y-y^2)}{n-p}$ is an unbiased estimator of σ^2, and k is defined as

$$k_{\text{HK}} = \hat{k}_{\text{HK}} = \frac{\hat{\sigma}^2}{\hat{\alpha}_{max}^2}$$

where $\hat{\alpha}_{max}$ denotes the maximum element of the vector $\hat{\alpha}$. Note that when σ^2 and α are known, $\hat{k}_{HK}$ results in a smaller MSE than the OLS.

B. KIBRIA ESTIMATOR

By generalizing the idea of Hoerl and Kennard [1], Kibria [9] proposed some new estimators based on the geometric mean, arithmetic mean and median of $\hat{k}_i$. The resulted estimators are

$$k_{K1} = \hat{k}_{GM} = \frac{\hat{\sigma}^2}{(\prod_{i=1}^{p} \hat{\alpha}_i^2)^{1/p}}$$

$$k_{K2} = \hat{k}_{AM} = \frac{1}{p}\sum_{i=1}^{p} \frac{\hat{\sigma}^2}{\hat{\alpha}_i^2}$$

$$k_{K3} = k_{MED} = Median\left[\frac{\hat{\sigma}^2}{\hat{\alpha}_i^2}\right], \quad i = 1, 2, \ldots, p, \text{ for } p \geq 3$$

C. SHUKUR ESTIMATOR

Realizing the importance of the matrix X′X, Shukur *et al.* [12] proposed a new method for estimating the ridge parameter k. The suggested estimator is given by

$$k_S = \hat{k}_{KS} = \frac{t_{max}\hat{\sigma}^2}{(n-p)\hat{\sigma}^2 + t_{max}\hat{\alpha}_{max}^2}$$

where t_{max} is the maximum eigenvalue of X′X matrix.

D. ALKHAMISI ESTIMATOR

In the context of generalized ridge regression approach [10], Alkhamisi and Shukur [26] generalizes the idea of Shukur *et al.* [12] by applying the geometric mean, median and maximum value approaches to estimate the k using the following equations.

$$k_{S2} = \hat{k}_{gm}^{KS} = (\prod_{i=1}^{p} \frac{t_i\hat{\sigma}^2}{(n-p)\hat{\sigma}^2 + t_i\hat{\alpha}_i^2})^{1/p}$$

$$k_{S3} = \hat{k}_{max}^{KS} = \max(\frac{t_i\hat{\sigma}^2}{(n-p)\hat{\sigma}^2 + t_i\hat{\alpha}_i^2})$$

$$k_{S4} = \hat{k}_{md}^{KS} = \text{median}(\frac{t_i\hat{\sigma}^2}{(n-p)\hat{\sigma}^2 + t_i\hat{\alpha}_i^2})$$

E. MUNIZ AND KIBRIA ESTIMATOR

Defining $m_i = \sqrt{\frac{\hat{\sigma}_i^2}{\hat{\alpha}_i^2}}$, [2] proposed the following estimators based on the square root transformations.

$$KM2 = \max(\frac{1}{m_i})$$

$$KM3 = \max(m_i)$$

$$KM4 = (\prod_{i=1}^{p} \frac{1}{m_i})^{1/p}$$

$$KM5 = (\prod_{i=1}^{p} m_i)^{1/p}$$

$$KM6 = \text{median}(\frac{1}{m_i})$$

$$KM7 = \text{median}(m_i)$$

F. PROPOSED ESTIMATORS

Many researchers suggested principle component (PC) regression technique to replace $\hat{\boldsymbol{\beta}}$ with $\hat{\boldsymbol{\alpha}}$, which is a linear combination of uncorrelated variables. However, this method suffers from the fact that each PC is itself a linear combination of all the original variables, and thus it is often difficult to interpret the results. This article proposed an alternative method for the estimation of "k" by modifying some existing estimators by replacing $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\alpha}}$ with jackknife estimators. Jackknife is a useful method for estimating and compensating an estimator's bias and does not require knowledge of the theoretical form of standard error [27].

G. JACKKNIFE ALGORITHM

Jackknife (JK) is a computer-based method for estimating biases and standard errors [28]. Quenouille [29] proposed the idea of Jackknife for bias estimation. Later, Tukey [30] recognized the potential of jack-knife for estimating the standard errors. Further developments were presented by [17], [31]–[38]. In the linear regression settings, Shao and Wu [39] and Shao [40] presented general theoretical results on the deleted-d JK algorithm which is a commonly used method nowadays.

To understand the basic idea of the deleted-d Jackknife approach, suppose we have a vector $W_i = (Y_i, Z_{ij})'$ that contains observed values w_i for $i = 1, 2, \ldots, n$. where $Y_i = (y_1, y_2, \cdots, y_n)'$ contains the responses, and $Z_{ij} = (x_{j1}, x_{j2}, \ldots, x_{jn})$ is a matrix of dimension $n \times k$, where $j = 1, 2, \ldots, k$, $i = 1, 2, 3, \cdots, n$. Draw a random sample of size n from the population and label the elements $w_1, w_2, \ldots, w_n$. To compute the deleted-d Jackknife estimator, we proceed as follows.

Step 1: Divide the sample into ' s ' independent group of size d.

Step 2: Leave the first d observations from the sample at a time and estimate the OLS coefficients $\hat{\boldsymbol{\theta}}_{j1}$ using the n-d observation.

Step 3: Delete the second d observation set from the sample and compute again the OLS coefficients $\hat{\boldsymbol{\theta}}_{j2}$ from the remaining n-d observations.

Step 4: Repeating the above, delete each time d observations out of the n and estimate the OLS coefficients $\hat{\boldsymbol{\theta}}_{jk}$, where $\hat{\boldsymbol{\theta}}_{jk}$ denotes the kth JK regression coefficient after deleting of kth d observation set from the sample. Thus, the total number of delete-d JK sample are $S = \binom{n}{d}$.

Step 5: Finally, calculate the JK regression estimator as follows:

$$\hat{\theta}_J = \frac{\sum_{k=1}^{s} \hat{\boldsymbol{\theta}}_{jk}}{S}.$$

That is, the JK regression estimator $\hat{\theta}_J$ is the mean of the deleted d-JK estimates $\hat{\boldsymbol{\theta}}_{j1}, \hat{\boldsymbol{\theta}}_{j2}, \cdots, \hat{\boldsymbol{\theta}}_{j_s}$ [41].

H. PROPOSED ESTIMATORS

Using the delete-d Jackknife estimator, the first proposed estimator is given by

$$\text{IFS}_1 = \frac{\hat{\sigma}^2}{\min \hat{\boldsymbol{\theta}}_i^2}$$

The second proposed estimator is

$$\text{IFS}_2 = \frac{\hat{\sigma}^2}{(\prod_{i=1}^{p} \hat{\boldsymbol{\theta}}_i^2)^{1/p}}$$

The third proposed estimator is

$$\text{IFS}_3 = \frac{1}{p}\sum_{i=1}^{p} \frac{\hat{\sigma}^2}{\hat{\boldsymbol{\theta}}_i^2}$$

TABLE 1. Estimated MSE for p = 4 (superscript reports the MSE rank).

ρ	σ	OLS	HK	KGM	KAM	KMED	KSM	IFS_1	IFS_2	IFS_3	IFS_4
						n=25					
0.90	**1**	1.365^{10}	0.780^{9}	0.262^{5}	0.29^{6}	0.396^{7}	0.134^{3}	0.416^{8}	0.196^{4}	0.056^{1}	0.072^{2}
	3	12.043^{10}	6.280^{9}	1.331^{6}	2.004^{7}	2.507^{8}	1.083^{5}	0.409^{3}	0.752^{4}	0.192^{1}	0.197^{2}
	5	34.822^{10}	18.688^{9}	2.903^{5}	4.587^{6}	6.115^{8}	4.894^{7}	0.397^{3}	0.897^{4}	0.329^{1}	0.331^{2}
0.95	**1**	2.834^{10}	1.511^{9}	0.422^{6}	0.389^{5}	0.535^{7}	0.171^{3}	0.545^{8}	0.222^{4}	0.047^{1}	0.049^{2}
	3	25.044^{10}	12.972^{9}	2.132^{5}	2.478^{6}	3.200^{8}	2.812^{7}	0.387^{3}	0.776^{4}	0.184^{2}	0.163^{1}
	5	72.411^{10}	38.623^{9}	4.615^{5}	5.31^{6}	7.910^{7}	13.426^{8}	0.365^{3}	0.908^{4}	0.322^{2}	0.293^{1}
0.99	**1**	15.438^{10}	7.978^{9}	1.385^{7}	0.615^{5}	0.789^{6}	1.451^{8}	0.605^{4}	0.316^{3}	0.051^{1}	0.052^{2}
	3	136.800^{10}	70.411^{9}	6.693^{7}	2.751^{5}	5.363^{6}	31.204^{8}	0.236^{3}	0.838^{4}	0.222^{1}	0.225^{2}
	5	395.366^{10}	209.352^{9}	13.949^{6}	4.913^{5}	14.154^{7}	133.823^{8}	0.288^{1}	0.936^{4}	0.371^{2}	0.374^{3}
						n = 50					
0.90	**1**	0.510^{8}	0.368^{7}	0.139^{2}	0.185^{3}	0.243^{5}	0.134^{1}	0.238^{4}	0.966^{10}	0.658^{9}	0.312^{6}
	3	4.549^{10}	2.446^{9}	0.770^{4}	1.475^{7}	1.844^{8}	0.359^{2}	0.261^{1}	0.996^{6}	0.861^{5}	0.625^{3}
	5	12.911^{10}	6.963^{9}	1.635^{6}	3.678^{7}	4.566^{8}	0.594^{2}	0.232^{1}	0.999^{5}	0.913^{4}	0.744^{3}
0.95	**1**	1.066^{10}	0.652^{7}	0.224^{2}	0.27^{3}	0.379^{6}	0.105^{1}	0.352^{5}	0.968^{9}	0.657^{8}	0.320^{4}
	3	9.500^{10}	5.036^{9}	1.245^{6}	2.05^{7}	2.536^{8}	0.355^{2}	0.267^{1}	0.996^{5}	0.860^{4}	0.633^{3}
	5	26.971^{10}	14.444^{9}	2.611^{6}	4.783^{7}	6.238^{8}	1.205^{5}	0.213^{1}	0.999^{4}	0.912^{3}	0.751^{2}
0.99	**1**	5.887^{10}	3.149^{9}	0.755^{7}	0.519^{4}	0.706^{6}	0.133^{1}	0.482^{3}	0.977^{8}	0.669^{5}	0.374^{2}
	3	52.421^{10}	27.531^{9}	3.960^{7}	2.945^{5}	4.378^{8}	2.951^{6}	0.166^{1}	0.997^{4}	0.866^{3}	0.680^{2}
	5	148.839^{10}	79.068^{9}	7.984^{6}	5.547^{5}	11.680^{7}	15.651^{8}	0.150^{1}	0.999^{4}	0.916^{3}	0.786^{2}
						n = 100					
0.90	**1**	0.264^{10}	0.217^{9}	0.084^{3}	0.12^{4}	0.148^{7}	0.147^{6}	0.163^{8}	0.124^{5}	0.022^{1}	0.047^{2}
	3	2.314^{10}	1.260^{9}	0.437^{4}	0.938^{7}	1.199^{8}	0.456^{5}	0.261^{3}	0.686^{6}	0.104^{1}	0.118^{2}
	5	6.650^{10}	3.592^{9}	1.012^{6}	2.567^{7}	3.071^{8}	0.463^{4}	0.216^{3}	0.865^{5}	0.193^{2}	0.178^{1}
0.95	**1**	0.548^{10}	0.384^{9}	0.134^{4}	0.182^{6}	0.247^{7}	0.121^{3}	0.258^{8}	0.141^{5}	0.028^{1}	0.041^{2}
	3	4.809^{10}	2.526^{9}	0.709^{5}	1.408^{7}	1.790^{8}	0.317^{4}	0.275^{3}	0.711^{6}	0.138^{2}	0.086^{1}
	5	13.814^{10}	7.391^{9}	1.632^{6}	3.693^{7}	4.426^{8}	0.388^{4}	0.198^{2}	0.877^{5}	0.245^{3}	0.131^{1}
0.99	**1**	2.999^{10}	1.624^{9}	0.464^{7}	0.418^{5}	0.587^{8}	0.059^{2}	0.433^{6}	0.213^{4}	0.082^{3}	0.017^{1}
	3	26.298^{10}	13.649^{9}	2.276^{6}	2.584^{7}	3.346^{8}	0.422^{4}	0.173^{2}	0.783^{5}	0.303^{3}	0.042^{1}
	5	75.460^{10}	40.082^{9}	5.115^{6}	5.732^{7}	8.728^{8}	2.315^{5}	0.119^{2}	0.912^{4}	0.449^{3}	0.079^{1}

The fourth proposed estimator is

$$IFS_4 = \text{Median}(\frac{\hat{\sigma}^2}{\hat{\theta}_i^2})$$

III. SIMULATION STUDIES

For each simulation case, all covariates were standardized to have mean 0 and standard deviation 1. Then, the predictors are generated as follows:

$$X_{ij} = (1-\rho^2)^{\frac{1}{2}} Z_{ij} + \rho Z_i$$

where Z_{ij} are pseudo random numbers generated using the standard normal distribution and ρ represents the strong correlation existing between two explanatory variables. These explanatory variables are standardized so that $X'X$ and $X'y$ are strongly related.

TABLE 2. Estimated MSE for p = 8 (superscript reports the MSE rank).

ρ	σ	OLS	HK	KGM	KAM	KMED	KSM	IFS_1	IFS_2	IFS_3	IFS_4
						n=25					
0.90	1	2.909^{10}	1.773^{9}	0.305^{6}	0.397^{7}	0.718^{8}	0.117^{2}	0.090^{1}	0.145^{4}	0.130^{3}	0.147^{5}
	3	25.467^{10}	14.550^{9}	1.962^{5}	3.246^{7}	5.312^{8}	1.979^{6}	0.595^{4}	0.166^{1}	0.246^{2}	0.280^{3}
	5	73.160^{10}	43.014^{9}	4.529^{5}	7.91^{6}	13.718^{8}	9.722^{7}	0.809^{4}	0.361^{2}	0.338^{1}	0.378^{3}
0.95	1	6.058^{10}	3.559^{9}	0.538^{7}	0.534^{6}	1.030^{8}	0.251^{5}	0.082^{1}	0.097^{4}	0.084^{2}	0.085^{3}
	3	53.248^{10}	30.249^{9}	3.499^{5}	4.248^{6}	7.133^{8}	5.628^{7}	0.582^{4}	0.152^{2}	0.151^{1}	0.152^{2}
	5	153.040^{10}	89.503^{9}	8.101^{5}	9.979^{6}	18.912^{7}	26.923^{8}	0.801^{4}	0.361^{3}	0.222^{1}	0.224^{2}
0.99	1	33.181^{10}	19.194^{9}	2.162^{7}	0.923^{6}	1.731^{8}	3.043^{5}	0.077^{1}	0.028^{4}	0.02^{2}	0.026^{3}
	3	293.421^{10}	164.982^{9}	13.885^{5}	6.094^{6}	13.723^{8}	62.796^{7}	0.575^{4}	0.165^{2}	0.074^{1}	0.073^{2}
	5	843.586^{10}	488.936^{9}	31.912^{5}	12.326^{6}	37.401^{7}	270.543^{8}	0.797^{4}	0.400^{3}	0.145^{1}	0.138^{2}
						n = 50					
0.90	1	1.295^{10}	0.900^{9}	0.157^{6}	0.397^{7}	0.458^{8}	0.066^{5}	0.034^{3}	0.036^{4}	0.027^{1}	0.027^{1}
	3	11.13^{10}	6.477^{9}	1.002^{6}	3.246^{7}	3.736^{8}	0.254^{5}	0.198^{4}	0.185^{3}	0.097^{1}	0.101^{2}
	5	31.528^{10}	18.599^{9}	2.482^{6}	7.91^{7}	9.612^{8}	1.028^{5}	0.467^{4}	0.448^{3}	0.183^{1}	0.190^{2}
0.95	1	2.689^{10}	1.670^{9}	0.276^{6}	0.39^{7}	0.699^{8}	0.052^{5}	0.024^{3}	0.024^{3}	0.022^{2}	0.021^{1}
	3	23.050^{10}	13.321^{9}	1.774^{6}	3.125^{7}	5.225^{8}	0.563^{5}	0.188^{3}	0.191^{4}	0.101^{2}	0.090^{1}
	5	65.352^{10}	38.378^{9}	4.400^{6}	8.384^{7}	13.628^{8}	3.094^{5}	0.456^{3}	0.460^{4}	0.194^{2}	0.176^{1}
0.99	1	14.634^{10}	8.632^{9}	1.095^{6}	0.751^{7}	1.445^{8}	0.307^{5}	0.012^{3}	0.013^{3}	0.037^{2}	0.016^{1}
	3	124.741^{10}	71.440^{9}	6.984^{6}	5.349^{7}	10.271^{8}	8.236^{5}	0.186^{3}	0.232^{4}	0.171^{2}	0.084^{1}
	5	354.607^{10}	206.615^{9}	17.224^{6}	12.823^{7}	28.758^{8}	42.204^{5}	0.455^{3}	0.516^{4}	0.297^{2}	0.169^{1}
						n = 100					
0.90	1	0.598^{10}	0.474^{9}	0.084^{6}	0.173^{7}	0.267^{8}	0.086^{5}	0.023^{4}	0.037^{3}	0.022^{1}	0.030^{2}
	3	5.334^{10}	3.161^{9}	0.549^{6}	1.500^{7}	2.350^{8}	0.248^{5}	0.192^{4}	0.109^{3}	0.067^{1}	0.069^{2}
	5	15.284^{10}	9.100^{9}	1.369^{6}	4.206^{7}	6.255^{8}	0.315^{3}	0.463^{5}	0.318^{4}	0.125^{2}	0.114^{1}
0.95	1	1.252^{10}	0.858^{9}	0.145^{6}	0.259^{7}	0.427^{8}	0.061^{5}	0.017^{2}	0.024^{4}	0.016^{1}	0.021^{3}
	3	11.149^{10}	6.508^{9}	0.976^{6}	2.225^{7}	3.599^{8}	0.173^{4}	0.182^{5}	0.110^{3}	0.066^{2}	0.051^{1}
	5	31.954^{10}	18.910^{9}	2.419^{6}	6.113^{7}	9.237^{8}	0.547^{5}	0.450^{4}	0.327^{3}	0.131^{2}	0.091^{1}
0.99	1	6.894^{10}	4.020^{9}	0.575^{6}	0.57^{7}	1.111^{8}	0.050^{5}	0.009^{2}	0.008^{4}	0.024^{1}	0.008^{3}
	3	61.299^{10}	35.378^{9}	3.892^{6}	4.641^{7}	7.781^{8}	1.386^{4}	0.178^{5}	0.135^{3}	0.126^{2}	0.033^{1}
	5	175.714^{10}	102.904^{9}	9.459^{6}	11.595^{7}	21.050^{8}	8.138^{5}	0.446^{4}	0.377^{3}	0.231^{2}	0.071^{1}

TABLE 3. Estimated MSE for p = 16 (superscript reports the MSE rank).

ρ	σ	OLS	HK	KGM	KAM	KMED	KSM	IFS_1	IFS_2	IFS_3	IFS_4
						n=25					
0.90	1	14.665^{10}	8.536^{9}	0.501^{1}	0.544^{2}	1.258^{7}	0.748^{5}	1.087^{6}	1.528^{8}	0.712^{3}	0.735^{4}
	3	140.494^{10}	82.444^{9}	3.678^{5}	4.572^{6}	9.948^{7}	18.757^{8}	0.739^{1}	1.313^{2}	1.71^{3}	1.784^{4}
	5	383.551^{10}	227.613^{9}	9.884^{5}	12.732^{6}	26.714^{7}	75.115^{8}	0.457^{1}	0.842^{2}	2.296^{3}	2.406^{4}
0.95	1	31.047^{10}	17.836^{9}	0.942^{4}	0.74^{3}	1.823^{7}	2.083^{8}	1.121^{5}	1.616^{6}	0.546^{1}	0.685^{2}
	3	299.311^{10}	173.931^{9}	6.910^{6}	6.078^{5}	13.735^{7}	50.286^{8}	0.511^{1}	0.909^{2}	0.984^{3}	1.312^{4}
	5	814.389^{10}	478.076^{9}	18.539^{6}	16.638^{5}	38.022^{7}	195.093^{8}	0.294^{1}	0.509^{2}	1.190^{3}	1.612^{4}
0.95	1	173.502^{10}	97.654^{9}	4.220^{7}	1.324^{5}	3.412^{6}	22.299^{8}	0.740^{3}	1.096^{4}	0.106^{1}	0.329^{2}
	3	1687.368^{10}	964.936^{9}	30.686^{7}	9.912^{5}	28.175^{6}	479.467^{8}	0.161^{2}	0.253^{3}	0.133^{1}	0.415^{4}
	5	4569.486^{10}	2637.427^{9}	82.068^{7}	24.837^{5}	79.055^{6}	1730.558^{8}	0.116^{1}	0.149^{2}	0.167^{3}	0.467^{4}
						n = 50					
0.90	1	3.226^{10}	2.167^{9}	0.207^{5}	0.414^{6}	0.853^{8}	0.044^{1}	0.048^{2}	0.710^{7}	0.054^{3}	0.204^{4}
	3	28.890^{10}	18.394^{9}	1.540^{6}	3.628^{7}	7.312^{8}	0.849^{5}	0.106^{2}	0.469^{4}	0.088^{1}	0.373^{3}
	5	79.000^{10}	49.976^{9}	3.977^{5}	9.748^{7}	19.23^{8}	4.023^{6}	0.306^{3}	0.262^{2}	0.122^{1}	0.459^{4}
0.95	1	6.764^{10}	4.310^{9}	0.386^{5}	0.579^{6}	1.289^{8}	0.089^{3}	0.031^{2}	0.719^{7}	0.024^{1}	0.158^{4}
	3	60.602^{10}	38.247^{9}	2.863^{6}	4.962^{7}	10.387^{8}	2.623^{5}	0.094^{2}	0.284^{4}	0.047^{1}	0.235^{3}
	5	165.154^{10}	103.465^{9}	7.403^{5}	13.224^{7}	27.839^{8}	12.022^{6}	0.286^{4}	0.151^{2}	0.081^{1}	0.277^{3}
0.99	1	37.310^{10}	23.207^{9}	1.709^{7}	1.093^{5}	2.832^{8}	1.291^{6}	0.010^{1}	0.355^{4}	0.011^{2}	0.052^{3}
	3	334.459^{10}	207.879^{9}	12.551^{6}	8.766^{5}	21.483^{7}	33.692^{8}	0.086^{4}	0.067^{3}	0.059^{1}	0.065^{2}
	5	907.147^{10}	558.822^{9}	32.353^{6}	21.895^{5}	60.743^{7}	140.093^{8}	0.274^{4}	0.053^{1}	0.122^{3}	0.083^{2}
						n = 100					
0.90	1	1.427^{10}	1.073^{9}	0.104^{4}	0.278^{6}	0.518^{8}	0.040^{1}	0.051^{3}	0.510^{7}	0.050^{2}	0.131^{5}
	3	12.817^{10}	8.249^{9}	0.780^{6}	2.466^{7}	4.645^{8}	0.164^{3}	0.069^{1}	0.524^{5}	0.085^{2}	0.257^{4}
	5	35.638^{10}	22.939^{9}	1.972^{6}	6.752^{7}	12.436^{8}	0.605^{5}	0.224^{2}	0.305^{3}	0.114^{1}	0.319^{4}
0.95	1	3.007^{10}	2.033^{9}	0.193^{5}	0.404^{6}	0.818^{8}	0.029^{2}	0.035^{3}	0.601^{7}	0.023^{1}	0.105^{4}
	3	27.012^{10}	17.210^{9}	1.455^{6}	3.571^{7}	7.222^{8}	0.376^{5}	0.057^{2}	0.343^{4}	0.041^{1}	0.165^{3}
	5	75.159^{10}	47.941^{9}	3.669^{6}	9.689^{7}	18.677^{8}	1.922^{5}	0.204^{4}	0.174^{2}	0.067^{1}	0.195^{3}
0.99	1	16.719^{10}	10.444^{9}	0.851^{7}	0.845^{6}	2.108^{8}	0.172^{4}	0.011^{2}	0.410^{5}	0.007^{1}	0.036^{3}
	3	150.291^{10}	94.149^{9}	6.392^{6}	7.189^{7}	16.364^{8}	5.859^{5}	0.049^{3}	0.083^{4}	0.038^{1}	0.046^{2}
	5	418.68^{10}	262.843^{9}	16.126^{5}	18.922^{6}	44.629^{8}	28.061^{7}	0.190^{4}	0.048^{1}	0.085^{3}	0.058^{2}

The n observations on the dependent variable are computed as

$$y_i = \boldsymbol{\beta}_0 + \boldsymbol{\beta x_{i1}} + \boldsymbol{\beta} x_{i2} + \cdots + \boldsymbol{\beta} x_{ip} + \boldsymbol{\epsilon}_i$$

where $\boldsymbol{\epsilon}_i$ are random errors that are normally distributed with zero mean vector and covariance matrix σ^2 and $\boldsymbol{\beta}_0$ is considered to be identically zero. To see the effect of high multicollinearity in the model, we consider three different values for ρ, i.e., $\rho = 0.90, 0.95, 0.99$. To assess the effect of sample size, $n = 25, 50,$ and 100 are considered. Moreover, the number of regressor, p, is set to 4, 8, and 16 whereas the error variance $\sigma^2 = 1, 3,$ and 5 are used. Furthermore,

TABLE 4. Descriptive statistics for the Cruise-ship-info data.

	age	tonnage	passengers	length	cabins	pasgrden	crew
Min	4.00	2.32	0.66	2.79	0.33	17.70	0.59
Median	14.00	71.89	19.50	8.56	9.57	39.09	8.15
Mean	15.00	71.29	18.46	8.13	8.83	44.19	7.79
Max	48.00	220.00	54.00	11.82	27.00	71.43	21.00

TABLE 5. Correlation matrix for the Cruise-ship-info data.

	age	tonnage	passengers	length	cabins	pasgrden	crew
Age	1.00	-0.61	-0.52	-0.53	-0.51	-0.28	-0.53
Tonnage	-0.61	1.00	0.95	0.92	0.95	-0.04	0.93
passengers	-0.52	0.95	1.00	0.88	0.98	-0.29	0.92
length	-0.53	0.92	0.88	1.00	0.89	-0.09	0.90
cabins	-0.51	0.95	0.98	0.98	1.00	-0.25	0.95
pasgrden	-0.28	-0.04	-0.29	-0.09	-0.25	1.00	-0.16
crew	-0.53	0.93	0.92	0.90	0.95	-0.16	1.00

simulation studies are repeated 2,000 time and for each replication, the mean squared error (MSE) of the estimators is computed.

A. SIMULATION RESULTS AND DISCUSSION

This section presents the performance of the existing as well as the proposed estimators using the mean squared error (MSE) as the assessment measure.

To evaluate the performance of the different estimators on simulated data sets, different factors such as standard deviation of the error term σ, correlation coefficient ρ, sample size n, and the number of explanatory variables p are varied and the results are listed in Tables 1, 2 and 3. These tables report the MSEs and their ranking for different estimators used in this study. From these tables, one can observe that the correlation coefficient and the sample size significantly affect the MSE of the estimators. The MSE increased for all the estimators when we increase the degree of correlation. In general, when the degree of correlation ρ is increased, the performance of our proposed estimators, especially, IFS_1 and IFS_4 becomes more visible as compared to other estimators. With the increase in sample size, the MSE generally tends to decrease, which has been previously observed in the literature [3]. Increasing the standard deviation σ of errors also increases the MSEs of the estimators.

Considering the results for the number of explanatory variables p separately, Table 1 reports the results for $p = 4$ with the aforementioned specification of n, ρ, and σ. From this table, note that the OLS estimator performs worst for different specifications used. In the case of $n = 25$ and $\rho \in \{0.90, 0.95\}$, our proposed estimators IFS_3 and IFS_4 outperform the competitors. However, with the increase in the values of ρ and σ, IFS_1 improves its ranking. For $n = 50$,

TABLE 6. Variance inflation factor.

Variables	disp	hp	drat	wt	qsec
VIF-value	9.11	5.20	2.32	7.01	3.19

the performance of our proposed estimator IFS_1 is remarkable as it produces lower MSEs compared to the rest. On the other hand, for $\rho = 0.9$ and smaller σ values, KGM also performs relatively better. In the case of $n = 100$, our proposed estimator, especially IFS_4, outperforms the rest in terms of the lowest MSEs. It is worth mentioning that for any combination of n, ρ and σ, one of our proposed estimators produces the lowest MSE, ranked 1^{st}, that shows the significance of our proposed estimators.

In Tables 2 and 3, we only change the number of variables, that is, $p = 8$ and 16, respectively, and compute the MSEs for different estimators using the aforementioned specifications of n, ρ and σ. From these tables, one can observe that the correlation between the explanatory variables, sample size, and error standard deviation affect the MSE of the estimators. From the results of Table 2 we can see that when $\rho = 0.90$ and $\sigma = 1$, IFS_1 performs well, for $\sigma = 3, 5$, the performance of IFS_2 and IFS_3 are better than the rest of the estimators. Note that as the value of ρ increases, the MSEs produced by our proposed estimators are much smaller than the OLS and existing ridge regression estimators, indicating that our proposed estimators are capable to handle strong multicollinearity situation better than the rest.

The results in Table 3 for $p = 16$ show that for $\rho = 0.90$, $n = 25$, $\sigma = 1$, KGM and KAM perform better than the rest. However, as the value of σ increases, the proposed estimators outperform the rest. For other specifications of ρ, n, and σ, our proposed estimators perform well than the existing estimators. Among our proposed estimators, the performance of IFS_3 is evident as compared to the rest.

TABLE 7. Mean square error for different estimators for the Cruise-ship-info data.

OLS	HK	KGM	KAM	KMED	KSM	IFS_1	IFS_2	IFS_3	IFS_4
1.985^{10}	1.967^{8}	1.479^{4}	1.794^{5}	1.880^{6}	1.970^{9}	1.301^{3}	1.941^{7}	1.095^{1}	1.195^{2}

To conclude the simulation study results assuming idifferent combinations of parameters, the proposed estimators generally have the smallest MSEs than the ordinary least squared and the existing ridge regression estimators, indicating our proposed estimators' significance and superiority over the rest.

IV. REAL DATA APPLICATION

In the previous section, the Monte Carlo simulation study compares the proposed estimators' performance with the existing ridge regression estimators. As in simulation studies, some ideal conditions are considered. This section presents two real-data examples to compare our proposed estimators' performance in practical situations

A. CASE STUDY 1

To check the performance of the proposed estimators, we use Cruise-ship-info data set. The data set is freely available on University of Florida website.[1] The data contain 6 predictors namely, ship (name of the ship), age (age up to 2013), tonnage (weight of the ship in tonnage), passengers (passengers on board (in 100s)), length (length of the ship (in 100s of feet)), cabins (number of cabins (in 100s)), pasgrden, (passenger density) and an outcome variable crew (number of crew (in 100s)). The descriptive statistics for the data are given in Table 4. From the table, one can see that variables are on different scales, therefore before computing the results of different estimators, we standardized them. Table 5 shows the correlation matrix of the data and it is evident that most of the variables are highly correlated with each other. For example, there is a strong positive linear relationship between crew and cabins, length, passengers, and tonnage. The age variable has a moderate-weak relationship to the crew, while passenger density shows a weak relationship to the number of crews on board. The correlation coefficient of 0.95 between the number of cabins on the ship and the number of crews indicates that these two variables are strongly and positively related. Similarly, Tonnage and passengers has the correlation value of 0.95, whereas cabins are strongly correlated with passengers and length ($\rho = 0.98$) indicating that the data set has a strong multicollinearity problem. To investigate further, we calculated the variance inflation factors (VIF) that detect multicollinearity in the data. The VIF estimates show how much the variance of a regression coefficient is inflated due to multicollinearity in the model. The VIFs are calculated by taking a predictor and regressing it against every other predictor in the model. This gives the R-squared values, which can then be plugged into the VIF formula. Finally, VIF can be calculated as

$$\text{VIF}_i = \frac{1}{1 - R_i^2}$$

[1] http://users.stat.ufl.edu/~winner/datasets.html

TABLE 8. Descriptive statistics for motor trend car road tests data.

	mpg	disp	hp	drat	wt	qsec
Min	10.400	71.100	52.000	2.760	1.513	14.500
Median	19.200	196.300	123.000	3.695	3.325	17.710
Mean	20.090	230.700	146.700	3.597	3.217	17.850
Max	33.900	472.000	335.000	4.930	5.424	22.900

TABLE 9. Correlation matrix for motor trend car road tests data.

	mpg	disp	hp	drat	wt	qsec
mpg	1.000	-0.848	-0.776	0.681	-0.868	0.419
disp	-0.848	1.000	0.791	-0.710	0.888	-0.434
hp	-0.776	0.791	1.000	-0.449	0.659	-0.708
drat	0.681	-0.710	-0.449	1.000	-0.712	0.091
wt	-0.868	0.888	0.659	-0.712	1.000	-0.175
qsec	0.419	-0.434	-0.708	0.091	-0.175	1.000

In general, a VIF above 5 indicates a high correlation and an indication of multicollinearity in the data. The VIF values for Cruise-ship-info data are given in Table 6 where one can note that the VIF values of the variables disp, hp, and wt are greater than 5, and thus the data set has a multicollinearity problem.

In Table 7, for each estimator, the estimated MSEs with their ranking are given. From this table, it is evident that ridge estimators have smaller MSE as compare to OLS. In addition, our proposed estimators IFS_3, IFS_4 and IFS_2 are ranked 1, 2, and 3, respectively. This suggests that the proposed estimators outperform their competitors, suggesting the superiority of these estimators over the rest.

B. CASE STUDY 2

We examine another dataset to check the performance of existing and proposed estimators. The dataset is extracted from the 1974 US Motor Trend magazine, and it contains fuel consumption information and design features of a small (32) set of cars. The data contain 11 predictors namely mpg (Miles/(US) gallon), cyl (Number of cylinders), disp (displacement (cu.in.)), hp (gross horsepower), drat (rear axle ratio), wt (weight (in 1000 lbs)), qsec (1/4 mile time), vs (Engine (0 = v-shaped, 1 = straight)), am (transmission (0 = automatic, 1 = manual)), gear (Number of forward gears), and carb (Number of carburetors). We selected only five highly correlated predictors from this data: disp, hp, drat, wt, and qsec. Data set is freely available in R from the dplyr package.

The descriptive statistics of the data are given in Table 8 and one can see that variables are measured on different

TABLE 10. Variance inflation factor.

Variables	age	tonnage	passengers	length	cabins	pasgrden
VIF-value	1.90	31.93	34.54	6.86	25.92	3.04

TABLE 11. MSEs for the motor trend car road tests data.

OLS	HK	KGM	KAM	KMED	KSM	IFS_1	IFS_2	IFS_3	IFS_4
35.766^{10}	33.225^{7}	5.808^{4}	35.154^{8}	23.238^{6}	6.198^{5}	3.139^{3}	35.718^{9}	0.905^{1}	2.535^{2}

scales, for example, the displacement variable ranges from 71 to 472 cm, while the gross horsepower variable ranges from 52 to 335.

Pairwise correlation coefficients for different variables are listed in Table 9 and one can see that most of the variables are highly correlated with each other. For example, the correlation between weight and miles/(US gallon) is -0.868, indicating a high negative correlation. Similarly, the correlation between displacement and weight shows a high linear positive correlation of 0.888. The VIFs values for this data are listed in Table 10. From this table, note that the VIF values for the variables tonnage, passengers, length, and cabins are much higher than the threshold value of 5, indicating a strong correlation among predictors. Thus, there is a multicollinearity problem in the data. The MSEs for different estimators along with their ranking are listed in Table 11. From this table, one can see that the existing as well as the proposed ridge regression estimators have the smallest MSEs compared to the OLS estimator. Note that the MSEs for the proposed estimators are much smaller than the rest estimators. Moreover, the results indicate that the proposed estimators, IFS_3, IFS_4 and IFS_1 outperform their competitors with IFS_3 is the best in terms of MSE compared to all other estimators.

V. CONCLUSION

It is well documented in the literature that in the presence of multicollinearity, the OLS estimators are inconsistent, has large variances, and consequently, can lead to wrong inferences. Ridge regression is a well known technique used in the presence of multicollinearity in the data. However, this technique heavily depends on the estimation of ridge parameter k. Thus, the main aim of this article is to introduce some new estimators for estimating the ridge parameter. To this end, we used the Jackknife approach and proposed new estimators to estimate k. The performance of new estimators is evaluated by extensive Monte Carlo simulations and two real data studies. The results show that as the number of variables, standard deviation of the random error, and the correlation between the independent variables increase, the MSEs also increase. On the other hand, when the sample size increases, the MSEs decreased. From the simulation studies and real data examples, the results indicated that, in general, the proposed estimators IFS_3, IFS_4, and IFS_1 have the smallest MSEs than the ordinary least squared estimators as well as existing ridge regression estimators. Further, we conclude that the ridge parameter estimates computed by the Jackknife technique are better. As in this study, we consider only Gaussian error, in the future, the study can be extended by using errors from other distributions.

ACKNOWLEDGMENT

The research is partially supported by Artificial Intelligence and Data Analytics Lab CCIS Prince Sultan University. The authors would like to acknowledge the support of Prince Sultan University for paying the Article Processing Charges (APC) of this publication.

CONFLICT OF INTEREST

The authors have no conflict to declare for this publication.

REFERENCES

[1] A. E. Hoerl and R. W. Kennard, "Ridge regression: Applications to nonorthogonal problems," *Technometrics*, vol. 12, no. 1, pp. 69–82, Feb. 1970.
[2] G. Muniz and B. G. Kibria, "On some ridge regression estimators: An empirical comparisons," *Commun. Statist. Simul. Comput.*, vol. 36, no. 2, pp. 115–138, 2009.
[3] S. Ali, H. Khan, I. Shah, M. M. Butt, and M. Suhail, "A comparison of some new and old robust ridge regression estimators," *Commun. Statist. Simul. Comput.*, pp. 1–19, Apr. 2019, doi: 10.1080/03610918.2019.1597119.
[4] J. Zhang and M. Ibrahim, "A simulation study on SPSS ridge regression and ordinary least squares regression procedures for multicollinearity data," *J. Appl. Statist.*, vol. 32, no. 6, pp. 571–588, Aug. 2005.
[5] L. Firinguetti, G. Kibria, and R. Araya, "Study of partial least squares and ridge regression methods," *Commun. Statist. Simul. Comput.*, vol. 46, no. 8, pp. 6631–6644, Sep. 2017.
[6] M. A. Alkhamisi and G. Shukur, "Developing ridge parameters for SUR model," *Commun. Statist. Theory Methods*, vol. 37, no. 4, pp. 544–564, Jan. 2008.
[7] G. C. McDonald and D. I. Galarneau, "A Monte Carlo evaluation of some ridge-type estimators," *J. Amer. Stat. Assoc.*, vol. 70, no. 350, pp. 407–416, Jun. 1975.
[8] D. G. Gibbons, "A simulation study of some ridge estimators," *J. Amer. Stat. Assoc.*, vol. 76, no. 373, pp. 131–139, Mar. 1981.
[9] B. M. G. Kibria, "Performance of some new ridge regression estimators," *Commun. Statist. Simul. Comput.*, vol. 32, no. 2, pp. 419–435, Jan. 2003.
[10] M. Alkhamisi, G. Khalaf, and G. Shukur, "Some modifications for choosing ridge parameters," *Commun. Statist. Theory Methods*, vol. 35, no. 11, pp. 2005–2020, Nov. 2006.
[11] G. Muniz, B. Kibria, and G. Shukur, "On developing ridge regression parameters: A graphical investigation," *Statist. Operations Res. Trans.*, vol. 38, no. 3, pp. 621–630, 2009.
[12] G. Khalaf and G. Shukur, "Choosing ridge parameter for regression problems," *Commun. Statist. Theory Methods*, vol. 34, no. 5, pp. 1177–1182, May 2005.
[13] K. Månsson, G. Shukur, and B. M. G. Kibria, "A simulation study of some ridge regression estimators under different distributional assumptions," *Commun. Statist. Simul. Comput.*, vol. 39, no. 8, pp. 1639–1670, Aug. 2010.
[14] B. Kibria and S. Banik, "Some ridge regression estimators and their performances," *J. Modern Appl. Stat. Methods*, vol. 15, no. 1, p. 12, 2016.
[15] B. M. G. Kibria and A. F. Lukman, "A new ridge-type estimator for the linear regression model: Simulations and applications," *Scientifica*, vol. 2020, pp. 1–16, Apr. 2020.
[16] A. K. M. Ehsanes Saleh and B. M. Golam Kibria, "Performance of some new preliminary test ridge regression estimators and their properties," *Commun. Statist. Theory Methods*, vol. 22, no. 10, pp. 2747–2764, Jan. 1993.

[17] L. Wang and F. Yu, "Jackknife resample method for precision estimation of weighted total least squares," *Commun. Statist. Simul. Comput.*, pp. 1–18, Apr. 2019.
[18] M. Arashi and T. Valizadeh, "Performance of Kibria's methods in partial linear ridge regression model," *Stat. Papers*, vol. 56, no. 1, pp. 231–246, Feb. 2015.
[19] A. Clark and C. Troskie, "Ridge regression—A simulation study," *Commun. Statist.-Simul. Comput.*, vol. 35, no. 3, pp. 605–619, 2006.
[20] J. Cadima and I. T. Jolliffe, "Loading and correlations in the interpretation of principle compenents," *J. Appl. Statist.*, vol. 22, no. 2, pp. 203–214, Jan. 1995.
[21] E. Cule and M. De Iorio, "Ridge regression in prediction problems: Automatic choice of the ridge parameter," *Genet. Epidemiol.*, vol. 37, pp. 704–714, 2013, doi: 10.1002/gepi.21750.
[22] E. Cule and M. De Iorio, "A semi-automatic method to guide the choice of ridge parameter in ridge regression," 2012, *arXiv:1205.0686*. [Online]. Available: http://arxiv.org/abs/1205.0686
[23] K. Y. Wong and S. N. Chiu, "An iterative approach to minimize the mean squared error in ridge regression," *Comput. Statist.*, vol. 30, no. 2, pp. 625–639, Jun. 2015.
[24] V. K. Vijayan, K. R. Bindu, and L. Parameswaran, "A comprehensive study of text classification algorithms," in *Proc. Int. Conf. Adv. Comput., Commun. Informat. (ICACCI)*, Sep. 2017, pp. 1109–1113.
[25] G. Khalaf, K. Månsson, and G. Shukur, "Modified ridge regression estimators," *Commun. Statist. Theory Methods*, vol. 42, no. 8, pp. 1476–1487, Apr. 2013.
[26] M. A. Alkhamisi and G. Shukur, "A Monte Carlo study of recent ridge parameters," *Commun. Statist. Simul. Comput.*, vol. 36, no. 3, pp. 535–547, May 2007.
[27] H. Obiorailouno and J. Mbegbu, "A Jackknife approach to error-reduction in nonlinear regression estimation," *Amer. J. Math. Statist.*, vol. 3, no. 1, pp. 32–39, 2013.
[28] B. Efron and R. J. Tibshirani, *An Introduction to Bootstrap*. Boca Raton, FL, USA: CRC Press, 1994.
[29] M. H. Quenouille, "Approximate tests of correlation in time-series," *J. Roy. Stat. Soc., Ser. B, Methodol.*, vol. 11, no. 1, pp. 68–84, Jan. 1949.
[30] J. Tukey, "Bias and confidence in not quite large samples," *Ann. Math. Statist.*, vol. 29, no. 2, p. 614, 1958.
[31] R. G. Miller, "A trustworthy jackknife," *Ann. Math. Statist.*, vol. 35, no. 4, pp. 1594–1605, Dec. 1964.
[32] R. G. Miller, "The jackknife—A review," *Biometrika*, vol. 61, no. 1, pp. 1–15, 1974.
[33] H. L. Gray and W. R. Schucany, *The Generalized Jackknife Statistic*, vol. 1. New York, NY, USA: M. Dekker, 1972.
[34] D. V. Hinkley, "Jackknife confidence limits using student t approximations," *Biometrika*, vol. 64, no. 1, pp. 21–28, 1977.
[35] J. A. Reeds, "Jackknifing maximum likelihood estimates," *Ann. Statist.*, vol. 6, no. 4, pp. 727–739, Jul. 1978.
[36] D. Hinkley and B.-C. Wet, "Improvements of jackknife confidence limit methods," *Biometrika*, vol. 71, no. 2, pp. 331–339, 1984.
[37] P. K. Sen, "Functional jackknifing: Rationality and general asymptotics," *Ann. Statist.*, vol. 16, no. 1, pp. 450–469, Mar. 1988.
[38] C. F. J. Wu, "Jackknife, bootstrap and other resampling methods in regression analysis," *Ann. Statist.*, vol. 14, no. 4, pp. 1261–1295, Dec. 1986.
[39] J. Shao and C. F. J. Wu, "A general theory for jackknife variance estimation," *Ann. Statist.*, vol. 17, no. 3, pp. 1176–1197, Sep. 1989.
[40] J. Shao, "Consistency of jackknife variance estimators Jun," *Statistics*, vol. 22, no. 1, pp. 49–57, Jan. 1991.
[41] S. Sahinler and D. Topuz, "Bootstrap and Jackknife resampling algorithms for estimation of regression parameters," *J. Appl. Quant. Methods*, vol. 2, no. 2, pp. 188–199, 2007.

ISMAIL SHAH received the master's degree from Lund University, Sweden, and the Ph.D. degree from the University of Padova, Italy. He is currently an Assistant Professor with the Department of Statistics, Quaid-i-Azam University, Islamabad, Pakistan. His research interests include functional data analysis, time series analysis, regression analysis, energy economics, and applied and industrial statistics.

FAIZA SAJID received the M.Phil. degree in statistics from Quaid-i-Azam University (QAU), Islamabad, Pakistan. Her research interests include regression analysis, shrinkage methods, and applied statistics.

SAJID ALI received the Ph.D. degree in statistics from Bocconi University, Milan, Italy. He is currently an Assistant Professor with the Department of Statistics, Quaid-i-Azam University (QAU), Islamabad, Pakistan. His research interests include time series analysis, Bayesian inference, construction of new flexible probability distributions, and process monitoring.

AMJAD REHMAN (Senior Member, IEEE) received the Ph.D. degree (Hons.) in forensic documents analysis and security from the Faculty of Computing, Universiti Teknologi Malaysia, in 2010. He held the postdoctoral position with Universiti Teknologi Malaysia, in 2011. He is currently a Senior Researcher with the Artificial Intelligence and Data Analytics Laboratory, CCIS, Prince Sultan University, Riyadh, Saudi Arabia. He is also the PI in several funded projects and also completed projects funded from MOHE Malaysia, Saudi Arabia. He is the author of more than 200 ISI, Scopus journal articles, and conferences. His research interests include big data mining, health informatics, pattern recognition, and forecasting. He received the Rector Award for the 2010 Best Student from Universiti Teknologi Malaysia.

SAEED ALI BAHAJ received the Ph.D. degree from Pune University, India, in 2006. He is currently an Associate Professor with the Department of Computer Engineering, Hadhramout University, and also an Assistant Professor with Prince Sattam Bin Abdul-Aziz. His main research interests include artificial intelligence, information management, forecasting, information engineering, big data, and information security.

SULIMAN MOHAMED FATI (Senior Member, IEEE) received the Ph.D. degree from Universiti Sains Malaysia (USM), Malaysia, in 2014. He is currently an Assistant Professor with the Information Systems Department, CCIS, Prince Sultan University, Saudi Arabia. His research interests include the Internet of Things, machine learning, cloud computing, cloud computing security, and cyber security. He is an active member in different professional bodies as ISACA, IACSIT, IAENG, and the Institute of Research Engineers and Doctors, USA. He serves as reviewer in many international impact-factor journals, and a TPC in different international conferences.