Original Article

GENERALIZED RIDGE REGRESSION: MULTI-RIDGE AND INVERSE-RIDGE METHODS WITH AND WITHOUT MULTICOLLINEARITY

Ebikeme Tari Emmanuel, Michael James Bennett and Emily Claire Watson

Mathematics Department, Isaac Jasper Boro College of Education, Sagbama, Bayelsa State, Nigeria

DOI:https://doi.org/10.5281/zenodo.15436344

Abstract

The study presented Mult-, and Inverse-ridge regressions for data with or without multicollinearity for certain shrinkage factors. The study considered data of GDP of Nigeria as response, while exchange, unemployment, inflation and foreign direct investment were used as the predictors. The data were tested for outlier using Grubb's test and the VIF, condition number, correlation and t-values were used to assess how the OLS and Ridge regressions were related with the proposed mult-and inverse-ridge regressions. The study revealed that whether or not, there is outlier or multicollinearity in a data set, the mult or inverse-ridge gives the same estimate of model parameters with the respective shrinkage factors of 1.000006 and 0.999999. These methods, overcame the barrier of testing for outlier or multicollinearity in a data set, it is advised that instead of testing, use any of the methods, Ridge, Sub-Ridge, Multi-Ridge and Inverse-Ridge methods with their respective shrinkage penalty. The OLS was not condemned, rather, it was used as the basis for judging these proposed methods.

Keywords: Multi-ridge, Inverse-ridge, Ridge regressions, OLS, t-values.

1. Introduction

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. 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). In fact, the Gaus Markov theorem states that OLS produces estimates that are better than estimates from all other linear model estimation methods when the assumptions hold true. Ordinary Least Squares linear regression (OLS) is one of the most commonly and oldest used approaches in multiple regression. The estimator relates the dependent variable to a set of explanatory variables. In particular, if a model is constructed from variables with mean zero, then the estimator takes the covariance between the explanatory and dependent variables $X'X$, and scales it by the inverse of the variancecovariance matrix of the explanatory variables $(X'X)^{-1}$.

Original Article

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 model that is linear or nonlinear have been a major problem to decide as many will say that if the highest power of the unknown is one, it is linear and if the highest power is two, the model is quadratic and if more than two it is polynomial. Multiple linear regression is very sensitive to predictors that are in a configuration of near collinearity. When this is the case, the model parameters become unstable (large variances) and cannot be interpreted. From a mathematical standpoint, near-collinearity makes the $X'X$ matrix ill-conditioned (with X the data matrix), that is, the value of its determinant is nearly zero, thus, attempts to calculate the inverse of the matrix result in numerical snags with uncertain final values. Exact collinearity occurs when at least one of the predictors is a linear combination of other predictors. Therefore, X is not a full rank matrix, the determinant of X is exactly zero, and inverting $X'X$ is not simply difficult, it does not exist. When multicollinearity occurs, the least squares estimates remain unbiased and efficient. The problem is that the estimated standard error of the coefficient β_i tends to be inflated. This standard error has a tendency to be larger than it would be in the absence of multicollinearity because the estimates are very sensitive to any changes in the sample observations or in the model specification. In other words, including or excluding a particular variable or certain observations may greatly change the estimated partial coefficient. If bi is larger than it should be, then the tvalue for testing the significance of β_i is smaller than it should be. Thus, it becomes more likely to conclude that a variableX_i is not important in a relationship when, in fact, it is important. The Multiplicative ridge and Inverse ridge regressions, known as Mult-ridge and Inverse ridge were proposed as regression methods used in estimating parameters. This was as a result of the fact that Ordinary Least Square (OLS) was only better when the data is free from multicollinearity and outlier. Also, the data must be normally distributed. Ridge regression was introduced to handle such problem. This proposed methods, can estimate parameters with data with or without multicollinearity and outlier. The estimates give same results for some pronounced shrinkage factors.

Regression analysis can be explained as a function between interested response variable and explanatory variables thought to be related on response (Ari & Onder, 2013). Least square method (LS) is a common method to estimate parameters in the regression model (Uckardes et al., 2012). Besides, the LS method is an unbiased method that is not only estimate parameter but also minimizing the error of the model. However, the LS method needs some assumptions which should be provided for the model reliable. If assumptions aren't provided, the reliability of the model will decrease. Therefore, it will cause misinterpretations. To guarantee the usability of this method, the assumptions must be valid such as that the errors are independent and normally distributed, and independent among explanatory variables.

Ridge regression is a technique for analyzing multiple regression data that suffer from multicollinearity. When multi-collinearity occurs, least squares estimates are unbiased, but their variances are large so they may be far from the true value. By adding a degree of bias to the regression estimates, ridge regression reduces the standard errors. It is hoped that the net effect will be to give estimates that are more reliable. Another biased regression technique, known as principal components regression, but Ridge regression is the more popular of the two methods. Many procedures have been suggested in an attempt to overcome the effects of multicollinearity in regression analysis. Hoerl and Kennard (1970) proposed a class of biased estimator called ridge regression estimators as an alternative to the OLS estimator in the presence of collinearity.

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

Original Article

(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

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

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

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

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

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

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

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

$$H, M = \frac{N}{\left(\frac{1}{Y_1} + \frac{1}{Y_2} + \frac{1}{Y_3} + \dots + \frac{1}{Y_N}\right)} \quad (7)$$

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), 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 λ, 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 $|A - \lambda I|$

Original Article

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 ($\lambda_1, \lambda_2, \ldots, \lambda_k$) is referred to as the condition index:

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

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

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

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

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \varepsilon \quad (13)$$

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, β_0, β_1, β_2, β_3 *and* β_4 are the unknown model parameters while ε is the

Original Article

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:

$$Y = X\beta + \varepsilon \qquad (14)$$

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

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. From the model in (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 (13), 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 Multiplication and Inverse based Ridge Regressions for varying Shrinkage Penalty Values.

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

Another method to be tested in this research is the multiplicative ridge regression, it is given as

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

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

Also, the inverse based ridge is given as

$$\underline{\hat{\beta}} = (X'X + (KI)^{-1})^{-1}X'\underline{Y} \qquad (19)$$

$$\underline{\hat{\beta}} = (X'X - (KI)^{-1})^{-1}X'\underline{Y} \qquad (20)$$

$$\underline{\hat{\beta}} = (X'X \times (KI)^{-1})^{-1}X'\underline{Y} \qquad (21)$$

Original Article

where, Equation (19), (20) and (21) are the Inverse Ridge regression in term of Additive, Subtractive and Multiplicative respectively

Test of Significance of Combined Regression ANOVA for k Predictor Variables Multiple Linear Regression

We present the F-test provided by the method of analysis of variance (ANOVA). For the general case of k independent variables and the test is base on the F-ratio given as

$$F = \frac{SSR/k}{SSE/(n-k-1)} = \frac{MSR}{MSE}$$

The overal variance in dependent (Y) can be splitted into

$\Sigma(y_i - \bar{y})^2 = \Sigma(\hat{y_i} - \bar{y})^2 + \Sigma(y_i - \hat{y_i})^2$

$SST = SSR + SSE$ where

SST is the total variation in dependent (Y).

SSR is the regression variation in dependent (Y).

SSE is the error (residual) variation.

They are summarized in the table below.

Table 1: Analysis of Variance (based onk Predictor Variables)

Source	*Degree of Freedom*	*Sum of Squares*	*Mean Square*	*F-Ratio*	*P-Value*
Regression	k	SSR	MSR=SSR/k	F_{obs}=MSR/MSE	
Error	n-k-1	SSE	MSE=SSE/(n-k-1)	---	
Total	n-1	SST	---	---	

The F-ratio above described above is the test statistics for the null hypothesis

H_0: b_1 = ... = b_k = 0 (*Y* is not linearly associated with and of the independent variables) H_1: Not all b_j = 0 (At least one of the independent variables is associated with dependent variable.)

From F-distribution with $v_1$and v_2 degree of freedom with selected significance level. A null hypothesis is accepted if the test statistics is less than the table value. Otherwise, null hypothesis is rejected. That is *Fobs* □ *F*□,*k*,*n*−*k*−1

P-value: Area in the *F*-distribution to the right of $F_{obs.}$

Table 2: Results of Varying Shrinkage Penalty Values on the Selected Economic Variables Data

K	**OLS**	**Ridge**	**Sub-Ridge**	**Multi-Ridge**	**Inverse-Ridge**

Original Article

0.000000 13.132 3 -0.4061 -0.0372 -2.6536 0.0327	13.132 3 -0.4061 -0.0372 -2.6536 0.0327		13.1323 0.4061 -0.0372 -2.6536 0.0327	-	0.0000 0.0000 0.0000 0.0000 0.0000		0.0000 0.0000 0.0000 0.0000 0.0000	
0.000005		tx		tx		**t_x**		**tx**
	13.132 3 -0.4061 -0.0372 -2.6536 0.0327	5.61 - 0.73 - 0.84 - 3.60 2.55	13.1323 0.4061 -0.0372 -2.6536 0.0327	-5.61 -0.73 -0.84 -3.60 2.55	2626500 0 -812000 -74000 - 5307000 65000	1122435 -1468354 -16666667 - 7.1911056 5078125	0.000006566 - 0.000000020 3 -0.000000019 -0.000001327 -0.000000016	0.0000028 - 0.00000003 7 - 0.00000004 3 -0.00018 0.000125
0.000007	13.132 3 -0.4061 -0.0372 -2.6536 0.0327	- 0.73 - 0.84 - 3.60 2.56	13.1324 0.4062 -0.0372 -2.6536 0.00327	- -0.73 -0.84 -3.60 2.56	1876000 -58000 -5300 379100 4700	801709.4 04882.5 --119369.4 513685.6 367187.5	9.0926 1.2986 -0.0346 2.1647 0.0197	3.89 -2.35 0.78 -2.93 1.54
0.00005	13.131 9 -0.4061 -0.0372 -2.6535	5.61 - 0.73 - 0.83	13.1328 0.4062 -0.0372 -2.6537 0.0327	- 5.61 -0.73 -0.84 -3.60 2.55	262650 -8120 -740 53070 650	112243.6 --14683.5 -16666.7 -71910.6 50781.3	0.0006566 -0.0000203 -0.000019 0.0001327 0.0000016	0.00028 --0.000038 -0.000043 -0.00018 0.000125
0.0327 3.60 2.55	-							
0.99999	7.8091	3.34	44.3446		18.95	13.1325	5.61 13.1322	5.61

0.3414		0.62	-4.9423		-8.99	-0.4061		-0.73	-0.4061		-0.73	
-0.0566			0.0870		1.96	-0.0372		-0.84	-0.0372		-0.84	
1.27												
-1.3887			-13.83	-2.6536		-3.60	-2.6535		-3.60			
-	10.0989											
0.0213			7.88	0.0327		2.55	0.0327		2.55			
1.88												
0.1008												
1.66												
0.999999		7.8091		3.34	44.3455		18.95	13.1324		5.61	13.1323	5.61
0.3414		0.62	-4.9425		1.70	-0.4061		-0.73	-0.4061		-0.73	
-0.0566			0.0870		1.96	-0.0372		-0.84	-0.0372		-0.84	
-1.3887		1.27		-13.68	-2.6536		-3.60	-2.6536		-3.60		
-	10.0991											
0.0213			7.88	0.0327		2.55	0.0327		2.55			
1.88												
0.1008												
1.66												
1.000006		7.8091		3.34	44.3463		18.95	13.1323		561	13.1324	5.6
0.3414		0.62	-4.9426		-8.94	-0.4061		-0.73	-0.4061		-0.73	
-0.0566			0.0870		2.00	-0.0372		-0.84	-0.0372		-0.84	
1.27												
-1.3887			-13.68	-2.6536		-3.60	-2.6536		-360			
-	10.0993											
0.0213			7.88	-0.00327		2.56	0.0327		2.55			
1.88												
0.1008												
1.66												

Table 3: Comparison of known (Additive) and Multi-Ridge Results of Varying Shrinkage Penalty Values on the Selected Economic Variables Data.

K	OLS	Ridge	K	Multi-Ridge

Original Article

	0.000000	13.132 3	13.132 3		0.999999	13.1325 5.61	
-0.4061	-0.73 -0.4061 -0.0372 -2.6536 0.0327		-0.4061 -0.0372 -2.6536 0.0327			-0.0372 -2.6536 0.0327	-0.84 -3.60 2.55
	0.000005		13.132 3 -0.4061 -0.0372 -2.6536 0.0327	**tx** 5.61 - 0.73 - 0.84 - 3.60 2.55	1.000006	13.1323 -0.4061 -0.0372 -2.6536 -0.00327	**t_x** 5.61 -0.73 -0.84 -3.60 2.56
	0.000007		13.132 3 -0.4061 -0.0372 -2.6536 0.0327	- 0.73 - 0.84 - 3.60 2.56	1.000009	13.1322 0.4061 -0.0372 -2.6535 0.0327	-561 -0.73 -0.84 -3.60 2.56
	0.00005		13.131 9 -0.4061 -0.0372 -2.6535 0.0327	5.61 - 0.73 - 0.83 - 3.60 2.55	1.00001	13.1322 0.4061 -0.0372 -2.6535 0.0327	--0.73 -0.84 -3.60 2.56

0.00008	13.131 6 -0.4060 -0.0372 -2.6534 0.0327	- 0.73 - 0.84 - 3.60 2.56	1.00002	13.1321 0.4061 -0.0372 -2.6535 0.0327	- -0.73 -0.84 -3.60 2.56
0.0005	13.127 8 -0.4055 -0.0372 -2.6525 0.0327	5.61 - 0.73 - 0.83 - 3.59 2.55	1.00003	13.1319 -0.4061 -0.0372 -2.6535 0.0327	-0.73 -0.84 -3.60 2.56
0.0009	13.124 2 -0.4050 -0.0373 -2.6516 0.0327	- 0.73 - 0.84 - 3.59 2.55	1.00008	13.1313 0.4061 -0.0372 -2.6534 0.0327	- -0.73 -0.84 -3.60 2.56
0.005	13.087 0 -0.3997 -0.0374 -2.6428 0.0326	5.59 - 0.72 - 0.84 - 3.58	1.00009	13.1312 0.4061 -0.0372 -2.6533 0.0327 2.55	- -0.73 -0.84 -3.60 2.56

Discussion of Results

The results in table 2 revealed that for the shrinkage factor k=0.000000, the OLS, Ridge and Subridge regressions have equal coefficients, while the proposed Mult-Ridge and Inverse-Ridge have coefficients of zeros all through. For the shrinkage factor k=0.000005, the Ridge and Sub-Ridge regressions maintained the same coefficients as with the shrinkage of k=0.000000, while, the MultRidge and Inverse-Ridge regressions parameters increased out of bound, likewise its t-values. As the shrinkage factor k increased to k=0.000007, the Ridge and Sub-Ridge regressions parameters differed in the gradients of exchange rate from -0.4061 to -0.4062, also, they differed in the gradients of the FDI, while, the Ridge regression was equal to the OLS for the shrinkage k=0.000007. Hence, k=0.000007 was proposed as the shrinkage penalty for Ridge regression. That is to say, whether the data has multicollinearity or not, the Ridge regression will have equal estimation of parameters with the OLS when the shrinkage penalty k of the Ridge regression is 0.000007. This serves as a big advantage to researchers, especially those that deal on big data, because, needless of testing for multicollinearity, as it could be time consuming. As the k increases further, the Ridge and the Sub-Ridge estimates continue to differ, even more visibly away from the estimates of the OLS, while the Mult-Ridge and Inverse-Ridge estimates become closer to each other and tending towards the OLS estimates. As the k increased to k=0.999999, the Inverse-Ridge estimates became equal to the OLS, and the value k=0.999999 was proposed as the shrinkage penalty for Inverse-Ridge. As k increased to 1.000006, the Mult-Ridge estimates became equal to the OLS, as such, k=1.000006 was proposed as the shrinkage penalty for Mult-Ridge regression. The implications of these findings, was to provide diverse method of solving regression problem instead of using just OLS when there is no multicollinearity in the data or using Ridge when there is multicollinearity in the data set(s). Also, the methods, overcomes the barrier of testing for multicollinearity in a data, instead, use any of the methods, Ridge, Sub-Ridge, Multi-Ridge and Inverse-Ridge methods with their respective shrinkage penalty. The OLS was not condemned, rather, it was used as the basis for judging these methods.

Conclusion

The study concludes that multiplicative (mult-ridge) and Inverse-ridge regression methods should be applied with 1.000006 and 0.999999 shrinkage penalty respectively in order to overcome the extra work of testing the data for outlier and multicollinearity. With the proposed shrinkage value, the proposed methods yield same results with the OLS, whether with or without outlier or multicollinearity.

Contribution to knowledge

The study was able to show that, needless of testing for outlier or multicollinearity in a data set, when trying to use regression approach in estimating parameters of a model. Instead, used either of mult-ridge with 1.000006 shrinkage penalty or inverse-ridge with 0.999999 shrinkage penalty.

References

Arı, A.& Hasan Önder (2013). Regression models used for different data structures. *Journal of the Institute of Science and Technology,28 (3) 168-174*.

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*.

Hoerl, A.E. and Kennard, R.W. (1970). Ridge regression: Application to nonorthogonal problems. *Technometrics, 12*, 69-82.

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.

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

Original Article

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.

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.

Thompson, C. G., Kim, R. S., Aloe, A. M. &Becker, B. J. (2017) Extracting the variance inflation factor and other multicollinearity diagnostics from typical regression results. *Basic and Applied Social Psychology,39* (2) 81-90.

Uckarde, F., Efe, E., Narinc, D. & Aksoy, T. (2012). Estimation of the albumen index in Japan quails with ridge regression method. *Academic Journal of Agriculture, 1* (1) 12-20. Vinod, H. & Ullah, A. (1981). *Recent advances in regression models*. Marcel Dekker.