Comparing the Proposed Sub-Ridge Regression with Ridge and OLS for A Shrinkage Factor; Data with or Without Multicollinearity

Nelson, Maxwell[1], Biu, Oyinebifun Emmanuel[2] and Onu, Obineke Henry[3]
[1]Mathematics Department,
Isaac Jasper Boro College of Education, Sagbma, Bayelsa, State.
[2]Mathematics/Statistics
Department, University of Port Harcourt, Rivers State.
[3]Mathematics/Statistics Department,
Ignatius Ajuru University of Education, Rumuolumeni, Port Harcourt, Rivers State.
Corresponding author email: **onu.obineke-henry@iaue.edu.ng**

DOI: 10.56201/ijcsmt.v10.no2.2024.pg22.34

Abstract

1. *The study considered the comparisons of the proposed sub-ridge regression method with ridge and OLS for shrinkage factor, for data with or without multi-collinearity. The various values of the shrinkage factor k were simulated and tested for ridge and sub-ridge regression and the results were compared with the Ordinary Least Square analysis initially done for three different sample sizes, 50, 75 and 100 for five economic data and they were, Exchange rate, Unemployment, Inflation. The value of k=**0.0000005** was proposed as the convergent factor for which the Sub-Ridge becomes equal to the OLS parameters. As the k value decreases, the sum of square error of the Sub-Ridge regression parameters also decreases, for k=0.000007 the sum of square of the Sub-Ridge became equal to the sum of square error of the Ridge. The study recommended among other things that, even in the absence of multi-collinearity, Ridge regression and Sub-Ridge regression can still be used in obtaining equal parameter estimates and equal sum of square errors with the Ordinary Least Square, but for k value of 0.000007 for Ridge and k value of 0.0000005 for Sub-Ridge.*

Keywords: *OLS, Ridge, Sub-ridge regressions, determinant, condition index, VIF.*

1. Introduction

Regression analysis is known as a powerful analysis that can investigate multiple variables simultaneously to answer complex research questions. Nevertheless, if you do not satisfy the ordinary least square regression (OLS) assumptions, you might not be able to trust the results (Jim, 2020). Regression analysis is like other inferential methodologies with the goal of drawing a random sample from a population and use it to estimate the properties of that population. In regression analysis, the coefficients in the regression equation are estimates of the actual population parameters, it is expected that these coefficient estimates be the best possible estimates. Supposing one requests an estimate for the cost of a service that is being considered, a reasonable estimate can be defined by:

a) The estimates should tend to be right on target, which means that, they should not be systematically too high or too low. In a simplified way, it should be unbiased.
b) Recognizing that estimates are almost never exactly correct, you want to minimize the inconsistency between the estimated value and actual value, because large differences are not good.

The above two properties are precisely what we need for our coefficient estimates. If the linear regression model satisfies the OLS assumptions, the procedure generates unbiased coefficient estimates that tend to be relatively close to the true population values (minimum variance). As long as your model satisfies the OLS assumptions for linear regression, one can rest, knowing that the best estimates are being obtained. According to Onu, *et al.* (2021) and Shalabh (2012), a simple linear regression is an approach in statistics that is employed in the modeling of a linear surfaces. Regression analysis can be linear, nonlinear, second-order (quadratic or polynomial) regression.

The problem of multicollinearity in a data set has gone a long way in falsifying regression results, hence, introduction of ridge regression to be used when the data have been confirmed of having the presence of multicollinearity. To determine whether or not there is multicollinearity in a data set, statistical test have to be conducted. This test is also one of the setbacks in research work. In overcoming the testing of data for multicollinearity in this particular case of outlier or multicollinearity, the introduction of a proposed method is considered. Though, this research was in line with work of Nelson, et al. (2024) who proposed two of such methods (Mult- and Inverse-Ridge Regressions with similar approaches but different principles). The proposed method in this work is known as the Sub-ridge regression approach. This method gives equal or approximate equal results with the ridge and the ordinary least square regression. This is independent of whether the data has presence of multicollinearity or not. The ideal of testing data for this reason is a thing of the past.

Abubakari (2019) used Principal components as remedial to multicollinearity problem. Using a sample of six hundred participants, linear regression model was fitted and collinearity between predictors was detected using Variance Inflation Factor (VIF). After confirming the existence of high relationship between independent variables, the principal components were utilized to find the possible linear combination of variables that can produce large variance without much loss of information. The results show that Variance Inflation Factor values for each predictor ranged from 1 to 3 which indicates that multicollinearity problem was eliminated. Finally, another linear regression model was fitted using Principal components as predictors. The assessment of relationship between predictors indicated that no any symptoms of multicollinearity were observed.

Younker (2012) have studied Ridge Estimation and its Modifications for Linear Regression with Deterministic or Stochastic Predictors, the study used multiple regression problem where Statistical inference problems is often accompanied by the 'bias-variance trade-off'. For a fixed number of observations every additional explanatory variable typically causes coefficient estimates to become less certain. As confidence intervals around these estimates expand, in time, interpretation and forecasting power are risked.

2. Materials and Methods

Testing for Outliers in a Data set

Grubb's test was used to detect outlier since it detects one outlier at a time. It involves the following steps as seen in Nelson, et al. (2024).

(i) Order the data point from smallest to largest.
(ii) Find the mean and standard deviation of the data set.
(iii) Calculate the G-test statistic using one of the following equations.

In test for outliers in this study, Grubbs' test was employed and it is given as

$$G = \max_{i=1,\ldots,N} \frac{|Y_i - \bar{Y}|}{s} \quad (1)$$

Y_i is the sample data from a given population, here it represents any of GDP, FDI, Exchange rate, Inflation rate and Unemployment rate and $\bar{Y}$ is the sample mean, while s is the sample standard deviation.

The Grubbs test can also be given as a one-sided test as

$$G = \frac{\bar{Y} - Y_{min}}{s} \quad (2)$$

or

$$G = \frac{Y_{max} - \bar{Y}}{s} \quad (3)$$

The test is based on the assumption of normality. It detects one outlier at a time, the outlier detected is removed from the data set and the test is repeated until no more outlier is detected.

$$\hat{Y} = \frac{\Sigma Y_i}{n} \quad (4)$$

Where, $\hat{Y}$ is the arithmetic mean

Y_i is individual data value

n is the total number of data

$$S = \sqrt{\frac{\Sigma (Y_i - \hat{Y})^2}{n-1}} \quad \text{is the standard deviation} \quad (5)$$

Geometric mean is $\sqrt[N]{y_1 x y_2 x y_3 x.x.x.y_N}$ (6)

Harmonic mean is $\mathrm{H,M} = \frac{N}{(\frac{1}{Y_1} + \frac{1}{Y_2} + \frac{1}{Y_3} + \cdots + \frac{1}{Y_N})}$ (7)

Median is given as the size of $\frac{(N+1)^{th}}{2}$ $item$ (8)

Testing for the Presence of Multi-collinearity in the Data Set

Testing for multi-collinearity in the data sets, we employ the following methods.

Variance Inflation Factor (VIF)

Variance Inflation Factor according to Ayuya, (2021) and Deanna, (2018), the VIF is given as

$$VIF = \frac{1}{1 - R^2} \quad (9)$$

Where Coefficient of Determination (R^2) is the R-squared value obtained from the regression of X_i on the other independent variables. It is seen, if the R-squared in the denominator gets closer and closer to one, the VIF will get larger and larger. The rule of thumb cut-off value for VIF is 10. Solving backwards, this translates into an R-squared value of 0.90. Hence, whenever the R-squared value between one independent variable and the rest is greater than or equal to 0.90, you will have to face multi-collinearity.

According to Thompson, et al. (2017) and Nelson, et al. (2024), the coefficient of determination is given as

$$R^2 = \frac{SSR}{SST} = \frac{\Sigma(\hat{y}_i - \bar{y})^2}{\Sigma(y_i - \bar{y})^2} = 1 - \frac{SSE}{SST} = 1 - \frac{\Sigma(y_i - \hat{y}_i)^2}{\Sigma(y_i - \bar{y})^2}$$

Condition Number and Condition Index

In order to find the eigen values of a matrix, given a k x k matrix A, a k x k identity matrix I and an eigen value ⅄, the following steps are to be followed:

a) Be sure that the given matrix A is a square matrix k x k.
b) Estimate the matrix. That is $|A - \lambda I|$
c) Find the determinant of the matrix.
d) From the equation obtained $|A - \lambda I| = 0$
e) Calculate all the possible values of the equation.

The square root of the ratio between the maximum and each eigenvalue (λ_1, λ_2, …, λ_k) is referred to as the condition index:

$$k_s = \sqrt{\frac{\lambda_{max}}{\lambda_s}}, (s = 1,2, \dots, k) \tag{10}$$

The largest condition index is called the condition number and is the most widely used estimator to measure the strength of multi-collinearity called condition number by (Vinod & Uallh, 1981) is defined as $k = \sqrt{\frac{\lambda_{max}}{\lambda_{min}}}$ (11)

Where λ_{max} and λ_{min} are the largest and smallest eigenvalues of the matrix $X'X$ respectively. If λ_{min} is zero, then is k is infinite, means perfect multi-collinearity among the independent variables and if λ_{max} is equal to λ_{min}, then k is one, means the independent variables are said to be orthogonal. If k is between 30 to 100, it indicates a moderate to strong multi-collinearity. Any k value greater than 100 suggests severe multi-collinearity and larger value indicates serious multi-collinearity.

Correlation

This study is interested in the correlation that exist between two predictor variables as seen

$$r_{x_i x_j} = \frac{n\sum x_i x_j - (\sum x_i)(\sum x_j)}{\sqrt{(n\sum x_i{}^2 - (\sum x_i)^2)(n\sum x_j{}^2 - (x_j)^2)}} \tag{12}$$

Where x_i and x_j represent the i^{th} and j^{th} predictor variables, the higher value of r indicates higher presence of multicollinearity, while the lower value of r indicates reduced presence of multucollinearity. The formula of the correlation is as seen in Onu, et al. (2021).

Determinant of a Matrix

Key Points of determinant

a) Let A be an m×n matrix and k an integer with 0<k≤m, and k≤n. A k×k minor of A is the determinant of a k×k matrix obtained from A by deleting m-k rows and n-k columns.
b) The first minor of a matrix Mij is formed by removing the ith row and jth column of the matrix, and retrieving the determinant of the smaller matrix.
c) The cofactor of an element aij of a matrix A, written as Cij is defined as $(-1)^{i+j}$Mij.

Key Terms

a) **Cofactor**: The signed minor of an entry of a matrix.

b) Minor: The determinant of some smaller square matrix, cut down from matrix A by removing one or more of its rows or columns (Boundless, 2018).

The Parameter Estimates of Ordinary Least Square

This study will employ a five-parameter probabilistic model given as

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \varepsilon \tag{13}$$

Where Y is the Gross Domestic Product (GDP) of Nigeria used as the response variable, while X_1 is the Exchange rate, X_2 is the Unemployment rate, X_3 represents the Inflation rate, and X_4 is the Foreign Direct Investment (FDI) in Nigeria are the predictor variables, $\beta_0, \beta_1, \beta_2, \beta_3$ *and* β_4 are the unknown model parameters while ε is the stochastic disturbance or simply the error. The model in equation (13) is a multiple linear regression and it can be written in matrix form as:

$$Y = X\beta + \varepsilon \tag{14}$$

$$Y = \begin{pmatrix} y_1 \\ y_2 \\ . \\ . \\ . \\ y_n \end{pmatrix}, \quad X = \begin{pmatrix} 1 & x_{11} & x_{12} & . & . & . & x_{1p} \\ 1 & x_{21} & x_{22} & . & . & . & x_{2p} \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ 1 & x_{n1} & x_{n2} & . & . & . & x_{pp} \end{pmatrix}, \quad \beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ . \\ . \\ . \\ \beta_p \end{pmatrix}, \quad \varepsilon = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ . \\ . \\ . \\ \varepsilon_p \end{pmatrix}$$

where X is an $N \times P$ matrix, Y is an $N \times 1$ vectors of observed parameters and β is a $P \times 1$ vectors of unknown parameters and $\varepsilon \sim N(0, \delta^2)$ is the error term. The model in 3.13 we obtain the matrix X, the transpose of this matrix is obtained given as X'. The matrix X is multiplied by its transpose to obtain $X'X$ known as the information matrix. The inverse of $X'X$ is obtained by using the formula

$$(X'X)^{-1} = \frac{Adjoint(X'X)}{\det(X'X)} \qquad (15)$$

Where $\det(X'X)$ is the determinant of $X'X$.

The transpose of X is multiplied by the response variable Y to obtain $X'Y$. In order to obtain the parameters of the model in 1.3, the Ordinary Least Square formula is applied and given as seen in (Iwundu & Onu, 2017, Onu, et al. 2021 and Kutner, *et al.*2005).

$$\underline{\hat{\beta}} = (X'X)^{-1}X'\underline{Y} \qquad (16)$$

The Parameter Estimates of Ridge Regression for varying Values of Shrinkage Penalty

The Ridge Regression is like the Ordinary Least Square method; the only difference is the addition of the quantity KI to the information matrix to remove the effect of multi-collinearity in the analysis. K is a constant that takes on values not greater than 0.2 and the smaller the value of K, the better the Ridge parameters estimated and the higher the values of K above 0.2, the more the information matrix becomes singular matrix (Nduka & Ijomah, 2012).

It is given by the formula

$$\underline{\hat{\beta}} = (X'X + KI)^{-1}X'\underline{Y} \qquad (17)$$

Where I is an identity matrix.

The proposed Estimates of Subtraction based Ridge Regression for varying Shrinkage Penalty Values (Sub-ridge regression).

This is one of the approaches that was developed in this research to see how it can compete with the popularly known Ordinary Least Squares and the Ridge Regression, that is used with or without multi-collinearity in the data sets. It is given as

$$\underline{\hat{\beta}} = (X'X - KI)^{-1}X'\underline{Y} \qquad (18)$$

All the variables have their usual meaning.

3. Results and Discussion

Investigating the Presence of Multi-collinearity

We are to test the presence of multi-collinearity in the data of Exchange rate, Unemployment rate, Inflation rate and Foreign direct investment as predictors and the Gross Domestic Product in Nigerian using the following methods

Given that Matrix $X'X = \begin{pmatrix} 30 & 51.42 & 560.41 & 134.39 & 3999.50 \\ 51.42 & 129.63 & 1219.85 & 207.96 & 5508.88 \\ 560.41 & 1219.85 & 18624.82 & 2424.55 & 55538.91 \\ 134.39 & 207.96 & 2424.55 & 676.47 & 21657.64 \\ 3999.50 & 5508.88 & 55538.91 & 21657.64 & 800875.37 \end{pmatrix}$

The table 1 below shows the eigen values obtained. The largest condition index is called the condition number which is 4.9413 and is less than thirty, it indicates moderate or not serious multi-colliearity.

Ordinary Least Square Regression

Regression Analysis: GDP versus FDI, INFL, UNEMPL, EXCH

Table 2: Analysis of Variance (ANOVA) of the Selected Economic Variables Data

Source of variation	Degree of Freedom	Sum of Square	Mean of Squares	F-Value	P-Value
Regression	4	204.345	51.086	6.16	0.001
FDI	1	4.480	4.480	0.54	0.469
INFL	1	5.826	5.826	0.70	0.410
UNEMP	1	107.160	107.160	12.92	0.001
EXCH	1	54.359	54.359	6.55	0.017
Error	25	207.378	8.295		
Total	29	411.723			

Summary of the fitted model:

Root Square Error	R-squared	R-squared (adjusted)	R-squared (predicted)
2.88013	49.63%	41.57%	35.12%

$$GDP = 13.1323 - 0.406\,FDI - 0.0372\,INFL - 2.654\,UNEMPL + 0.0327\,EXCH$$

Table 3: Sum of Square Error, Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) Values.

	RIDGE			SUB-RIDGE		
K	**SSE**	**AIC**	**SBC**	**SSE**	**AIC**	**SBC**
0.000010	207.38	68.00	75.00	207.38	68.00	75.00
0.000005	207.38	68.00	75.01	207.38	68.00	75.01
0.000050	207.38	68.00	75.00	207.38	68.00	75.00
0.000070	207.38	68.00	75.01	903.11	112.14	119.15
0.000080	207.38	68.00	75.01	207.38	68.00	75.01
0.000320	207.38	68.00	75.00	207.38	68.00	75.00
0.000500	207.38	68.00	75.00	207.38	68.00	75.00
0.000900	207.38	68.00	75.01	207.38	68.00	75.01
0.002430	207.38	68.00	75.00	207.38	68.00	75.00
0.005000	207.38	68.00	75.00	207.38	68.00	75.00
0.006000	207.38	68.00	75.01	207.39	68.00	75.01

0.010240	207.38	68.00	75.00	207.38	68.00	75.00
0.031250	207.50	68.02	75.02	207.51	68.02	75.02
0.050000	207.67	68.04	75.05	207.72	68.05	75.06
0.077760	207.06	68.10	75.10	208.23	68.12	75.12
0.100000	20847	68.16	75.16	208.83	68.21	75.22
0.116000	208.82	68.21	75.21	209.39	68.29	75.30
0.168070	$1.48e^{10}$	$1.18e^{11}$	$1.18e^{11}$	8008.78	9453.92	391.85
0.200000	211.23	68.55	75.56	738.88	106.12	113.12
0.300000	215.07	69.09	76.10	225.62	70.53	77.54
0.327680	216.27	76.27	83.27	230.27	71.14	78.14
0.400000	219.59	69.72	76.72	246.54	73.19	80.20
0.500000	224.51	70.38	77.39	282.75	77.30	84.31
0.590490	229.16	71.00	78.00	336.96	82.56	89.56
0.600000	229.66	71.06	78.07	344.33	83.21	90.22
0.700000	234.90	71.74	78.74	449.91	91.24	98.24
0.800000	240.15	72.40	79.41	636.36	101.64	108.64
0.900000	245.35	73.04	80.05	983.70	114.70	121.71
0.999000	250.40	73.66	80.67	1679.05	130.74	137.74
0.999990	250.45	73.66	80.67	1689.09	130.92	137.92
1.000006	250.45	73.66	80.67	1689.25	130.93	137.93
1.000009	311.52	80.21	87.21	1689.25	130.93	137.93
1.000010	250.45	73.66	80.67	1689.29	130.93	137.93
1.000020	250.45	73.66	80.67	1689.29	130.93	137.93
1.000030	250.45	73.66	80.67	1689.29	130.93	137.93
1.000040	250.45	73.66	80.67	1689.29	130.93	137.93
1.000050	250.45	73.66	80.67	1689.29	130.93	137.93
1.000060	250.45	73.66	80.67	1689.29	130.93	137.93
1.000070	250.45	73.66	80.67	1689.29	130.93	137.93
1.000080	250.45	73.66	80.67	1690.57	130.95	137.95
1.000090	250.45	73.66	80.67	1689.99	130.94	137.94

1.000100	250.45	73.66	80.67	1690.21	130.94	137.95
1.000200	928.77	112.98	119.99	1691.22	130.94	137.95
1.000300	250.46	73.66	80.67	1692.25	130.94	137.95
1.000400	311.57	80.21	87.22	1693.27	131.00	138.00
1.000500	250.47	73.66	80.67	1694.28	131.01	138.01
1.000600	250.48	73.66	80.67	1695.30	131.03	138.03
1.000700	250.48	73.66	80.67	1696.33	131.04	138.94
1.000800	250.49	73.67	80.67	1697.36	131.07	138.07
1.000900	250.49	73.67	80.67	1698.35	131.07	138.07
1.009000	250.90	73.72	80.72	1784.32	132.57	139.57

Table 4: The Proposed Convergent Points

K	OLS	Ridge		Sub-Ridge	
0.000000	**13.1323**	**13.1323**		**13.1323**	
	-0.4061	**-0.4061**		**-0.4061**	
	-0.0372	**-0.0372**		**-0.0372**	
	-2.6536	**-2.6536**		**-2.6536**	
	0.0327	**0.0327**		**0.0327**	
0.000005			**t_x**		**t_x**
		13.1323	5.61	**13.1323**	5.61
		-0.4061	-0.73	**-0.4061**	-0.73
		-0.0372	-0.84	**-0.0372**	-0.84
		-2.6536	-3.60	**-2.6536**	-3.60
		0.0327	2.55	**0.0327**	2.55
0.000007		**13.1323**		13.1324	
		-0.4061	-0.73	-0.4062	-0.73
		-0.0372	-0.84	-0.0372	-0.84
		-2.6536	-3.60	-2.6536	-3.60
		0.0327	2.56	0.00327	2.56

0.999999	7.8091	3.34	44.3455	18.95
	0.3414	0.62	-4.9425	1.70
	-0.0566	-1.27	0.0870	1.96
	-1.3887	-1.88	-10.0991	-13.68
	0.0213	1.66	0.1008	7.88
1.000006	7.8091	3.34	44.3463	18.95
	0.3414	0.62	-4.9426	-8.94
	-0.0566	-1.27	0.0870	2.00
	-1.3887	-1.88	-10.0993	-13.68
	0.0213	1.66	0.1008	7.88

Discussion of Results

Test for Outliers in the Data Set

The analysis of the result for test of outliers reveals that not all the variables have outlier in the data used in this study (not severe). These results show that Ordinary Least Square regression will be a better technique to be applied. The analysis of the results indicated that from the five selected economic variables used, using Grubb's method test for outlier detection, one outlier was detected on gross domestic products, there was no outlier detected on unemployment rate, only one outlier was equally detected on inflation rate, two outliers were detected on foreign direct investment and no outlier was detected on exchange rate. Each of the selected economic variables gives the same result, irrespective of the type of mean used. Grubb's test was not performed on some variables due to no mode. Since, there are some outliers that may cause multi-collinearity in a data set. Therefore, Ridge regression would be a better approach to use.

Test for Multi-collinearity in a Data Set

From the given information matrix and the result of determinant of information matrix shows that there is no presence of multi-collinearity or multi-collinearity is not a problem since the result is far from zero. This result has not contradicted the result of the condition index. The test for multi-collinearity, using condition number and condition index, shows that there is presence of multi-collinearity among the predictor's variables but mild or not serious, since the largest condition index is 4.9413 and is less than 100. Also, the correlation between exchange rate and unemployment rate has a high value of 0.838, which shows that multi-collinearity is a problem between these two variables. But generally, the correlation reveals that multi-collinearity is not a problem. Because the pair of other variables had smaller correlation value.

Simulation of the Various Shrinkage Penalty Values and its Effect on the Parameters

The study discovered that for k-value of 0.0000, the ordinary least square (OLS) regression, Ridge regression and Sub-Ridge regression, gave equal values of parameters estimates. This study also suggests that needless of testing for multi-collinearity in a dataset, because this can lead to further academic stress, instead, some values of k were proposed for use to ensure that the ordinary least square (OLS) parameters, the Ridge and Sub-Ridge regression parameters are equal, whether there is multi-collinearity or not in the dataset. For k range of 0 to 0.00005, the Ridge and Sub-Ridge regression equations can be used to estimate parameters of a linear model and this result will give equal value of parameter estimate with the ordinary least square (OLS), irrespective of the presence of multi-collinearity or not in the data. The simulated result shows that the Ridge regression and the Sub-Ridge regression parameters are equal and equal to the OLS regression parameters for k=0.0000005, 0.0000001, 0.0000002, 0.0000003, 0.0000004, 0.0000006. The k values of 0.000006 and 0.000007, gave equal parameter estimates for Sub-Ridge regression, but differed in the estimation of the intercept or grand mean parameter of the Ridge regression.

Sum of Square Error for Varying Shrinkage Penalty Values

The sum of square error for the k-values from 0.00001 to 0.01024 gave a sum of square error equal or approximately equal to 207.38 for both Ridge regression and Sub-Ridge regression equations, The higher the value of k, the more the ordinary least square (OLS) regression parameters differ appreciable from the Ridge and tends to zero as the k value increases further. In addition, in the analysis of the sum of square error for Ridge regression, it was observed that as the k value decreases, the sum of square error of the Ridge regression decreases, the smaller the sum of square error, the better the regression equation. As the k value decreased up to **0.000007**, the sum of square error of the Ridge and that of the Sub-Ridge regressions became equal.

Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) Values.

From table 4.20 it was observed that the difference between Akaike Information Criterion and Bayesian Information Criterion Values is a constant or approximately with a value of seven, except for the shrinkage penalty of 0.168070 that produced a sum of square error of $1.48e^{10}$. In this particular value of k, Akaike information criterion and Bayesian information criterion are equal, which means that their difference is zero.

Parameters of Sub-Ridge Regression for Varying Shrinkage Penalty Values

The k-value of **0.0000005** was proposed as the convergent point for which the Sub-Ridge regression parameters and ordinary least square (OLS) regression give equal parameter estimates. It was observed that for k=0.006, the values of the Sub-Ridge regression parameters are found greater than those of the Ordinary Least Square parameters. As the value of k decreases to 0.0009, the values of the Sub-Ridge parameters also decrease closer to OLS value. As the k value decreases further to 0.00008 the Sub-Ridge regression value further decreases, at a certain k value of 0.000007, the parameters of the Sub-Ridge regression decrease more and more towards the OLS parameters. This was the value of k for which the Ridge regression became equal to the OLS, but such was not the case for Sub-Ridge regression, but for k=0.0000005, the Sub-Ridge regression became equal to the OLS parameters. The value of k=**0.0000005** was proposed as the convergent factor for which the Sub-Ridge becomes equal to the OLS parameters. As the k value decreases, the sum of square error of the Sub-Ridge regression parameters also decreases, for k=0.000007 the sum of square of the Sub-Ridge became equal to the sum of square error of the Ridge.

Conclusion

It is obvious that ordinary least square regression is better in the estimation of parameter for data without multi-collinearity or when multi-collinearity is not a problem, but in such a situation, Ridge regression can also be used but for k value of 0.000007, while Sub-Ridge regression can be used for k value of 0.0000005. All the above stated regression approaches will give equal estimation of parameters with the OLS and which also gives equal sum of square errors. It also concludes that the Sub-Ridge regression was somewhat better than the Ridge on the area that for increasing values of k, the Ridge regression will first have its parameter value equal to zero before the Sub- Ridge regression will become zero at some further increasing values of k. The both regression methods are better for decreasing values of k. Among the several methods of identifying multi-collinearity considered in this work, the Variance Inflation Factor, Correlation and the determinant of the information matrix were proved to be the best according to the order listed above from VIF as the best to the determinant of information matrix. This was because, the VIF and the Correlation revealed where the source of the little multi-collinearity in the data which was not a problem and the determinant showed that there was no multi-collinearity, since it was not a problem but condition number stated otherwise.

Recommendations

In course of this study, it was recommended to statisticians, researchers, government agencies and other agencies that;

1) Even in the absence of multi-collinearity, Ridge regression and Sub-Ridge regression can still be used in obtaining equal parameter estimates and equal sum of square errors with the Ordinary Least Square, but for k value of 0.000007 for Ridge and k value of 0.0000005 for Sub-Ridge.
2) The proposed ridge regression can be used in a data set with or without multi-collinearity for k value of 0.000005.
3) Apart from arithmetic mean, any other measurements of central tendency can be used in the Grubb's method in detecting outliers.

References

Abubakari, S. G (2019). Principal components to overcome multicollinearity problem. *Oradea Journal of Business and Economics 4* (1) 79-91.

Ayuya, C. (2021). How to detect and correct multicollinearity in regression models. *Engineerng Education community.*

Boundless (2018) Boundless algebra. *Lumen learning.com/contact*.

Deanna, S. (2018). Ridge regression and multicollinearity:An in depth review. *The Henry M Jackson foundation for advancement of military medicine*.

Iwundu, M. P. & Onu, O. H. (2017). Preferences of equiradial designs with changing axial distances, design sizes and increase center points and their relationship to the N-point central composite design. *International Journal of Advanced Statistics and Probability, 5*(2), 77-82.

Jim, F. (2020). *Regression Analysis: An Intuitive Guide or Using and Interpreting Linear Models.*Statistics by Jim Publishing.

Kutner, M. H., Nschtsheim, C. J., Neter, J. and Li, W. (2005). Applied linear statistical model, fifth edition, McGraw-Hilit:a Irwin.

Nduka, E. C., & Ijomah, M. A., (2012). The effect of perturbing eigenvalues in the presence of multi-collinearity. *Electronic Journal of Applied Statistical Analysis*, *5*(2), 304-311.

Nelson, M., Biu, O. E., & Onu, O. H. (2024). Multi-Ridge, and Inverse- Ridge Regressions for Data with or Without Multi-Collinearity for Certain Shrinkage Factors. International Journal of Applied Science and Mathematical Theory, 10(2), 29-41.

Onu, O. H., George, D. S., Uzoamaka, E. C., & Okerengwu, B. (2021). The statistical bias in genetic model analysis with varying model parameters. *International Journal of Research, 8*(6), 178-185.

Shalabh, A., Sahon, B., Kalyan, M. & Asoke, N. (2012). Design and analysis of e-journal management systems. *International Journal of Advanced Computing Sciences,* (IJACS), *8,* 913-918.

Vinod, H. & Ullah, A. (1981). *Recent advances in regression models*. Marcel Dekker.

Younker, J. (2012). Ridge estimation and its modifications for linear regression with deterministic or stochastic predictors. *Ottawa- Carleton Institute of Mathematics and Statistics.*