DOI: https://doi.org/10.26682/sjuod.2020.23.2.2

Journal of University of Duhok, Vol. 23, No.2 (Pure and Eng. Sciences), Pp 9-24, 2020

ESTIMATE THE PARAMETERS IN PRESENCE OF MULTICOLLINEARITY AND OUTLIERS USING BISQUARE WEIGHTED RIDGE LEAST MEDIAN SQUARES REGRESSION (WRLMS)

KAFI DANO PATI

College of Science, University of Duhok, Duhok, Kurdistan Region, Iraq –

Dept Of Computer Science, College of Science, University of Duhok, Kurdistan Region–Iraq

(Received: March 8, 2020; Accepted for Publication: September 29, 2020)

ABSTRACT

The presence of multicollinearity and outliers are classical problems of data within the linear regression framework. We are going to present a proposal of a new method which can be a potential candidate for robust ridge regression as well as a robust detection of multicollinearity. This proposal arises as a logical combination of principles used in the ridge regression and the Bisquare weighted function. The technique of the Least Median of Squares (LMS) is used for the sake of overcoming the resulting regression problems. This paper investigates the non-resistance of Ordinary Least Square (OLS) to multicollinearity and outliers and proposes the utilization of robust regression for instance, Least Median Squares LMS to detect non-normality of residuals, the use of robust methods yields more reliable trend estimations and outlier detection. LMS is introduced as a robust regression technique and through medical application its effect on regression is discussed. The numerical example and simulation study shows that the outcome of the Weighted Ridge Least Median Squares (WRLMS) is better than other estimators in terms of its efficiency. This has been done by utilizing both Standard Error (SE) and the Root Mean Squared Error criterion for the numerical example and simulation study, respectively as far as a lot of combinations of error distribution and degree of multicollinearity are concerned.

KEYWORDS: Multicollinearity; Outliers; Ridge Regression; Robust LMS Estimation and Weighted Ridge Least Median Squares.

1. INTRODUCTION

Multicollinearity and outliers are two important problems considered in regression analysis. The coefficients' OLS estimators usually have particular ideal properties if explanatory variables are not intercorrelated, and the unsetteling influences of the regression equation are independent, indistinguishably allocated typical random variables.

The existence of intercorrelationship may be conducive to wrong information regarding the coefficients of regression. On the other hand, in the presence of nonnormal disturbance distributions the OLS estimator may produce extremely poor estimates.

Consequently, to address these problems independently, various remedial techniques have been suggested. One of these techniques is the ridge regression technique which is used to tackle multicollinearity, as well as the robust estimation techniques which are not as barely influenced by nonnormal disturbances. Be that as it may, in spite of the fact that we by and large think of these two problems independently, but in practical circumstances, they happen at the same time.

Montgomery and Peck (1982) have suggested that it is likely that either robust or ridge estimation methods alone are appropriate for managing the combined problem. Therefore, in order to handle these two problems at the same time, some robust ridge regression estimators have been introduced; estimators that are much less influenced by multicollinearity and outliers.

Askin and Montgomery (1980), on the other hand, have suggested that the ridge and the Least Absolute Deviation (LAD) robust regression techniques be joined together. In the current paper, as a starting point, there is a need to create a more robust technique to deal with these two problems. The suggestion being made is that the ridge regression and the weighted function be combined with the high breakdown point estimator, which is the LMS-estimator.

kafi.pati@uod.ac

This modified method is called in this study, the weighted robust ridge regression on the basis of the LMS-estimation WRLMS.

It is anticipated that this modified method would be less susceptible to the multicollinearity and outliers presence. Thus, this paper aims to test some estimators which resist the problems of multicollinearity and outliers when these problems are of a combined nature. So the question that may be put is whether it is possible to combine the ridge estimators and some robust estimation techniques in order to come out with a robust ridge regression estimator. The outline of the present paper is as follows: section (2) explains the ridge estimator; section (3) clarifies the seek for the robust estimation techniques; sections (4) and (5) explain the alternative combined estimators of ridge and weighted robust regression; section (6) is an application of the suggested method and available methods; section (7) presents the results and discussion of a Monte Carlo simulation study to examine the performance of such estimators; and lastly, section (8) gives the conclusions.

2. THE ESTIMATORS OF RIDGE REGRESSION

See this model of linear regression:

$$y = X\beta + \varepsilon, \tag{1}$$

here y is an (*nx1*) vector of observations relevant to the dependent variable, X is an (*n x p*) matrix of observations pertinent to the explanatory variables, β is a (*px1*) vector of regression coefficients that need estimation, and ε is (*nx1*) vector of disturbances distribution. It is possible to write the LS estimator of β as:

$$\hat{\beta}_{LS} = \left(X'X\right)^{-1} X'Y \tag{2}$$

The advantage of the present method is that it gives a variance which is unbiased and minimum amidst all unbiased linear estimators on the condition that the errors are independent as well as normally distributed in an identical manner. But when multicollinearity is present, the singularities which exist in $(X'X)$ matrix can make this ill-conditioned X matrix liable to the production of bad estimates.

Ridge regression use LS estimator because of it is slightly biased. The idea is that a biased estimator with a small standard error is often preferable over an unbiased estimator with a large standard error. Ridge regression uses the correlation transformation along with a biasing constant to obtain the ridge estimators for the transformed model (see Kutner et al., 2004). The biasing constant is most commonly chosen based on the Variance Inflation Factor (VIF). The VIF's will also decrease and eventually stabilize as the biasing constant increases. So, the "best" biasing constant is the one whose value stabilizes the VIF's. The degree of multicollinearity is often specified by VIF

Hoerl and Kennard (1970a) designed a method for ridge regression for the purpose of estimating a proper parameter where it is possible to add a constant k i to the $\mathbf{X'X}$. In the prsent case, the following gives the ridge estimator of the parameter β:

$$\hat{\beta}_{Rid} = \left(X'X + kI\right)^{-1} X'Y, \tag{3}$$

here $0 < k < 1$ is the is the biasing constant, $\mathbf{I}$ is the *pxp* identity matrix and *p* is the number of parameters. In accordance with the researcher's judgment, the constant k which relies on a trace would be established on sunjective bases (Hoerl and Kennard, 1970b) and (Gibbons, 1981) as mentioned in (Montgomery and Peck, 1992). Various methods for the determination of k value have been mentioned in the literature. Consider, for example the following by Hoerl and Kennard (1970):

$$k_{HK} = \frac{p S^2_{LS}}{\hat{\boldsymbol{\beta}}'_{LS}\hat{\boldsymbol{\beta}}_{LS}} \tag{4}$$

where $S^2_{LS} = \dfrac{\left(\mathbf{y}\text{-}\mathbf{x}\hat{\boldsymbol{\beta}}_{LS}\right)'\left(\mathbf{y}\text{-}\mathbf{x}\hat{\boldsymbol{\beta}}_{LS}\right)}{\text{n-p}}$ (5)

where the ridge regression procedure of β is utilized instead of the LS estimator in the computation of the k and S^2 values so as to minimize the influence of nonnormality on the value, then $\hat{\boldsymbol{\beta}}_{Rid}$ shows bias; however, it is more steady and exact in comparison to the LS estimator and when $k \to \infty, \hat{\boldsymbol{\beta}}_{Rid} \to 0$. Hoerl and Kennard (1970) have stated that a value $k > 0$ always exists in such away that $MSE(\hat{\boldsymbol{\beta}}_{Rid})$ becomes less than $MSE(\hat{\beta}_{LS})$.

3. THE ESTIMATORS OF ROBUST REGRESSION

It has been proven that Robust regression estimators are more effective and solid than LS estimator in two cases in particular: when disturbances are nonnormal and when disturbance distributions that possess heavy or fatter tails in comparison to the normal distribution are thus susptible to the production of outliers. In view of the fact that outliers enormously influence the estimated coefficients, standard errors, and test statistics, then there is a possibility that the regular statistical procedur be most ineffective because the precision of the estimator has been influenced. There exists lot of various classifications of robust regression. The LMS estimator is one of the important members of regression.

It is possible to define the LMS estimator $\hat{\beta}_{LMS}$ as the solution to the following minimization problem:

$$med\ (y_i - x_i' \beta)^2 \quad (6)$$

Here the sum of the absolute values of the residuals is minimized instead of minimizing the sum of squared residuals as in LS estimation. In this way, the impact of outliers on the LMS estimates becomes less than the one on LS estimates. It is possible to use the procedure of Bisquare weighted function for the sake of computing the weighted LMS estimates as follows:

To calculate the weights, consider a situation with *n* observations of y_i a response variable and *p* explanatory variables.

$$\mathbf{y_i} = \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + ... + \beta_p x_{pi} + \boldsymbol{\varepsilon_i} \quad (7)$$

We adopt the following procedure for all competing estimators being considered in the study.

Step 1: Choose initial estimates of β, we use robust estimator as the starting value.

Step 2: Compute the residuals $\hat{\varepsilon}_i = y_i - x_i \hat{\beta}$ and also the corresponding weights based on the weight function from the previous iteration. Usually the residuals are scaled by applying some suitable scale estimate *u*. In practice the variance σ^2 is unknown. A very good choice for scale estimate due to a high degree of robustness is the Median Absolute Deviation (MAD) and is given by

$$\text{MAD} = \frac{1}{0.6745} \text{median} \left\{ \left| \varepsilon_i - \text{median} \left\{ \varepsilon_i \right\} \right| \right\} \quad (8)$$

Step 3: Calculate new estimates of the regression coefficients by performing weighted function, that scaled residual is $u = \frac{\varepsilon_i}{MAD}$ many weight functions have been proposed for dampening the influence of outlying cases or down weight the influence of outliers.

$$w_B(u) = \begin{cases} \left[1 - \left(\frac{u}{c} \right)^2 \right]^2 & |u| \le c \\ 0 & |u| > c \end{cases} \quad (9)$$

As before, w_B denotes the Bisquare weighted function, and u denotes the scaled residual to be defined shortly. The Bisquare estimators' c value is known as a tuning constant. If c exhibits smaller values, more resistance to outliers is brought about, but this happens at the cost of lower efficiency when there is a regular distribution of errors.

The tuning constant is commonly chosen in order to provide high efficiency which is reasonable under normal conditions; especially, c = 4.685 if the Bisquare is to produce 95% efficiency when the errors are normal, and still offer protection against outliers (Fox and Weisberg (2011); (Neter, 2004) and (Alaminger, 2013).

4. ROBUST RIDGE REGRESSION ESTIMATORS

This is a combination of ridge and robust regression to handle the multicollinearity and outliers problems at the same time. In this way, the effects of both problems will be dampen in a classical linear regression model. The formula for the comoutation of Robust Ridge Estimator runs as follows:

$$\hat{\boldsymbol{\beta}}_{Robustridge} = \left(X'X + k_{Robust} I \right)^{-1} X'Y \quad (10)$$

Where k_{Robust} is called the robust parameter. It is obtained from robust regression methods instead of using OLS. This will be computed as given in equation (4) and (5), only that k_{HK} and S^2_{LS} are replaced with k_{Robust} and S^2_{Robust} respectively.

5. WEIGHTED ROBUST RIDGE REGRESSION ESTIMATORS

In the current section, some combinations of ridge and robust regression estimation are presented. These were discussed in sections (2) and (3), respectively with weight function. In this regard, it is possible to compute the robust ridge estimator, known as weighted robust ridge estimator $\hat{\beta}_{Wrobustridge}$, by utilizing the following formula:

$$\hat{\boldsymbol{\beta}}_{Wrobustridge} = \left(X'WX + k_{WR}I\right)^{-1} X'WY \qquad (11)$$

where the value of $k_{WRobust}$ is identified as

$$k_{WRobust} = \frac{ps^2_{WRobust}}{\hat{\boldsymbol{\beta}}'_{Robust}\hat{\boldsymbol{\beta}}_{Robust}},$$ and where

$$s^2_{WRobust} = \frac{\left(\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}_{WRobust}\right)'\left(\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}_{WRobust}\right)}{n-p} \qquad (12)$$

the weighted w_B can be determined from equation (9). Moreover, there is another weighted ridge estimator which was formerly utilized by another researcher; it is called weighted ridge estimator $\hat{\boldsymbol{\beta}}_{WRID}$ The following formula computes this estimator:

$$\hat{\boldsymbol{\beta}}_{WRID} = \left(\mathbf{X'WX} + k\mathbf{I}\right)^{-1} \mathbf{X'WY} \qquad (13)$$

here the values of k and s^2 are identified as (4) and (5), respectively. The same values of weighted in equation (9) used to estimate (Zahari et al., 2012) and compared with this proposed estimator LMS.

The estimator of weighted robust ridge estimator will be designated by the WRLMS estimator and it is possible to write it as

$$\hat{\boldsymbol{\beta}}_{WRLMS} = \left(X'WX + k_{WLMS}I\right)^{-1} X'WY \qquad (14)$$

where the value of k_{WLMS} is ascertained on the basis of data by using:

$$k_{WLMS} = \frac{PS^2_{WLMS}}{\hat{\boldsymbol{\beta}}'_{WLMS}\hat{\boldsymbol{\beta}}_{WLMS}} \qquad (15)$$

and

$$S^2_{WLMS} = \frac{\left(Y - X\hat{\boldsymbol{\beta}}_{WLMS}\right)'\left(Y - X\hat{\boldsymbol{\beta}}_{WLMS}\right)}{n-p} \qquad (16)$$

$\hat{\beta}_{WRLMS}$ represents the estimator of RLMS which can be defined as equation (14) solution. It is important to note that the value of k_{WLMS} represents the estimator of k_{HK} shown by equation (4) with the presence of two changes.

This process was firstly applied instead of the LS estimate. Afterwards, the estimator applied in equation (5) is modified by other estimates, namely, s^2_{WLMS} which were also applied instead of the LS estimate. The aim behind these changes is to minimize the impact of extreme points on the value selected for the biasing parameter. In the end, it is possible to exploit the ridge regression estimator in order to decide on k biasing parameter.

6. APPLICATIONS

6.1 Numerical Example

Body fat data, containing 252 observations taken from Penrose et al. (1985), are exploited to assess the estimators' performance. The data involved fourteen explanatory variables. The response variable is y= PCTBF (percentage of body fat), which were regressed to the x1= Density, x2= Age, x3= Weight, x4= Height, x5= Neck, x6= Chest, x7=Abdomen, x8= Hip, x9= Thigh, x10= Knee, x11= Ankle, x12= Biceps, x13= Forearm and x14= Wrist.

Table 1 presents the VIF of the parameters, this condition can be used to detect multicollinearity using the following formula, $\text{VIF} = \frac{1}{1-R^2}$, where R^2 is the determinant of the matrix $\mathbf{X'X}$. Similarly, it is possible to determine the outliers in the data through the computation of the residuals connected with LMS regression, $s = 1.4826\left(1+\frac{5}{n-p}\right)\sqrt{med\left(\varepsilon_i^2\right)}$, where med is the median of the squared residuals, p is the number of predictors. The points $\left(y_i, x_{i1}, ..., x_{ip}\right)$ are designated as regression outliers if the equivalent standardized residual is large.

Rousseeuw and Van Zomeren (1990) designated the i^{th} vector as regression outlier if $|r_i|/s > 2.5$ suggests that the value is outlier. The most frequently used measures for the detection of outliers are ordinary or simple residuals (observed - predicted values).

kafi.pati@uod.ac

Table (1): Variance Inflation Factor VIF for the body fat data set

Var.	X1	X2	X3	X4	X5	X6	X7
VIF	3.8183	2.2747	**34.0317**	1.6778	4.3965	9.4722	**18.1199**
Var.	**X8**	**X9**	**X10**	**X11**	**X12**	**X13**	**X14**
VIF	**14.9610**	7.8877	4.6123	1.9200	3.6516	2.2370	3.5215

The results of the analysis in Table 1 indicate that when the real data belonged to the body fat application note from which the VIF for some variables value has exceeded the value of 10 since the computation of VIF is highly dependent on the calculation of R^2. So it is clear that multicollinearity problem exists between variables.

Thus, the VIF can help to identify which explanatory variables are involved in the multicollinearity, as the maximum value is 34.0317 in X3 variable in Table 1. Additionally, the data comprised 13 outliers which influenced the data and produced poor results. On the other hand, Table 2 indicates that the proposed methods when applied statistical criteria standard errors for this data indicate that, the standard error of WRLMS less than existing methods (OLS, RIDGE, RLMS and WRID).
Table 2 below su

mmarizes estimates of the parameter as well as the standard error for the methods in question, i.e., the proposed and existing ones.

Table (2): Estimation of the parameters and standard error SE of $\hat{\beta}_1 \ldots \hat{\beta}_{14}$ with respect to the proposed and existing methods WRLMS estimators used weighted Tukey function for the body fat data set.

Coef.	Estimate	OLS	RIDGE	RLMS	WRID	WRLMS
$\hat{\beta}_1$	parameter	-411.2000	-0.9296	-0.9301	-1.0027	-1.0038
	S.E.	8.2580	0.0203	0.0197	0.0057	**0.0026**
$\hat{\beta}_2$	parameter	0.0126	0.0194	0.0193	0.0006	0.0000
	S.E.	0.0158	0.0157	0.0143	0.0041	**0.0019**
$\hat{\beta}_3$	parameter	0.0101	0.0321	0.0323	0.0077	0.0011
	S.E.	0.0578	0.0575	0.0528	0.0165	**0.0073**
$\hat{\beta}_4$	parameter	-0.0080	-0.0036	-0.0035	-0.0052	-0.0045
	S.E.	0.0284	0.0135	0.0131	0.0038	**0.0020**
$\hat{\beta}_5$	parameter	-0.0285	-0.0088	-0.0088	0.0021	0.0015
	S.E.	0.0694	0.0219	0.0212	0.0060	**0.0026**
$\hat{\beta}_6$	parameter	0.0268	0.0269	0.0269	0.0002	0.0042
	S.E.	0.0307	0.0316	0.0294	0.0090	**0.0039**
$\hat{\beta}_7$	parameter	0.0186	0.0312	0.0305	-0.0025	0.0021
	S.E.	0.0433	0.0431	0.0395	0.0117	**0.0053**
$\hat{\beta}_8$	parameter	0.0192	0.0155	0.0156	0.0099	0.0114
	S.E.	0.0434	0.0392	0.0381	0.0120	**0.0049**
$\hat{\beta}_9$	parameter	-0.0168	-0.0094	-0.0095	-0.0138	-0.0121
	S.E.	0.0430	0.0292	0.0283	0.0081	**0.0035**
$\hat{\beta}_{10}$	parameter	-0.0046	-0.0012	-0.0012	0.0017	-0.0004
	S.E.	0.0716	0.0224	0.0217	0.0064	**0.0027**
β_{11}	parameter	-0.0857	-0.0169	-0.0170	0.0004	0.0015
	S.E.	0.0658	0.0146	0.0141	0.0037	**0.0020**
β_{12}	parameter	-0.0551	-0.0192	-0.0192	-0.0004	-0.0017
	S.E.	0.0509	0.0200	0.0194	0.0052	**0.0024**
β_{13}	parameter	0.0339	0.0087	0.0087	-0.0009	0.0004
	S.E.	0.0595	0.0157	0.0152	0.0040	**0.0020**
β_{14}	parameter	0.0073	-0.0002	-0.0001	-0.0015	-0.0018
	S.E.	0.1617	0.0197	0.0191	0.0054	**0.0024**

The results of numerical example in Table 2 shown as the standard error SE for the WRID less than OLS, RIDGE and RLMS. Likewise, the SE for the RLMS less than OLS, RIDGE the reason of this is due to the presence of autocorrelation and outliers in the data. It is obvious that the OLS affected if there is a single observation of outlier and presence of multicollinearity in the data and also another estimators affected in the presence of these problems. On the other, noticed that SE for the WRLMS less than WRID for all the parameters.

6.2 Simulation Study

A Monte Carlo simulation study was designed for the sake of comparing the performance of some alternative combined estimators in question to confirm the results that acquired in numerical example. The simulation is developed in such a way as to permit multicollinearity and outliers at the same time. Multicollinearity various degrees are stated. In the same way, the distributions which are not normal are utilized to produce outliers.

The five estimators involved in the study are:

(1) Ordinary least squares (OLS).
(2) Ridge regression (RIDGE).
(3) Weighted ridge regression (WRID).
(4) Ridge least median squares (RLMS).
(5) Weighted robust ridge LMS (WRLMS).

The OLS, RIDGE, WRID, RLMS and WRLMS estimators were defined above. If one assumes that the following linear regression model is available:

$$\mathbf{y_i} = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \boldsymbol{\varepsilon}_\mathbf{i} \quad \text{where } i = 1, 2, \ldots, n \qquad (17)$$

The parameter values $\beta_0, \beta_1, \beta_2 \; and \; \beta_3$ are put equal to one (Dempster, 1977). Then, the

kafi.pati@uod.ac

explanatory variables x_{i1}, x_{i2} and x_{i3} are generated as:

$$\mathbf{x_{ij}} = (1-\rho^2)\mathbf{z_{ij}} + \rho\mathbf{z_{ij}} \quad \text{where} \quad i=1, 2,\ldots, n, \quad j=1, 2, 3 \qquad (18)$$

where $\mathbf{z_{ij}}$ are independent standard normal random numbers generated by the R-normal. ρ constitutes the correlation between the two explanatory variables. Its values were selected as: 0.0, 0.5 and 0.99. For a given sample size *n*, the explanatory variables values were generated earlier. In this study, the sample sizes investigated are: 25, 50, and 100 and the fundamental factor is the disturbance distribution. The following three disturbance distributions are used:

1. Standard normal distribution.
2. Cauchy distribution along with median zero and scale parameter one.
3. Student-t distribution along with three degrees of freedom.

Generally speaking, all the acquired random numbers are generated by exploiting the R programming. Disturbances are generated in all cases in isolation from the explanatory variables. The simulations that were executed used programs written in R.

Three different distributional forms for the ε_i disturbances are considered, to statement an inquiry of model performance in a broad range of circumstances. The distributional forms are:

The results used are based on 500 Mote Carlo trials. Consider these statistics:

1. The mean squared error is $MSE = \frac{1}{500}\sum_{i=1}^{500}\left(\hat{\boldsymbol{\beta}}_{\mathbf{j}}\text{-}\boldsymbol{\beta}_{\mathbf{j}}\right)^2$, therefore,
2. The RMSE is given by $[MSE(\hat{\boldsymbol{\beta}}_{\mathbf{j}})]^{1/2}$

   where j= 0, 1, 2, 3

7. RESULTS AND DISCUSSION

The comparison of the three robust ridge estimators WRID, RLMS, and WRLMS are considered with two non-robust which are ridge regression and OLS estimate.

The results from the Table 3 and Table 9 we see that the value of the RMSE for the OLS estimator is less than all the estimators when no multicollinearity and no outliers in the data and the disturbances are normal and S-student. Otherwise the RMSE for OLS is greater than estimators used when the disturbance is cauchy distribution.

Whilst Tables 4-11 presents the value of the RMSE for the proposed method WRLMS is less than all estimators for different sample size and different degree of multicollinearity and outliers estimates except Table 3 and Table 9 the value of the RMSE of the RLMS greater than OLS when the no outliers and no multicollinearity for the sample size 25, 50 and 100 when the error are normal and S-student disturbance distribution and However, the value of the RMSE of the RLMS less than all estimator when the no outliers and no multicollinearity for the sample size 25, 50 and 100 when the disturbance is cauchy and S-student distribution. Occasionally, the value of the RMSE of the RLMS less than WRID when the ρ =0.5 and 20% of outliers, also when ρ =0.99 10% and 20% for all disturbance distribution.

As a conclusion and taking into consideration the different estimators' properties, it should be noted that the results obtained from comparing the WRLMS estimator with the WRID, RLMS, RIDGE and LS estimators are not totally unanticipated. Accordingly, the most fundamental result acquired from these comparisons is that the WRLMS estimator is better than the RLMS estimator. This is evident over a wide range of estimators' values for the disturbance distributions in question as the ridge regression, and in some cases, it is even likely to work well.

Table (3): RMSE values for the parameters $\hat{\beta}_1$, $\hat{\beta}_2$ and $\hat{\beta}_3$ of the Bisquare weighted function led to the generation of Normal disturbance distribution, with different sample sizes and 0.0 of ρ and 0% of outliers.

Collinearity and Outliers	Sample size	Par.	OLS	RIDGE	RLMS	WRID
	25	$\hat{\beta}_1$	**0.2268**	0.5175	0.5065	0.5386
		$\hat{\beta}_2$	**0.2211**	0.5147	0.5128	0.5330
		$\hat{\beta}_3$	**0.2288**	0.5246	0.5172	0.5430
		$\hat{\beta}_1$	**0.1469**	0.5067	0.5035	0.5056
ρ =0.0 outliers=0%	50	$\hat{\beta}_2$	**0.1438**	0.5138	0.5080	0.5103
		$\hat{\beta}_3$	**0.1526**	0.5076	0.5025	0.5156
		$\hat{\beta}_1$	**0.0996**	0.5047	0.5033	0.5056
	100	$\hat{\beta}_2$	**0.1092**	0.5037	0.5008	0.5043
		$\hat{\beta}_3$	**0.1028**	0.5022	0.4996	0.5087

Table (4): RMSE values for the parameters $\hat{\beta}_1$, $\hat{\beta}_2$ and $\hat{\beta}_3$ of the Bisquare weighted function led to the generation of Normal disturbance distribution, with different sample sizes and 0.5 of ρ and 5%, 10%, and 20% of outliers.

Collinearity and Outliers	Sample size	Par.	OLS	RIDGE	RLMS	WRID	WRLMS
		$\hat{\beta}_1$	7.6662	1.0146	0.9802	0.9589	**0.9567**
	25	$\hat{\beta}_2$	7.4358	1.0134	0.9777	0.9589	**0.9585**
		$\hat{\beta}_3$	7.7912	1.0091	0.9736	0.9585	**0.9571**
		$\hat{\beta}_1$	4.9159	0.9816	0.9697	0.9575	**0.9567**
ρ =0.5 outliers=5%	50	$\hat{\beta}_2$	4.9135	0.9834	0.9682	0.9576	**0.9570**
		$\hat{\beta}_3$	5.0025	0.9705	0.9678	0.9573	**0.9566**
		$\hat{\beta}_1$	3.6112	0.9822	0.9697	0.9615	**0.9598**
	100	$\hat{\beta}_2$	3.7634	0.9698	0.9656	0.9617	**0.9613**
		$\hat{\beta}_3$	3.7729	0.9796	0.9711	0.9617	**0.9613**
		$\hat{\beta}_1$	10.3840	1.0206	0.9835	0.9712	**0.9689**
	25	$\hat{\beta}_2$	10.3675	1.0060	0.9810	0.9705	**0.9699**
		$\hat{\beta}_3$	10.3284	1.0302	0.9812	0.9745	**0.9683**
		$\hat{\beta}_1$	7.5537	1.0007	0.9797	0.9728	**0.9715**
ρ =0.5	50	$\hat{\beta}_2$	7.8356	0.9997	0.9848	0.9746	**0.9719**

kafi.pati@uod.ac

outliers=10%		$\hat{\beta}_3$	7.1996	0.9905	0.9816	0.9751	**0.9716**
		$\hat{\beta}_1$	5.2223	0.9923	0.9828	0.9754	**0.9710**
	100	$\hat{\beta}_2$	5.1746	0.9867	0.9804	0.9803	**0.9715**
		$\hat{\beta}_3$	5.1154	0.9838	0.9771	0.9723	**0.9716**
		$\hat{\beta}_1$	14.6464	1.0322	0.9882	1.0240	**0.9790**
	25	$\hat{\beta}_2$	15.6775	1.0046	0.9883	0.9925	**0.9791**
		$\hat{\beta}_3$	15.1694	1.0259	0.9888	1.0218	**0.9803**
		$\hat{\beta}_1$	10.2780	1.0091	0.9874	1.0036	**0.9782**
ρ=0.5 outliers=20%	**50**	$\hat{\beta}_2$	10.2605	1.0029	0.9884	0.9878	**0.9798**
		$\hat{\beta}_3$	9.8772	1.0007	0.9883	0.9951	**0.9775**
		$\hat{\beta}_1$	6.7804	1.0015	0.9885	0.9923	**0.9784**
	100	$\hat{\beta}_2$	6.9266	0.9856	0.9831	0.9856	**0.9787**
		$\hat{\beta}_3$	6.8933	0.9921	0.9873	0.9829	**0.9794**

Table (5): RMSE values for the parameters $\hat{\beta}_1$, $\hat{\beta}_2$ and $\hat{\beta}_3$ of the Bisquare weighted function led to the generation of Normal disturbance distribution, with different sample sizes and 0.99 of ρ and 5%, 10%, and 20% of outliers.

Collinearity and Outliers	Sample size	Par.	OLS	RIDGE	RLMS	WRID	WRLMS
		$\hat{\beta}_1$	36.2995	16.9471	2.2958	2.0718	**1.7265**
	25	$\hat{\beta}_2$	36.5225	17.0882	2.1072	2.2286	**1.8383**
		$\hat{\beta}_3$	37.0392	17.2163	2.1122	2.1627	**1.7050**
		$\hat{\beta}_1$	23.1791	11.4812	2.9258	1.6163	**1.2125**
ρ=0.99 outliers=5%	**50**	$\hat{\beta}_2$	23.3869	11.4224	3.0880	1.6619	**1.2753**
		$\hat{\beta}_3$	24.3679	11.9025	2.9641	1.6373	**1.2263**
		$\hat{\beta}_1$	17.2393	7.7543	3.0266	2.0837	**1.0251**
	100	$\hat{\beta}_2$	17.7965	7.9083	2.8387	2.1194	**1.0894**
		$\hat{\beta}_3$	18.3800	8.1809	3.0398	2.1017	**1.0976**
		$\hat{\beta}_1$	48.5662	16.3733	1.9724	3.5504	**1.4511**
	25	$\hat{\beta}_2$	50.2358	16.9561	2.1970	3.1679	**1.4637**
		$\hat{\beta}_3$	49.0428	16.2526	1.8728	3.4291	**1.5974**
		$\hat{\beta}_1$	36.7919	11.5193	1.7268	3.2345	**1.1874**
ρ=0.99	**50**	$\hat{\beta}_2$	37.2747	11.9873	1.9073	3.2399	**1.2362**

outliers=10%		$\hat{\beta}_3$	34.9526	11.0492	1.7914	3.1326	**1.1414**
		$\hat{\beta}_1$	25.4084	8.1956	2.0064	4.0558	**1.0216**
	100	$\hat{\beta}_2$	24.7682	8.0134	1.8948	4.1744	**1.0639**
		$\hat{\beta}_3$	24.8356	8.0676	1.9224	3.9335	**1.0793**
		$\hat{\beta}_1$	71.7362	15.9207	2.1173	14.2355	**1.1232**
	25	$\hat{\beta}_2$	74.4900	16.9513	2.4359	15.0788	**1.1066**
		$\hat{\beta}_3$	73.5528	16.5358	1.9955	14.8189	**1.1634**
		$\hat{\beta}_1$	49.3527	11.8731	1.5982	9.2223	**1.2091**
ρ=0.99 outliers=20%	**50**	$\hat{\beta}_2$	49.4907	12.0011	1.6354	9.4697	**1.2517**
		$\hat{\beta}_3$	47.4601	11.5370	1.4223	9.3366	**1.1937**
		$\hat{\beta}_1$	33.5315	8.1319	1.5714	7.9758	**1.1120**
	100	$\hat{\beta}_2$	32.3816	7.9985	1.3693	7.2893	**1.1092**
		$\hat{\beta}_3$	33.6884	8.1593	1.4098	7.5249	**1.1335**

Table (6): RMSE values for the parameters $\hat{\beta}_1$, $\hat{\beta}_2$ and $\hat{\beta}_3$ of the Bisquare weighted function led to the generation of Cauchy disturbance distribution, with different sample sizes and 0.0 of ρ and 0% of outliers.

Collinearity and outliers	Sample size	Par.	OLS	RIDGE	RLMS	WRID
		$\hat{\beta}_1$	76.0338	0.8760	**0.8655**	0.8743
	25	$\hat{\beta}_2$	19.5653	0.8665	**0.8565**	0.8571
		$\hat{\beta}_3$	32.1690	0.8780	**0.8557**	0.8660
		$\hat{\beta}_1$	73.5600	0.9028	**0.8936**	0.8982
ρ=0.0 outliers=0%	**50**	$\hat{\beta}_2$	82.0571	0.9064	**0.8783**	0.8882
		$\hat{\beta}_3$	65.0662	0.9015	**0.8839**	0.8908
		$\hat{\beta}_1$	25.0922	0.9179	**0.9096**	0.9098
	100	$\hat{\beta}_2$	30.7459	0.9147	**0.9114**	0.9115
		$\hat{\beta}_3$	21.8587	0.9133	**0.9038**	0.9095

Table (7): RMSE values for the parameters $\hat{\beta}_1$, $\hat{\beta}_2$ and $\hat{\beta}_3$ of the Bisquare weighted function led to the generation of Cauchy disturbance distribution, with different sample sizes and 0.5 of ρ and 5%, 10%, and 20% of outliers.

Collinearity and Outliers	Sample size	Par.	OLS	RIDGE	RLMS	WRID	WRLMS
ρ=0.5 outliers=5%	25	$\hat{\beta}_1$	13.4442	1.0247	0.9809	0.9820	**0.9641**
		$\hat{\beta}_2$	55.7264	1.0064	0.9819	0.9683	**0.9627**
		$\hat{\beta}_3$	65.4795	1.0157	0.9814	0.9745	**0.9650**
	50	$\hat{\beta}_1$	35.1494	1.0028	0.9775	0.9773	**0.9649**
		$\hat{\beta}_2$	20.4408	0.9852	0.9768	0.9649	**0.9646**
		$\hat{\beta}_3$	99.3026	0.9805	0.9752	0.9702	**0.9646**
	100	$\hat{\beta}_1$	54.0001	0.9821	0.9766	0.9724	**0.9693**
		$\hat{\beta}_2$	33.0903	0.9845	0.9827	0.9721	**0.9696**
		$\hat{\beta}_3$	29.2753	0.9792	0.9733	0.9698	**0.9682**
ρ=0.5 outliers=10%	25	$\hat{\beta}_1$	35.5010	1.0404	0.9867	0.9854	**0.9730**
		$\hat{\beta}_2$	55.7800	1.0159	0.9916	0.9788	**0.9718**
		$\hat{\beta}_3$	65.6834	1.0159	0.9884	0.9814	**0.9757**
	50	$\hat{\beta}_1$	35.5010	1.0404	0.9867	0.9825	**0.9752**
		$\hat{\beta}_2$	55.7800	1.0159	0.9916	0.9778	**0.9762**
		$\hat{\beta}_3$	65.6834	1.0159	0.9884	0.9819	**0.9739**
	100	$\hat{\beta}_1$	34.4850	0.9990	0.9827	0.9788	**0.9753**
		$\hat{\beta}_2$	20.8670	0.9931	0.9895	0.9821	**0.9767**
		$\hat{\beta}_3$	99.6775	0.9910	0.9892	0.9764	**0.9760**
ρ=0.5 outliers=20%	25	$\hat{\beta}_1$	13.3415	1.0416	0.9912	1.0322	**0.9821**
		$\hat{\beta}_2$	53.3649	1.0105	0.9968	0.9985	**0.9783**
		$\hat{\beta}_3$	66.2287	1.0413	0.9944	1.0310	**0.9865**
	50	$\hat{\beta}_1$	35.0602	1.0214	0.9883	1.0085	**0.9813**
		$\hat{\beta}_2$	20.2306	0.9932	0.9893	0.9871	**0.9816**
		$\hat{\beta}_3$	10.1989	0.9946	0.9934	0.9885	**0.9797**
	100	$\hat{\beta}_1$	34.3565	0.9891	0.9877	0.9859	**0.9803**
		$\hat{\beta}_2$	21.9355	0.9952	0.9843	0.9888	**0.9814**
		$\hat{\beta}_3$	28.6493	0.9944	0.9862	0.9834	**0.9813**

Table (8): RMSE values for the parameters $\hat{\beta}_1$, $\hat{\beta}_2$ and $\hat{\beta}_3$ of the Bisquare weighted function led to the generation of Cauchy disturbance distribution, with different sample sizes and 0.99 of ρ and 5%, 10%, and 20% of outliers.

Collinearity and Outliers	Sample size	Par.	OLS	RIDGE	RLMS	WRID	WRLMS
ρ=0.99 outliers=5%	25	$\hat{\beta}_1$	539.3470	17.6219	3.2915	7.3765	**2.2822**
		$\hat{\beta}_2$	309.8400	16.9365	3.2369	7.9873	**2.3988**
		$\hat{\beta}_3$	399.8710	17.0689	3.4521	8.3870	**2.3454**
	50	$\hat{\beta}_1$	260.6930	11.9031	2.8967	6.2531	**1.4235**
		$\hat{\beta}_2$	187.8640	11.8749	2.9503	6.3613	**1.4540**
		$\hat{\beta}_3$	161.4990	11.4368	2.8232	5.4830	**1.4779**
	100	$\hat{\beta}_1$	24.0010	7.8427	2.7373	3.6954	**1.1295**
		$\hat{\beta}_2$	204.9350	8.2270	2.5330	3.6380	**1.1398**
		$\hat{\beta}_3$	128.7050	8.0030	2.5646	3.3723	**1.1173**
ρ=0.99 outliers=10%	25	$\hat{\beta}_1$	540.5340	17.8739	2.6284	6.8147	**2.2857**
		$\hat{\beta}_2$	308.7860	16.2799	2.9147	6.7048	**2.2827**
		$\hat{\beta}_3$	400.6470	17.2466	2.9667	7.1761	**2.3026**
	50	$\hat{\beta}_1$	258.1350	11.0295	1.8995	5.8956	**1.3689**
		$\hat{\beta}_2$	183.7840	10.9748	2.0777	5.9038	**1.4312**
		$\hat{\beta}_3$	164.9560	10.7422	2.1277	5.4044	**1.3615**
	100	$\hat{\beta}_1$	247.5590	7.9864	1.9910	3.5608	**1.1237**
		$\hat{\beta}_2$	206.2420	8.0351	1.8062	3.3314	**1.1536**
		$\hat{\beta}_3$	130.5290	8.2137	1.8650	2.9793	**1.1148**
ρ=0.99 outliers=20%	25	$\hat{\beta}_1$	538.2300	16.9860	2.7484	15.5190	**1.4903**
		$\hat{\beta}_2$	297.1080	16.6398	2.5092	15.0558	**1.5353**
		$\hat{\beta}_3$	402.7270	17.2583	2.8482	15.7592	**1.3967**
	50	$\hat{\beta}_1$	259.0720	11.2609	1.8286	9.3057	**1.3471**
		$\hat{\beta}_2$	182.0790	10.8335	1.8424	8.6640	**1.4859**
		$\hat{\beta}_3$	168.5210	11.0509	1.7711	9.0069	**1.4449**
	100	$\hat{\beta}_1$	178.6700	8.2138	1.4967	5.9508	**1.2290**
		$\hat{\beta}_2$	118.6620	7.8324	1.5636	6.1447	**1.2817**
		$\hat{\beta}_3$	131.7760	8.0231	1.7366	6.2444	**1.2483**

Table (9): RMSE values for the parameters $\hat{\beta}_1$, $\hat{\beta}_2$ and $\hat{\beta}_3$ of the Bisquare weighted function led to the generation of Student-t disturbance distribution, with different sample sizes and 0.0 of ρ and 0% of outliers.

Collinearity and Outliers	Sample size	Par.	OLS	RIDGE	RLMS	WRID
ρ=0.0 outliers=0%	25	$\hat{\beta}_1$	**0.3842**	0.6165	0.5924	0.6138
		$\hat{\beta}_2$	**0.3897**	0.6121	0.5841	0.6126
		$\hat{\beta}_3$	**0.3705**	0.6088	0.6001	0.6149
	50	$\hat{\beta}_1$	**0.2623**	0.5951	0.5850	0.5972
		$\hat{\beta}_2$	**0.2569**	0.6026	0.5901	0.5969
		$\hat{\beta}_3$	**0.2436**	0.5982	0.5854	0.5992
	100	$\hat{\beta}_1$	**0.1686**	0.5932	0.5816	0.5902
		$\hat{\beta}_2$	**0.1691**	0.5905	0.5851	0.5889
		$\hat{\beta}_3$	**0.1712**	0.5895	0.5830	0.5938

Table (10): RMSE values for the parameters $\hat{\beta}_1$, $\hat{\beta}_2$ and $\hat{\beta}_3$ of the Bisquare weighted function led to the generation of Student-t disturbance distribution, with different sample sizes and 0.5 of ρ and 5%, 10%, and 20% of outliers.

Collinearity and Outliers	Sample size	Par.	OLS	RIDGE	RLMS	WRID	WRLMS
ρ=0.5 outliers=5%	25	$\hat{\beta}_1$	7.7562	1.0188	0.9775	0.9597	**0.9581**
		$\hat{\beta}_2$	7.5081	1.0173	0.9753	0.9596	**0.9561**
		$\hat{\beta}_3$	7.7566	1.0013	0.9765	0.9596	**0.9591**
	50	$\hat{\beta}_1$	5.0686	0.9957	0.9720	0.9580	**0.9573**
		$\hat{\beta}_2$	5.0127	0.9857	0.9740	0.9575	**0.9560**
		$\hat{\beta}_3$	5.0385	0.9714	0.9677	0.9581	**0.9565**
	100	$\hat{\beta}_1$	3.8059	0.9740	0.9714	0.9626	**0.9610**
		$\hat{\beta}_2$	3.7875	0.9739	0.9709	0.9624	**0.9614**
		$\hat{\beta}_3$	3.8512	0.9726	0.9706	0.9625	**0.9611**
ρ=0.5 outliers=10%	25	$\hat{\beta}_1$	10.5781	1.0227	0.9887	0.9782	**0.9707**
		$\hat{\beta}_2$	10.3531	1.0358	0.9828	0.9731	**0.9694**
		$\hat{\beta}_3$	10.3936	1.0134	0.9862	0.9702	**0.9686**
	50	$\hat{\beta}_1$	7.7285	1.0119	0.9847	0.9754	**0.9715**
		$\hat{\beta}_2$	7.4220	0.9996	0.9854	0.9768	**0.9708**
		$\hat{\beta}_3$	8.0386	0.9867	0.9786	0.9730	**0.9707**

	100	$\hat{\beta}_1$	5.3747	0.9895	0.9806	0.9742	**0.9718**
		$\hat{\beta}_2$	5.2116	0.9783	0.9771	0.9735	**0.9713**
		$\hat{\beta}_3$	5.3658	0.9845	0.9810	0.9759	**0.9714**
ρ=0.5 outliers=20%	25	$\hat{\beta}_1$	15.9433	1.0249	0.9897	1.0162	**0.9830**
		$\hat{\beta}_2$	15.6591	1.0577	0.9918	1.0386	**0.9789**
		$\hat{\beta}_3$	15.4914	1.0123	0.9872	1.0055	**0.9759**
	50	$\hat{\beta}_1$	10.7110	1.0032	0.9893	0.9922	**0.9792**
		$\hat{\beta}_2$	9.9938	1.0175	0.9900	0.9998	**0.9779**
		$\hat{\beta}_3$	10.3199	0.9935	0.9842	0.9896	**0.9794**
	100	$\hat{\beta}_1$	7.0414	0.9909	0.9882	0.9867	**0.9790**
		$\hat{\beta}_2$	7.3687	0.9906	0.9846	0.9824	**0.9773**
		$\hat{\beta}_3$	7.1057	0.9960	0.9857	0.9952	**0.9795**

Table (11): RMSE values for the parameters $\hat{\beta}_1$, $\hat{\beta}_2$ and $\hat{\beta}_3$ of the Bisquare weighted function led to the generation of Student-t disturbance distribution, with different sample sizes and 0.99 of ρ and 5%, 10%, and 20% of outliers.

Collinearity and Outliers	Sample size	Par.	OLS	RIDGE	RLMS	WRID	WRLMS
ρ=0.99 outliers=5%	25	$\hat{\beta}_1$	36.1524	17.0805	3.6399	2.1425	**1.9098**
		$\hat{\beta}_2$	36.1496	16.8572	3.1818	2.0859	**1.7869**
		$\hat{\beta}_3$	37.8469	17.3168	3.8969	2.0963	**1.9191**
	50	$\hat{\beta}_1$	24.0905	11.9968	3.0933	1.8662	**1.3742**
		$\hat{\beta}_2$	24.1853	11.7035	3.4139	1.7744	**1.3193**
		$\hat{\beta}_3$	24.2879	11.7290	3.0674	1.7857	**1.3833**
	100	$\hat{\beta}_1$	18.9644	8.2189	3.1698	1.9303	**1.0921**
		$\hat{\beta}_2$	18.4052	8.1423	2.9061	1.9060	**1.1121**
		$\hat{\beta}_3$	18.2177	8.2767	3.2336	2.0042	**1.0969**
ρ=0.99 outliers=10%	25	$\hat{\beta}_1$	51.1991	16.9592	2.2003	4.8924	**1.9464**
		$\hat{\beta}_2$	50.6350	16.7245	1.9134	4.5262	**1.7856**
		$\hat{\beta}_3$	50.2193	16.8400	2.1204	4.5201	**1.8535**
	50	$\hat{\beta}_1$	37.9301	11.8750	1.9374	3.2397	**1.2629**
		$\hat{\beta}_2$	36.3352	11.6201	2.0478	3.3999	**1.1791**
		$\hat{\beta}_3$	38.6808	12.2641	2.0806	3.4389	**1.3015**
		$\hat{\beta}_1$	25.2224	8.4940	2.0950	3.5808	**1.0824**

	100	$\hat{\beta}_2$	25.1905	8.1902	1.9087	3.5896	**1.0696**
		$\hat{\beta}_3$	26.2743	8.5006	2.1842	3.5527	**1.0936**
		$\hat{\beta}_1$	77.6328	17.3533	2.2137	15.3698	**1.3229**
	25	$\hat{\beta}_2$	74.3532	16.7828	2.1874	14.8895	**1.2371**
		$\hat{\beta}_3$	75.7421	17.2904	2.0472	15.3778	**1.2914**
		$\hat{\beta}_1$	52.8298	12.5645	1.7418	9.4074	**1.2621**
ρ =0.99 outliers=20%	50	$\hat{\beta}_2$	48.6979	11.7374	1.6013	8.7445	**1.2819**
		$\hat{\beta}_3$	50.0828	11.9296	1.6593	9.2481	**1.2784**
		$\hat{\beta}_1$	34.0494	8.4791	1.5152	7.0952	**1.1584**
	100	$\hat{\beta}_2$	35.2928	8.5728	1.3684	7.0510	**1.1075**
		$\hat{\beta}_3$	34.9378	8.4390	1.6890	7.0143	**1.1766**

8. CONCLUSIONS

In regression analysis, two problems are more frequently faced, namely, multicollinearity and outliers. Apparently separate, both problems occur at the same time in actual practice. In order to address them, a numerical example and Monte Carlo simulation were developed for the sake of comparing some combining weighted ridge and robust regression estimators' performance.

The results of the comparisons in question showed that the OLS surpassed all estimators for different sample size, when there was neither multicollinearity nor outliers in the data and the disturbances are normal and S-student. But, when disturbances were Cauchy, the estimator OLS showed less efficiency than the all other estimators.

Be that as it may, the resultant loss is large as far as efficiency is concerned. Data were generated in order to test and permit generalizations to practical circumstances. For that matter, one particular form of the weighted ridge estimator WRID was compared to the WRLMS estimator.

Many other possible weighting forms are likely to be utilized to construct the WRLMS estimator. Hampel (1972) suggested some of them and Askin and Montgomery (1980) discussed their function. However, in this study the Bisquare weighted function was used, giving the result that WRLMS is better than all other estimators where multicollinearity and outliers are present.

REFERENCE

Askin, R. G. & Montgomery, D. C. (1980). "Augmented robust estimators". Technometrics, Taylor & Francis, 22, 333-341.

Alamgir, A. A.; Khan, S. A.; Khan, D. M. & Khalil, U. (2013)." A New Efficient Redescending M-Estimator" Alamgir Redescending M-estimator. Research Journal of Recent Sciences ISSN, 2277, 2502.

Dempster, A.P. Schatzoff, M., and N. Wermuth (1977). "A simulation study of alternatives to Ordinary Least Squares" J.A.S.A., 72, p.99.

Fox, J. and Weisberg, S. (2011). “An R Companion to Applied Regression”. Sage, Thousand Oaks, CA, second edition.

Gibbons, D. G. A. (1981). “Simulation study of some ridge estimators”. Journal of the American Statistical Association, Taylor & Francis Group. 76, 131-139.

Hampel F.R., Andrews D.F., Bickel P.J., Huber P.J., Rogers W.H. and Tukey J.W., (1972). Robust Estimates of Location: Survey and Advances, Princeton University Press, Princeton, New Jersey.

Hoerl, A. E. & Kennard, R. W. (1970a). ” Ridge regression: Biased estimation for non-orthogonal problems. Technometrics, Taylor & Francis Group, 12, 55-67.

Hoerl, A.E. & Kennard, R.W. (1970b). “Ridge regression: Applications to non-orthogonal problems”. Technometrics. 12(1):68-82.

Kutner, M. H. (2004). “Applied Linear Regression Models”. 4th edition. Mc Graw- Hill/I, 1221 Avenue of the Americas, New York, NY 10020.

Montgomery, D.C. & Peck, E.A. (1992). “Introduction to Linear Regression Analysis”. Canada: John Wiley & Sons.

Montgomery, D., & Peck, E. (1982). “Introduction to linear regression analysis”. New York: Wiley.

Penrose, K. W., Nelson, A. and Fisher, A. G. (1985). “Generalized Body Composition Prediction Equation For Men Using Simple Measurement Techniques”. Medicine and Science in Sports and Exercise 17(2), 189.

Rousseeuw, P. J. and Van Zomeren, B. C. (1990). “Unmasking multivariate outliers and leverage point.

Zahari, S. M.; said Zainol, M. & bin Ismail, M. I. a. B. (2012). ” Weighted ridge mm-estimator in robust ridge regression with multicollinearity”. Mathematical Models and Methods in Modern Science.

 kafi.pati@uod.ac