

***Handling multicollinearity and outliers using Weighted Ridge Least Trimmed Squares***

<sup>1\*</sup>KAFI DANO PATI, <sup>2</sup> ROBIAH BINTI ADNAN, <sup>3</sup> SEYED EHSAN SAFFARI  
AND <sup>4</sup> BELLO ABDULKADIR RASHEED

<sup>1,2,4</sup>Department of Mathematical Sciences, Faculty of Science  
UniversitiTeknologi Malaysia,  
81310 UTM Johor Bahru, Johor, Malaysia

<sup>3</sup>Centre of Education  
Sabzevar  
University of Medical Sciences  
Sabzevar, Iran

<sup>1</sup>[kafi\\_dano@yahoo.com](mailto:kafi_dano@yahoo.com), <sup>2</sup>[robiaha@utm.my](mailto:robiaha@utm.my), <sup>3</sup>[ehsanreiki@yahoo.com](mailto:ehsanreiki@yahoo.com),  
<sup>4</sup>[arasheedbello@yahoo.com](mailto:arasheedbello@yahoo.com)

**Abstract** Common problems in multiple linear regression models are multicollinearity and outliers. In this paper, we will propose a robust ridge regression. It is based on weighted ridge least trimmed squares (WRLTS). The proposed method (WRLTS) has been compared to some different estimation methods, namely the Ordinary Least Squares (OLS), Ridge Regression (RR), Robust Ridge Regression (RRR) such as Ridge Least Median Squares (RLMS), Ridge Least Trimmed Squares (RLTS) regression based on LTS estimator and Weighted Ridge (WRID) with respect to Standard Error. Two examples are used to illustrate the proposed method. In both examples, WRLTS is found to be the best estimator among the other methods in this paper.

**Keywords** Multicollinearity; Outliers; Ridge Regression; Robust Estimator; Weighted Ridge Least Trimmed Squares (WRLTS).

## 1.0 INTRODUCTION

One of the main problems in multiple linear regression is multicollinearity. Multicollinearity is the term used to define a case in which the predictor variables are themselves highly correlated with each other and one goal of a regression model is to discover to what extent the outcome (dependent variable) can be predicted by the independent variables. The power of the prediction is indicated by  $R^2$ , also known as the variance explained by the power of determination[6].

The presence of multicollinearity has a number of possible serious effects on the least-squares estimates of the regression coefficients. Some of these effects may be easily confirmed [5]. The ridge regression is one of the techniques that remedied the multicollinearity problem [17]. The technique of Ridge regression (RR) is one of the most popular and best performing instead of the ordinary least squares (LS) estimation in the existence of a multicollinearity problem [9],[10]. Ridge regression was also proposed for selecting the  $k$  ridge parameter added to the diagonal of the identity matrix. The  $k$  chosen should be small enough that the MSE of the ridge estimator is less than the MSE of the OLS estimator[9], [10]. Many other works on ridge regression have been suggested and some of them are [11] and[15] they show that these techniques are immune to the deviation from the normal assumption, that is a heavy-tailed distribution which may arise as a result of outliers[8]. Outlier is the other common problem in regression. It is the extreme observations in the data and may have an unsuitable effect on OLS the estimate of the parameter. Robust estimators reduce the effects of outliers in the data. The problem is more difficult when both multicollinearity and outliers exist. Robust ridge regression analysis has always been the interest of some researchers in the literature.

[8]used the robust ridge regression based on MM estimators (RMM)with the highly efficient and high breakdown point estimator. [3]and[21] suggested to combine the properties of the ridge estimator and the Least Absolute Value (LAV) robust estimator to remedy multicollinearity and outliers simultaneously. Their estimates were required robust estimates, using an appropriately chosen ridge regression. [12] gave the formulas for deviation of ridge regression methods when weights are associated with each observation, and proposed the combination of ridge regression with robust regression methods. In addition, [12]

suggested a procedure for choosing the optimal value of the biasing parameter  $k$  adaptively [13], [22] introduced, the most robust estimator having the highest possible breakdown point, that is 50% which is known as Least Median Squares (LMS) and Least Trimmed Squares (LTS). [2] proposed two alternative ridge type GM estimators to handle simultaneously multicollinearity and the existence of outliers. [20] reduced the effect of outlier by computing a robust estimates for  $k$ , and used these estimates to obtain robust ridge estimates for the regression coefficients. [16] proposed using the Weighted Ridge (WRID). In addition, [29] improved and evaluated new robust regression procedures and compared their performance to the best alternatives currently available, in terms of breakdown, bounded influence and efficiency.

[30] applied a robust regression estimator that achieves well, irrespective of the amount and configuration of outliers. They show that the best available estimators are weak when the outliers are extreme in the regressor space (high leverage). They suggested compound estimator revise recently published methods with an enhanced initial estimate and measure of leverage [13].

In this study, robust ridge regression methods based on RLMS and RLTS estimators are examined in the presence of both multicollinearity and outliers. The performance of the robust ridge estimators examines two examples by using real data to evaluate the standard errors on the data of a hospital manpower [18] and the data of body fat [19]. The propose methods combining the robust ridge regression with the high breakdown point estimator, namely (RLMS) and (RLTS). [4] suggested new biased estimators based on the least trimmed squares (LTS) ridge regression that possesses a high breakdown point, 50% and expected that this method modified and be less sensitive to the presence of multicollinearity and outliers. And we develop of a robust ridge regression estimator based on the weighted ridge least trimmed squares (WRLTS) and is believed to have abilities to remedy the problems of multicollinearity and outliers. We expect that this new statistical technique is not so easily affected by these problems when we compare with several existing estimators, since it removed the influence of multicollinearity and outliers. The efficiency of the this method relative to the alternatives has been examined using SE. In general, it has been found that the WRLTS estimator is the best estimator since it will have a small standard error against the five existing estimators when the error term is not normal.

## 2.0 PROBLEM STATEMENT

### 2.1 Ridge Regression

Consider the following multiple linear regression models

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \quad (1)$$

where  $\mathbf{y}$  is a vector of  $nx1$  response values,  $\mathbf{X}$  is an  $n \times p$  matrix of the regressor variables,  $\boldsymbol{\beta}$  is a  $px1$  vector of unknown parameter, and  $\boldsymbol{\varepsilon}$  is an  $nx1$  vector of random errors, such that  $E(\boldsymbol{\varepsilon}) = 0$ , and  $\text{Var}(\boldsymbol{\varepsilon}) = \sigma^2 \mathbf{I}_n$ . It will be convenient to assume that the regressor variables are standardized. Consequently,  $\mathbf{X}'\mathbf{X}$  is a  $pxp$  matrix of correlations between the regressors and  $\mathbf{X}'\mathbf{y}$  is a  $px1$  vector of correlation between the regressors and the response. When the columns of the design matrix  $\mathbf{X}$  have a near-linear dependence, the ordinary least-squares (OLS) estimate  $\hat{\boldsymbol{\beta}}$  becomes highly sensitive to random errors in the observed response  $\mathbf{y}$  with large variances [13].

If the columns of  $\mathbf{X}$  are multicollinearity, then the least-squares estimator of  $\boldsymbol{\beta}$ , namely

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{y} \quad (2)$$

is an unreliable estimator due to the large variances associated with its elements. The most popular of the methods that can be used to handle multicollinearity is the ridge regression and the estimate can be obtained by using equation (3).

$$\hat{\boldsymbol{\beta}}_R = (\mathbf{X}'\mathbf{X} + k\mathbf{I}_n)^{-1} \mathbf{X}'\mathbf{y} \quad (3)$$

This method, is based on adding a positive constant  $k$  to the diagonal element of  $\mathbf{X}'\mathbf{X}$  [9], [10]. This leads to a biased estimator of  $\boldsymbol{\beta}$ , called the ridge estimator and given by [4] where  $\mathbf{I}$  is the  $(p \times p)$  identity matrix. Note that in Equations (2) and (3), are different from the OLS estimator and the ridge regression estimator and  $k$  is the ridge parameter,  $k > 0$ . [11] suggested that the chosen ridge parameter  $k$ , should be small positive value and in order MSE of the Ridge regression estimator is less than the MSE of the OLS estimator and added to the diagonal elements of the  $\mathbf{X}'\mathbf{X}$  matrix. They proposed another criterion for choosing the  $k$  ridge parameter as given in Equation (4).

$$k_{LS} = \frac{pS_{LS}^2}{\hat{\boldsymbol{\beta}}'_{LS} \hat{\boldsymbol{\beta}}_{LS}} \quad (4)$$

where  $p$  is the number of independent variables, and

$$S_{LS}^2 = \frac{(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}_{LS})'(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}_{LS})}{n-p} \quad (5)$$

when  $k=0$ ,  $\hat{\boldsymbol{\beta}}_R = \hat{\boldsymbol{\beta}}_{LS}$ , when  $k>0$ ,  $\hat{\boldsymbol{\beta}}_R$  is biased, but precise and more stable than the LS estimator and when  $K \rightarrow \infty$ ,  $\hat{\boldsymbol{\beta}}_R \rightarrow 0$ . [9] have presented that there always exists a value  $k>0$  such that the MSE of the ridge estimator,  $MSE(\hat{\boldsymbol{\beta}}_R)$  is less than the MSE of the OLS estimator,  $MSE(\hat{\boldsymbol{\beta}}_{LS})$ .

## 2.2 Robust Regression Methods

[20] Illustrates that robust regression analysis provides an alternative to a least squares regression when essential assumptions are unfulfilled by the nature of the data [26]. The properties of efficiency, breakdown and high leverage points are used to define robust techniques. One objective of robust estimators is a high breakdown point  $\varepsilon$  defined by [6]. The breakdown point is simply the initial point due to contaminated data and it can be defined as the point or limiting the percentage of contamination in the data at which any test statistics first becomes swamped. Hence, some regression estimators have the smallest possible breakdown point of  $1/n$  or  $0/n$ . In other words, only one outliers would cause the regression equation to be rendered useless. Other regression estimators have the highest possible breakdown point of  $n/2$  or 50%. If robust estimation technique has a 50% breakdown point, then 50% of the data could contain outliers and the coefficients would remain usable [22],[1] state takes any sample of  $n$  data points, then robust regression estimators have been proven to be more reliable and efficient than least squares estimator especially when disturbances are nonnormal. Nonnormal disturbances are disturbance distributions that have heavy or fatter tails than the normal distribution and are prone to produce outliers. Since outliers greatly influence the estimated coefficients. There are several different classifications of robust regression exist, to use to reduce the outliers such as least median squares (LMS) and (LTS).

## 2.2.1 Least Median Squares LMS Estimator

It can be defined as the solution to the following minimize  $\text{med } r_i^2$  rather than minimizing the sum of squared residuals as in least squares estimation, the minimize med of the residuals is minimized. Thus, the effect of outliers on the LMS estimates will be less than that on LS estimates.

Least median of squares regression procedure has a high breakdown point this reason makes to interest in it and has strength. Approximately, the breakdown point of a statistical estimator is the smallest percentage of contamination which may cause the estimator to take on arbitrary large values [7].

It now appears that breakdown point has risen as one of the basic criteria for judging the robustness of an estimator [20].

LMS has a breakdown point of 50% and this is obviously the highest possible breakdown point of any reasonable estimator [22] and [23].

## 2.2.2 Least Trimmed Squares LTS Estimator

When problems with parameter estimation occur and assumptions do not hold under the linear model, the least trimmed squares is the alternative to OLS regression and is very common in statistical modeling [23]. [25] explained that the objective function of the LTS method is the smallest trimmed of squared residuals as follows.

$\varepsilon_{(1)}^2 \leq \varepsilon_{(2)}^2 \leq \varepsilon_{(3)}^2 \leq \dots \leq \varepsilon_{(n)}^2$  denotes the order statistics of a set of residuals, from smallest to largest. LTS are calculated by minimizing the  $h$  ordered squares residuals, where  $h = [n/2] + [(p+1)/2]$ , with  $n$  and  $p$  being sample size and number of parameters, respectively. The largest squared residuals are excluded from the summation in this method, which allows those outlier data points to be excluded completely. Depending on the value of  $h$  and the outlier data configuration, LTS can be very efficient. In fact, if the exact numbers of outlying data points are trimmed, this method is computationally equivalent to OLS. However, if there are more outlying data points are trimmed, this method is not as efficient.

On the contrary, if there is more trimming than there are outlying data points, then some good data will be excluded from the computation. In terms of

breakdown, LTS reaches the maximal possible value of the breakdown point of 50% [22] and [24].

Then the estimate of  $\hat{\beta}_{LTS}$  is computed by solving the  $\binom{n}{h}$  total least squares results for all subsets of size  $h$ . Hence, the solution that minimizes  $\sum_{i=1}^h \varepsilon_i^2$  exists and is obtained computationally [24].

### 2.2.3 Weighted Ridge Estimators

[9] and [10] developed a ridge regression method in which a constant  $k$  is added to the  $\mathbf{X}'\mathbf{X}$  matrix in order to estimate the coefficient of parameter that is computed by using the following form

$\hat{\beta}_R = (\mathbf{X}'\mathbf{X} + k\mathbf{I}_n)^{-1} \mathbf{X}'\mathbf{y}$  instead of  $\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{y}$ . The constant  $k$  which is based on a trace would be objectively determined based on the judgment of the researcher [14].

The robust ridge regression estimator proposed by [3] first time introduced a weighted least squares estimator, and the procedure of the Weighted Least Squares estimator can be used to compute the WRLTS estimates. Then the Weighted Least Squares estimator can be written as:

$$\hat{\beta}_{WLS} = (\mathbf{X}'\mathbf{W}\mathbf{X})^{-1} \mathbf{X}'\mathbf{W}\mathbf{Y} \quad (6)$$

Then WRID estimator is computed using the formula where  $W$  defines a new diagonal matrix with diagonal elements  $w_{ii}$ . Now, we will define a new diagonal matrix with square roots of  $w_{ii}$  along the main diagonal.

This matrix is symmetrical where  $\mathbf{W}^{1/2}\mathbf{W}^{1/2} = \mathbf{W}$ , and  $\mathbf{W}^{-1/2}\mathbf{W}^{-1/2} = \mathbf{W}$ .

Now we can define the matrix of regression coefficients for this weighted model.

$$\begin{aligned} b_w &= [(\mathbf{W}^{1/2}\mathbf{X})' \mathbf{W}^{1/2}\mathbf{X}]^{-1} (\mathbf{W}^{1/2}\mathbf{X})' \mathbf{W}^{1/2}\mathbf{Y} \\ &= (\mathbf{X}'\mathbf{W}^{1/2}\mathbf{W}^{1/2}\mathbf{X})^{-1} \mathbf{X}'\mathbf{W}^{1/2}\mathbf{W}^{1/2}\mathbf{Y} \\ &= (\mathbf{X}'\mathbf{W}\mathbf{X})^{-1} \mathbf{X}'\mathbf{W}\mathbf{Y} \end{aligned} \quad (7)$$

The diagonal elements of  $W$  matrix are set equal to

$$w_{ii} = \begin{cases} \frac{1}{|\hat{\varepsilon}_i|} & \text{if } \hat{\varepsilon}_i \leq \text{zero} \\ 1 & \text{if } \hat{\varepsilon}_i > \text{zero} \end{cases} \quad (8)$$

Where the  $\hat{\varepsilon}_i$  are residuals from an initial LS fit to the data. The weights  $w_{ii}$  are applied to the observations and are intended to downweight the extreme observations. Thus, [6] used weighted least squares estimates and computed by applying least squares to the transformed observations  $\sqrt{w_{ii}}y_i$  and  $\sqrt{w_{ii}}x_i$ .

This produces an estimate equal to  $\hat{\beta}_{WLS}$  in equation (8). In this case, the estimation procedure can be iterated to produce what we called the iteratively reweighted least squares estimates [6].

### 2.2.4 Ridge Least Trimmed Squares LTS Estimator

[4] proposed new estimator based on the (LTS) ridge estimator. They let  $\hat{\beta}_{LTS}$  be the estimate of the LTS parameter. They consider a robust estimate of the LTS estimator,  $\hat{\beta}_{LTS-RIDGE}$ , based on a ridge estimator  $\hat{\beta}$ .

$$\hat{\beta}_{LTS-RIDGE} = (\mathbf{X}'\mathbf{X} + k_{LTS}\mathbf{I})^{-1}\mathbf{X}'\mathbf{X}\hat{\beta}_{LTS} \quad (9)$$

where  $\mathbf{X}'\mathbf{X}$  has the form of a correlation matrix and  $k_{LTS}$  is the robust choice of the  $k$  parameter and can be calculated as

$$k_{LTS} = \frac{pS_{LTS}^2}{\hat{\beta}'_{LTS}\hat{\beta}_{LTS}} \quad (10)$$

Here,  $p$  is the number of independent variables and

$$S_{LTS}^2 = \frac{(\mathbf{y} - \mathbf{X}\hat{\beta}_{LTS})'(\mathbf{y} - \mathbf{X}\hat{\beta}_{LTS})}{n-p} \quad (11)$$

where  $S^2$  is the estimated variance,  $n$  is the sample size and  $p$  is the number of estimated parameters.

Robust Ridge Regression based on Weighted Ridge LTS estimators (WRLTS). These are practical tools for parameter estimations in the presence of multicollinearity and outliers [9].

To estimate the standard errors (SE) of the coefficient for the WRLTS written as  $S^2(\mathbf{A}'\mathbf{A} + \mathbf{\Gamma}'\mathbf{\Gamma})^{-1}\mathbf{A}'\mathbf{A}(\mathbf{A}'\mathbf{A} + \mathbf{\Gamma}'\mathbf{\Gamma})^{-1}$  (12)

The notation above follows the wiki notation for ridge regression. Specifically,  $\mathbf{A}$  is the covariate matrix,

$S^2$  is the error variance.

$\Gamma$  is the Tikhonov matrix chosen suitably in ridge regression

### 2.2.5 Weighted Ridge Least Trimmed Squares WRLTS Estimator

To compute weighted robust ridge estimators, the formula used is

$$\hat{\beta}_{WRID} = (\mathbf{X}'\mathbf{W}\mathbf{X} + k\mathbf{I})^{-1} \mathbf{X}'\mathbf{W}\mathbf{Y} \quad (13)$$

The value of  $k$  (the biasing parameter), could be obtained from the untransformed observations through the transformed observations could also be used. The estimator

$\hat{\beta}_{WRID}$  With the biasing parameter  $k$  and weights  $w_{ii}$  determined from the data is denoted to as the weighted ridge (WRID) estimator. An estimator of  $k$ , has been suggested by [11].

The combination of weighted robust LTS estimator with ridge regression can be written in the following manner.

$$\hat{\beta}_{WRLTS} = (\mathbf{X}'\mathbf{W}\mathbf{X} + k\mathbf{I})^{-1} \mathbf{X}'\mathbf{W}\mathbf{Y} \quad (14)$$

where  $k$  obtained from the LTS estimator while  $W$  is robust weighted. In order to evaluate the robustness of this estimator, we compared the results of the numerical index of (SE) of parameter estimates using real data error distribution as well as multicollinearity and outliers.

### 3.0 REAL DATA AND RESULTS

Following [14], [8], [28]. The proposed estimation technique WRLTS is compared to several existing estimators, namely: Ordinary least squares (OLS), Ridge regression (RIDGE), Ridge LMS estimator (RLMS), Ridge LTS estimator (RLTS) and Weighted ridge (WRIDGE), using two real data sets. Table 1 Contains a portion of the data for a study of the relation of the amount of body fat ( $y$ ) to several possible predictor variables, based on a sample of 20 healthy females 25-34 years old. The possible predictor variables are triceps skinfold thickness ( $x_1$ ), thigh circumference ( $x_2$ ), and midarm circumference ( $x_3$ ). The amount of body fat on the table for each of the 20 persons was obtained by a cumbersome and expensive procedure requiring the immersion of the person in the water, and Table 2 consists of the measures taken at 17 U.S. Naval Hospitals

and the goal is to predict the required monthly man hours for staffing purposes with 17 observations on the following 6 variables. Hours monthly man hours (response variable), X1 Load average daily patient load, X2, X-ray monthly X-ray exposures, X3 Bed Days monthly occupied bed days, X4 Area Popeligible population in the area in thousands, X5 Stay average length of a patient's stay in days. SE used for this purpose to evaluating the results when the error term is nonnormal.

**Table 1: -Value of Standard Errors of the estimators for the real data (body fat)**

| Parameter | OLS    | RIDGE    | RLMS     | RLTS     | WRIDGE   | WRLTS    |
|-----------|--------|----------|----------|----------|----------|----------|
| $\beta_1$ | 3.016  | 0.035369 | 0.060926 | 0.050689 | 0.019977 | 0.015555 |
| $\beta_2$ | 2.582  | 0.039837 | 0.068622 | 0.057092 | 0.020084 | 0.014997 |
| $\beta_3$ | 1.5954 | 0.051537 | 0.088777 | 0.073861 | 0.023509 | 0.017153 |

**Table 2: -Value of Standard Errors of the estimators for the real data (manpower)**

| Parameter | OLS      | RIDGE    | RLMS     | RLTS     | WRIDGE   | WRLTS    |
|-----------|----------|----------|----------|----------|----------|----------|
| $\beta_1$ | 9.77E+01 | 0.005995 | 0.035707 | 0.0182   | 0.005995 | 0.004499 |
| $\beta_2$ | 2.13E-02 | 0.009223 | 0.054937 | 0.028001 | 0.009223 | 0.004953 |
| $\beta_3$ | 3.09E+00 | 0.006068 | 0.036141 | 0.018421 | 0.006068 | 0.004608 |
| $\beta_4$ | 7.18E+00 | 0.008479 | 0.050501 | 0.02574  | 0.008479 | 0.006251 |
| $\beta_5$ | 2.10E+02 | 0.012707 | 0.075685 | 0.038577 | 0.012707 | 0.00706  |

From Table 1 we can see that the SE for the LS is relatively greater than the other estimators when the errors are normally distributed in the presence of multicollinearity and outliers, likewise, the results in Table 2 also show that the SE for WRLTS is least compared to all existing methods indicating that the WRLTS is more efficient than the other methods. Evaluation of the efficiency of the estimators, measured in terms of the SE ratios and is exhibited in Table 1, Table 2 SE of an estimator under evaluation WRLTS by the SE indicates more efficiently than a benchmark estimator. For the comparison between WRLTS and WRIDGE estimators, results showed that WRLTS and WRIDGE outperform better than the RLTS in the presence of multicollinearity and outliers, On the contrary RIDGE, WRIDGE seemed to be more efficient as compared for all cases.

#### **4.0 SUMMARY AND CONCLUSION**

In summary, the OLS estimator outperforms the other estimators when there is no multicollinearity and outliers. But when multicollinearity and outliers exist in the dataset the WRLTS and WRID estimators outperformed the other estimators. Conversely, when the error term is non normal, the WRLTS estimator was found to be less efficient as compared to the RIDGE, WRIDGE, RLMS, RLTS estimators.

Multicollinearity datasets with outliers are very common in practice. In order to solve both problems, robust ridge regression estimators biased estimation methods are applied. It is concluded that the best model is obtained for the WRLTS model with the minimum SE value.

Real data results above show that the SE values magnitudes are affected by the type of outlier.

The SE value for the WRLTS is obtained as the minimum value compared with the other methods. This result is expected since there are both multicollinearity and outliers in the data set. The LTS is better than the OLS in terms of their SE values.

Consequently, in this study, it is shown that for the dataset with both multicollinearity and outliers, robust and biased estimation methods give better results than those that depend on the OLS.

## ACKNOWLEDGMENTS

We would like to acknowledge the financial support from universititeknologimalaysia for the research university Grant (Q.J130000.2626.07J71) and Ministry of Higher Education (MOHE) of Malaysia.

## REFERENCES

- [1] A. Christmann, Least median of weighted squares in logistic regression with large strata. *Biometrika* 81, 413-417, 1994.
- [2] Arslan, O., and Billor, N. Robust ridge regression estimation based on the GM-estimators. *Journal of Mathematical and Computational Science*, 9 (1), 1-9, 1996.
- [3] Asking, G. R., and Montgomery, D. C. Augmented robust estimators *Technometrics*, 22, 333-341, 1980.
- [4] B. Kan, O. Alpu and B. Yazıcıa, Robust ridge and robust Liu estimator for regression based on the LTS estimator, *Journal of Applied Statistics*, vol. 40, No. 3, 644–655, 2013.
- [5] Draper, N. R., Smith, H. *Applied regression analysis*, 3rd edition, Wiley, New York, 2003.
- [6] D.C. Montgomery and E.A. Peck. *Introduction to Linear Regression Analysis*. Wiley, New York, 1982.
- [7] F. R. Hampel. A general qualitative definition of robustness, *Ann. Math. Statist.* 42, 1887-1896, 1971.
- [8] Habshah Midi and Marina Zahari. A Simulation Study on Ridge Regression Estimators in the Presence of Outliers and Multicollinearity. *Jurnal Teknologi*, 47 (C): 59-74, 2007.
- [9] Hoerl, A. E., and Kennard, R. W. Ridge regression: Biased estimation for nonorthogonal problems *Technometrics*, 12, 55- 67, 1970a.
- [10] Hoerl, A. E., and Kennard, R. W. Ridge regression: Applications to nonorthogonal problems *Technometrics*, 12, 69-82, 1970b.
- [11] Hoerl, A. E., Kennard, R. W., and Baldwin, K. F. Ridge regression: Some simulations. *Communications in Statistical Theory*, 4 (2), 105-123, 1975.
- [12] Holland, P. W. Weighted ridge regression: Combining ridge and robust regression methods, *NBER Working Paper Series*, #11, 1-19, 1973.
- [13] H. Samkar and O. Alpu, Ridge Regression Based on Some Robust Estimators,

- Journal of Modern Applied Statistical Methods, volume 9, Issue 2, 2010.
- [14] K. Lawrence and J. Arthur (Eds.), In *Robust regression: Analysis and applications*, New York: Marcel Dekker, 243-270, 1990.
- [15] Marquard, D. W. and R. D. Snee. Ridge Regression in Practice. *The American Statistician*. 29:3-20. 1975.
- [16] Montgomery, D.C. and Peck, E.A. *Introduction to Linear Regression Analysis*, Canada: Canada: John Wiley & Sons, 1992.
- [17] M. El-Fallahand A. El-Salam, The Efficiency of Some Robust Ridge Regression For Handling Multicollinearity and Non-Normals Errors, *Problems, Applied Mathematical Science*, vol. 7, no. 77, 3831 – 3846, 2013.
- [18] Myers, R. *Classical and Modern Regression with Applications*, Boston, MA: Duxbury, 1990.
- [19] N.kutner, *Applied statistical models*, page 408-409, 2004.
- [20] O .G. Alma, Comparison of Robust Regression Methods in Linear Regression, *Int.J. Contemp. Math. Sciences*, vol. 6, no. 9, 409 – 421, 2011.
- [21] Pfaffenberger, R. C. and T. E. Dielman. A Modified Ridge Regression Estimator Using the Least Absolute Value Critrion in the Multiple Linear Regression Model of the American Institute for Decision Sciences. Toronto, 791-793, 1984.
- [22] P.J. Rousseeuw and A.M. Leroy, *Robust Regression and Outlier Detection*, Wiley, New York, 1987.
- [23] P.J. Rousseeuw, Least median of squares regression, *J. Amer. Statist. Assoc.* 79 -871-880, 1984.
- [24] P.J. Rousseeuw and K. Van Driessen, Computing LTS regression for large datasets, Tech. Rep., Statistics Group, University of Antwerp, Antwerp, Belgium, 1999.
- [25] P.J. Rousseeuw, Multivariate estimation with high breakdown point, *The Fourth Pannonian Symposium on Mathematical Statistics and Probability*, Bad Tatzmannsdorf, Austria, 1983.
- [26] R. Marona, R. Martin, V.J. Yohai, *Robust Statistics Theory and Methods*, John Wiley & Sons Ltd., England, 2006.
- [27] Rousseeuw, P. J., and Yohai, V. J. Robust regression by means of S-estimators, in *Robust and nonlinear time series analysis*, 1984.
- [28] S. M. Zahari, M. S. Zainol, and M. I. AL .bin Ismail, Weighted Ridge MM-Estimator in Robust Ridge Regression with Multicollinearity, *Mathematical Models and Methods in Modern Science*, 2012.

- [29] Simpson, J. R., & Montgomery, D. C. The development and evaluation of alternative generalized M estimation techniques. *Communications in Statistical Simulations*, 27 (4), 999-1018, 1998b.
- [30] Wisnowski, J. W., Simpson, J. R., and Montgomery, D. C. An improved compound estimator for robust regression. *Communications in Statistical Simulations*, 31(4), 653-672, 2002.