Using Ridge Least Median Squares to Estimate the Parameter by Solving Multicollinearity and Outliers Problems

Kafi Dano Pati, Robiah Adnan, Bello Abdulkadir Rasheed

Department of Mathematical Sciences, Universiti Teknologi Malaysia, 81310 Skudai, Johor, Malaysia
kafi_dano@yahoo.com

Abstract: In the multiple linear regression analysis, the ridge regression estimator is often used to address multicollinearity. Besides multicollinearity, outliers are also a problem in the multiple linear regression analysis. We propose new biased estimators robust ridge regression called the ridge least median squares (RLMS) estimator in the presence of both outliers and multicollinearity. For this purpose, a simulation study is conducted in order to see the difference between the proposed method and the existing methods in terms of their effectiveness; the mean square error. In our simulations, the performance of the proposed method RLMS is examined for different number of observations, and the different percentage of outliers. The results of different illustrative cases are presented. This paper also provides the results of the RLMS on a real-life data set. The results show that RLMS is better than the existing methods. (OLS, RLAV and Ridge Regression) in the presence of multicollinearity and outliers.
[Pati K.D., Adnan R., Rasheed B. A. **Using Ridge Least Median Squares to Estimate the Parameter by Solving Multicollinerity and Outliers Problems.** *Nat Sci* 2014;12(11):122-128]. (ISSN: 1545-0740). http://www.sciencepub.net/nature. 19

Keywords: Multicollinearity, ridge regression, outliers, robust estimation, robust ridge methods

1. Introduction

It is well known that the classical least squares (LS) estimation has many desirable statistical properties, particularly when errors are normally distributed. The variance of LS estimations are the smallest among all sorts of unbiased estimator class. Nevertheless, LS estimators frequently are influenced by two types of problem, outliers and multicollinearity which in turn lead to highly unstable LS estimates.

To overcome the influence of multicollinearity, many biased estimators were introduced to handle the mlticollinearity problem in regression model. (Hoerl and Kennard, 1970 a, b), first proposed a useful tool called ridge regression for improving prediction in regression situations with highly correlated regressors variables. This method has been successfully applied to many fields. However, the ridge regression method is not robust, it is sensitive to observations that is considered as contaminated data usually called outliers.

The robust ridge regression estimation has been investigated, to remedy the effect of multicollinearity and outliers. The parameter estimation is more difficult when both problems are present in the data set. Different robust estimation methods have been used in the literature to estimate the parameter of regression model in the presence of outliers and multicollinearity. (Pariente, 1977; Askin and Montgomery, 1980; Pfaffenberger and Dielman, 1984; Pfaffenberger, 1985 and Nikolai Krivulin, 1992), among others presented several new ideas concerning the LMS estimator so as to provide some theoretical framework for efficient regression algorithms. (Desire and Leonard, 1986) introduced a procedure to determine the minimum redundancy of a measurement set and the non-redundant samples of measurement that is robust against outliers and multicollinearity.

This paper propose a new estimation method of a robust ridge estimator that has a high breakdown point called the robust ridge regression estimation method based on least median squares (RLMS). The least median squares estimator has 50% breakdown point and was proposed by (Rousseeuw, 1984, 1985), who replaced the least sum of squares (LS) that has 0% breakdown point with the least median of squares LMS. (Hampel, 1975), explains the breakdown point as the smallest percentage of contaminated data that can cause the estimator to take a random large abnormal value.

The aim of evaluating these alternatives biased robust methods is to find the estimator of the parameters of the model that is highly efficient and effective in the presence of outliers and multicollinearity. The performance of the robust ridge estimators is examined by using the standard errors (SE) on a hald data set. The remainder of the paper is organized as follows. After a brief review of ridge regression in Section 2, we define a robust regression and ridge robust estimators in Section 3. Numerical example is given in section 4, and simulation is presented in Section 5, in order to compare between the proposed and the existing methods. It will illustrate that the new estimation method is better than the existing methods when multicollinearity and outliers occur simultaneously. Finally, in section 6 conclusions are presented.

2. Material and Methods

2.1. Ridge Regression

When the least-squares method is applied to non-orthogonal data, we obtained weak estimates of the regression coefficients (Hoerl and Kennard, 1976).

The assumptions of the least squares method that $\hat{\boldsymbol{\beta}}$ is an unbiased estimator of $\boldsymbol{\beta}$.

The Gauss-Markov property assures that the least-squares estimator has a minimum variance in the class of unbiased linear estimators in the presence of multicollinearity (Marquardt and Snee, 1975).

One way to reduce this problem is to descend the requirement that the estimator of $\boldsymbol{\beta}$ be unbiased. Assume that a biased estimator of $\boldsymbol{\beta}$ is found to say $\hat{\boldsymbol{\beta}}$ that has smaller variance than the unbiased estimator $\hat{\boldsymbol{\beta}}$. The mean square error of $\boldsymbol{\beta}$ is defined as

$$\mathbf{MSE}(\hat{\boldsymbol{\beta}}) = \mathrm{VAR}(\boldsymbol{\beta}) + \mathrm{E}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})^2 \qquad (1)$$

Assuming a small amount of bias in $\hat{\boldsymbol{\beta}}$, the variance of $\hat{\boldsymbol{\beta}}$ can be made small such that the MSE of $\hat{\boldsymbol{\beta}}$ is less than the variance of the unbiased estimator $\boldsymbol{\beta}$.

A number of procedures have been improved in obtaining biased estimates of regression coefficients. One of these procedures is ridge regression, primarily proposed by (Hoerl and Kennard, 1970). More specifically, this estimator is defined as the solution to $\hat{\boldsymbol{\beta}} = (\mathbf{X'X} + K\mathbf{I})^{-1}\mathbf{X'y}$ where $K \geq 0$ is a constant selected by the analyst. The constant K is usually refers to the biasing parameter, it is obvious that if $K = 0$ then the ridge estimator is the least squares estimator. Hence, on the other hand, when $K > 0$ the bias $\hat{\boldsymbol{\beta}}_{\mathrm{ridge}}$ increases. However, the variance decreases as K increases.

In this ridge regression, we would like to select a value of K such that the decrease in the variance term is greater than the increase in the squared bias. If this can be done, the mean square error of the ridge estimator $\hat{\boldsymbol{\beta}}_{\mathrm{ridge}}$ will be less than the variance of the least square estimator $\hat{\boldsymbol{\beta}}$. (Hoerl and Kennard, 1976), proved that there exists a non-zero value of K for which the mean squares error (MSE) is less than the variance of the least squares estimator $\hat{\boldsymbol{\beta}}$. We can select a reasonably small value of K at which the ridge estimates $\hat{\boldsymbol{\beta}}_{\mathrm{ridge}}$ are steady. Mostly, this will produce a set of estimates with smaller MSE than the least-squares estimates. (Hoerl and Kennard, 1976), have suggested that an appropriate choice of K is where $\hat{\boldsymbol{\beta}}$ and $\hat{\sigma}^2$ are found by the least squares solution.

2.2. Robust Regression Estimators

Robust regression analysis available is an alternative to a least squares regression model when essential assumptions are unaccomplished by the nature of the data. When the researchers estimate a regression model and test their assumptions, it is always observed that the assumptions are dramatically violated. Sometimes, the analyst can transform the variables to conform to those assumptions. But often, a transformation will not reduce the leverage of effective outliers that bias of the prediction. Under these situations, robust regression that is resistant to the influence of outliers may be the only reasonable recourse.

There are several different classifications of robust estimates that exist to handle these violations. One important estimator is called the least median squares estimator (Rousseeuw and Leroy, 1987), which has the advantage of minimizing the influence of the residuals. According to (Venables and Ripley, 1999), this algorithm minimizes the median of ordered squares of residuals in order to get the regression coefficients $\boldsymbol{\beta}$ and can be written as equation 2.

$$\text{Least Median of Squares} = \text{Min median } \left|\mathbf{y} - \mathbf{x_i}\boldsymbol{\beta}_{\mathrm{LMS}}\right|^2 \qquad (2)$$

(Martin, 2002) describes the median squared residuals lack a smooth squared residual function and the asymptotic rate is $n^{\frac{1}{3}}$ to convergence efficiently under normality. Also, it takes a long time to converge. Least median of squares (LMS) estimator is one of the true high breakdown point estimators that reached the above mentioned upper boundary of the breakdown point. (Rousseeuw, 1984).

2.2.1. Robust Ridge Regression Estimators (RRR)

(Pfaffenberger and Dielman 1985, 1990), proposed robust ridge regression by extending the development of their technique by performing Monte Carlo simulation studies to compare various approaches. The proposed method in this paper combines the LMS estimator with the ridge regression estimator which is referred to as the RLMS estimator. So, RLMS robust ridge estimators will be resistant to multicollinearity problems and less affected by outliers. The RLMS estimator can be written as equation 3.

$$\hat{\boldsymbol{\beta}}_{\mathrm{RLMS}} = \left(\mathbf{X'X} + K_{\mathrm{LMS}}\mathbf{I}\right)^{-1}\mathbf{X'Y} \qquad (3)$$

The value of K is determined from data using equation 4.

$$K_{LMS} = \frac{pS_{LMS}^{2}}{\hat{\beta}_{LMS}^{\prime}\hat{\beta}_{LMS}} \quad \text{and} \quad S_{LMS}^{2} = \frac{\sum_{i=1}^{n} \varepsilon_{i\,LMS}^{2}}{n-p} \qquad (4)$$

where p is the number of independent variables, n is the number of observations in the data, $\hat{\beta}_{LMS}$ is the estimates of β and ε_{LMS}^{2} is the residuals from LMS method.

3. Results

3.1. Numerical Result

A real data set from (Hald, 1952), is considered to assess the effectiveness of the proposed robust ridge regression method. This data set consists of 4 variables and 13 observations with 3 outliers. The response variable (y) is the heat evolved for a particular mixture of cement, and the covariates are tricalcium aluminate (x1), tricalcium silicate (x2), tetracalcium alumina ferrite (x3), and dicalcium silicate (x4).

Table 1 contains the parameter estimates, the standard error of the parameters estimates and VIF analysis of the Hald Data. The VIF for each of these predictors are all extremely high, indicating troublesome multicollinearity with the presence of outliers in the data, the use of robust method provides more stable parameter estimates.

With this aim, initial robust regression estimates was first calculated to obtain robust ridge estimates in the presence of both multicollinearity and outliers; these estimates are given in Table 1.

The standard error of the proposed RLMS method is lower than the existing RLAV in the presence of outliers. Table 1 shows that this data have a high VIF for all variables with three outliers.

Also, the standard errors of the parameter estimates for RLMS are less than all. But the mentioned existing methods except for the ridge regression the difference is small.

Table 1. Estimated parameter and SE of $\hat{\beta}_1$, $\hat{\beta}_2$, $\hat{\beta}_3$ and $\hat{\beta}_4$ for the different methods with the variance inflation factor.

Coef.	Estimate	VIF	LS	RIDGE	RLMS	RLAV
$\hat{\beta}_1$	**parameter**	38.4962	1.5511	0.5081	0.3964	0.6466
	s.e.		0.7448	0.0791	0.1036	0.1325
$\hat{\beta}_2$	**parameter**	254.423	0.5102	0.3102	0.3545	0.6082
	s.e.		0.7238	0.0916	0.1070	0.2726
$\hat{\beta}_3$	**parameter**	46.8684	0.1019	-0.0605	-0.0480	0.1253
	s.e.		0.7547	0.0777	0.1012	0.1392
$\hat{\beta}_4$	**parameter**	282.513	-0.1441	-0.3879	-0.3548	-0.0859
	s.e.		0.7091	0.0925	0.1064	0.2856

3.2. Simulation Study

We carry out a simulation study to compare the performance of the different methods LS, RR and RLAV with the proposed estimator RLMS. The simulation is designed to allow both multicollinearity and non-normality to exist simultaneously. The non-normal distributions are used to generate outliers.

Suppose, we have the following linear regression model 20. (Siti Meriam et al., 2012)

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + e_i$$

where i=1, 2, 3 (5)

The parameter values β_0 , β_1 , β_2 and β_3 are set equal to one. The explanatory variables x_{i1} , x_{i2} and x_{i3} are generated using equation (6)

$$x_{ij} = (1-\rho^{2}) + \rho z_{ij} \qquad i=1, 2,\ldots, n,\ j=1, 2, 3 \quad (6)$$

where, z_{ij} are independent standard normal random numbers generated by the normal distribution.

The explanatory variable values were generated for a given sample size n. The sample sizes used were 50 and 100. The value of ρ representing the correlation between any two explanatory variables, and its values were chosen as: 0.0, 0.5 and 0.99. The percentage of outliers present in this data set is 20%.

The number of replications used is 500. The statistics computed are the bias, root of mean squared error (RMSE), standard error (SE), and 6 pairwise MSE ratios of the estimators.

The bias and MSE are given as: Bias $=\bar{\beta}_i - \beta_i$

where $\bar{\beta}_i = \frac{\sum_{i=1}^{k} \beta_i}{k}$ $k = 500$, and the mean squared error (MSE) is $MSE = \frac{1}{500}\sum_{i=1}^{500}(\hat{\beta}_i - \beta_i)^2$, therefore, the RMSE is given by $[MSE(\hat{\beta}_j)]^{1/2}$

where j= 0, 1, 2, 3

The VIF for the simulated data are shown in Table 2.

Table 2. The VIF for the simulated data

$\rho = 0.99$			
Var	**X1**	**X2**	**X3**
VIF (N=50)	126.1541	204.3971	112.0508
VIF (N=100)	133.8203	238.6387	141.2589

Here the maximum VIF is 238.6387 when the correlation between independent variables was very high with different sample size. So it is clear that the multicollinearity problem exists.

Table 3 Bias, RMSE and SE of $\hat{\beta}_1$, $\hat{\beta}_2$ and $\hat{\beta}_3$ with error normal (0,1) distribution of the sample size n=50 correlation 0.0, 0.5, 0.99 outliers 0% and 20%

0% outliers						
Value of ρ	**Coef.**	**Parameter**	**LS**	**RIDGE**	**RLAV**	**RLMS**
	$\hat{\beta}_1$	**Bias**	-0.0015	-0.4994	-0.5016	-0.4929
		RMSE	0.1491	0.5064	0.5084	0.4998
		S.e	0.1491	0.0835	0.0831	0.0831
0.0	$\hat{\beta}_2$	**Bias**	-0.0046	-0.5018	-0.5039	-0.5032
		RMSE	0.1467	0.5092	0.5112	0.5098
		S.e	0.1466	0.0866	0.0863	0.0814
	$\hat{\beta}_3$	**Bias**	0.0008	-0.5008	-0.5029	-0.5023
		RMSE	0.1537	0.5071	0.5091	0.5084
		S.e	0.1537	0.0795	0.0791	0.0786
20% outiers						
	Coef.	**Parameter**	**LS**	**RIDGE**	**RLAV**	**RLMS**
	$\hat{\beta}_1$	**Bias**	1.7531	-0.9020	-0.9701	-0.9696
		RMSE	51.923	1.7526	0.9862	0.9796
		S.e	51.893	1.5027	0.1774	0.1401
0.5	$\hat{\beta}_2$	**Bias**	-1.4783	-1.0238	-0.9686	-0.9704
		RMSE	67.532	2.2031	0.9772	0.9761
		S.e	67.515	1.9507	0.1293	0.1057
	$\hat{\beta}_3$	**Bias**	-0.2814	-0.9854	-0.9757	-0.9694
		RMSE	48.457	1.7307	0.9896	0.9793
		S.e	48.456	1.4227	0.1655	0.1388
20% outiers						
	Coef.	**Parameter**	**LS**	**RIDGE**	**RLAV**	**RLMS**
	$\hat{\beta}_1$	**Bias**	2.1057	-0.9066	-0.9751	-0.9744
		RMSE	62.365	1.7566	0.9907	0.9842
		S.e	62.330	1.5046	0.1755	0.1383
	$\hat{\beta}_2$	**Bias**	-1.7756	-1.0289	-0.9733	-0.9745
0.99		**RMSE**	81.113	2.2077	0.9817	0.9800
		S.e	81.094	1.9533	0.1281	0.1045
	$\hat{\beta}_3$	**Bias**	-0.3379	-0.9904	-0.9804	-0.9741
		RMSE	58.203	1.7349	0.9940	0.9837
		S.e	58.202	1.4245	0.1638	0.1370

Table 4 Bias, RMSE and SE of $\hat{\beta}_1$, $\hat{\beta}_2$ and $\hat{\beta}_3$ with error normal (0,1) distribution of the sample size n=100 correlation 0.0, 0.5, 0.99 outliers 0% and 20%

			0% outliers			
Value of ρ	**Coef.**	**Parameter**	**LS**	**RIDGE**	**RLAV**	**RLMS**
		Bias	0.0055	-0.4987	-0.4991	-0.4984
	$\hat{\beta}_1$	**RMSE**	0.1032	0.5018	0.5022	0.5015
		S.e	0.1031	0.0562	0.0561	0.0557
0.0		**Bias**	0.0146	-0.496	-0.4964	-0.4951
	$\hat{\beta}_2$	**RMSE**	0.1078	0.4991	0.4995	0.4982
		S.e	0.1068	0.0557	0.0556	0.0551
		Bias	-0.0040	-0.5003	-0.5007	-0.4953
	$\hat{\beta}_3$	**RMSE**	0.1023	0.5039	0.5043	0.4988
		S.e	0.1022	0.0604	0.0603	0.0596
			20% outiers			
	Coef.	**Parameter**	**LS**	**RIDGE**	**RLAV**	**RLMS**
		Bias	-1.5448	-1.0143	-0.9750	-0.9695
	$\hat{\beta}_1$	**RMSE**	33.827	1.4107	0.9969	0.9856
		S.e	33.7918	0.9806	0.2077	0.1774
0.5		**Bias**	3.2966	-0.8810	-0.9654	-0.9651
	$\hat{\beta}_2$	**RMSE**	48.607	1.6624	0.9812	0.9776
		S.e	48.495	1.4098	0.1752	0.1559
		Bias	-1.8926	-1.0194	-0.9753	-0.9683
	$\hat{\beta}_3$	**RMSE**	32.248	1.3915	0.9941	0.9851
		S.e	32.193	0.9472	0.1923	0.1810
			20% outiers			
	Coef.	**Parameter**	**LS**	**RIDGE**	**RLAV**	**RLMS**
		Bias	-1.8554	-1.0192	-0.9798	-0.9743
	$\hat{\beta}_1$	**RMSE**	40.630	1.4151	1.0008	0.9895
		S.e	40.588	0.9816	0.2041	0.1725
0.99	$\hat{\beta}_2$	**Bias**	3.9596	-0.8858	-0.9706	-0.9703
		RMSE	58.383	1.6661	0.9857	0.9822
		S.e	58.2486	1.4112	0.1719	0.1522
	$\hat{\beta}_3$	**Bias**	-2.2732	-1.024	-0.9800	-0.9820
		RMSE	38.734	1.3959	0.9980	0.9977
		S.e	38.667	0.9482	0.1886	0.1759

Table 5 MSE ratios of 6 pairwise estimators of $\hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3$ with errors normal (0,1) distribution and 0% and 20% of outliers.

		$\hat{\beta}_1$			$\hat{\beta}_2$			$\hat{\beta}_3$		
Estimator 1 vs Estimator 2 vs Estimator 3		**Values of** ρ								
		0.0	**0.5**	**0.99**	**0.0**	**0.5**	**0.99**	**0.0**	**0.5**	**0.99**
	LS	11.24	0.00	0.00	12.08	0.00	0.00	10.94	0.00	0.00
		23.59	**0.00**	**0.00**	**21.37**	**0.00**	**0.00**	**23.77**	**0.00**	**0.00**
	RIDGE	0.97	0.31	0.31	1.00	0.20	0.20	1.01	0.32	0.32
RLMS		**1.00**	**0.49**	**0.49**	**1.00**	**0.35**	**0.35**	**0.98**	**0.50**	**0.51**
	RLAV	0.97	0.99	0.99	0.99	1.00	1.00	1.00	0.98	0.98
		1.00	**0.98**	**0.98**	**0.99**	**0.99**	**0.99**	**0.98**	**0.98**	**1.00**
	LS	11.63	0.00	0.00	12.15	0.00	0.00	10.97	0.00	0.00
RLAV		**23.66**	**0.00**	**0.00**	**21.48**	**0.00**	**0.00**	**24.30**	**0.00**	**0.00**
	RIDGE	1.01	0.32	0.32	1.01	0.20	0.20	1.01	0.33	0.33
		1.00	**0.50**	**0.50**	**1.00**	**0.35**	**0.35**	**1.02**	**0.51**	**0.51**
RIDGE	**LS**	11.53	0.00	0.00	12.05	0.00	0.00	10.89	0.00	0.00
		23.62	**0.00**	**0.00**	**21.45**	**0.00**	**0.00**	**24.26**	**0.00**	**0.00**

The measure of convergence was computed as the number of times estimator 2 or 3 to the true parameter while the value in Tables 3 and Table 4 show the summary statistics such as bias, root of mean squared error (RMSE) and standard error (SE) of the estimators of the normal distributions for sample size 50 and 100 with 0% and 20% of outliers and different value of ρ.

On the other hand, when we apply these methods to the simulated data with different sample size in the presence of different percentage of outliers and different levels of multicollineaity, we obtained that the standard errors for proposed method are less than the standard errors of the mentioned existing methods.

Table 5 shows the efficiency of the estimators by looking at the MSE ratios of the estimators written as follows.

$$MSE_{ratios} = \frac{\text{RMSE (proposed method)}}{\text{RMSE (existing method)}} \quad (7)$$

MSE_{ratios} less than 1 denote that the estimator is more efficient, however, values more than 1 denote that the other estimator is more efficient.

From Tables 3 and Table 4 we can see that the RMSE of the LS is relatively smaller than the other estimators when the errors are normally distributed that is, without outliers and no multicollinearity.

As expected, the LS gives the best result in the ideal situation. Also, the result in Table 5 is in favor of LS. However, we see in Table 5 that the of RLMS to LS is greater than 1.00 denoting that the LS is more efficient than the RLMS when no outliers and no multicolliearity.

On the other hand, we can see from the Tables 3 and Table 4 that the RMSE of the RIDGE is relatively smaller than the RLAV. Also, the MSE ratios of the estimators the values of ridge less than 1 indicates that the estimator is less efficient than RLAV and RLMS when the errors are normally distributed without outliers and no multicollinearity.

While, for non-normal error distribution and when correlation and outliers are present in the data, RLMS is better than the LS, RIDGE and RLAV. RLAV its almost as good as RIDGE and LS.

The MSE in Table 5 supported the result obtained from Tables 3 and Table 4. These ratios show the efficiency of RLMS relative to other estimators. Values less than one indicate that RLTS is more efficient, however, values greater than one denote that the other estimators are more efficient than
RLMS.

Consequently, we can see that the RMSE of the RLMS is relatively smaller than the other estimators when the errors are normally distributed in the presence of outliers and multicollinearity. It obviously shows that RLMS is more efficient than RLAV and RIDGE but certainly much more efficient than LS when in the presence of outliers and multicollinearity.

The simulation results for larger samples, that is for n=100 are consistent with the results of the smaller sized samples. The results also denoted that the estimator for larger samples are more efficient than those of smaller samples since the values of RMSE are smaller.

4. Discussions

Multicollinearity data sets with outliers are very common in real life endeavors. In order to remedy both problems, robust biased estimation methods are applied. The best model is chosen by looking at the RMSE value.

The simulation study in section 5 provide RMSE values, bias and standard errors of the LS, ridge, RLAV and RLMS estimators. It can be observed that the RMSE obtained from RLMS is the minimum. Thus, RLMS is the best method. When there is no outliers and multicollnearity, ridge regression has the least RMSE value, thus the best method. But when there is multicollinearity and outliers in the data, then RLMS has the least RMSE value, thus it is considered the best method.

We also use a the real-life data set to study the effect of multicollinearity and outliers. The results obtained from our proposed method of RLMS are better than the OLS, Ridge and RLAV in terms of their RMSE values. Consequently, in this study, it is shown that results obtained from the numerical data set and simulation study with both multicollinearity and outliers, RLMS gives better results followed by RLAV method. This result is true for ρ=0.0, 0.5 and 0.99 with the sample size of n=50 and 100.

Acknowledgements:

This work was supported by the Universiti Teknologi Malaysia, under the Research University Grant (QJ130000.2526.6 H 68) and Ministry of Higher Education (MOHE) of Malaysia.

Corresponding Author:
Ascot. Dr. Robiah Adnan
Department of Mathematical Sciences, Faculty of Science/Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Johor, Malaysia.
E-mail: robiaha@utm.my

References

1. Hoerl, A E and Kennard R W. Ridge Regression: Biased Estimation for Non-orthogonal Problems Technometrics.1970a; 12 (55).
2. Hoerl, A E and Kennard R. W. Ridge Regression: Applications for Non- orthogonal Problems Technometrics. 1970b; 12 (69).
3. Pariente S and Welsch R E. Ridge and Robust Regression Using Parametric Linear Programming Working Paper MIT Alfred P Sloan School of Management. 1977.

4. Askin R G and Montgomery D. C. Augmented robust estimators technometrics. 1980; 22: 333-341.
5. Pfaffenberger, R. C. and T. E. Dielman. A Modified Ridge Regression Estimator Using the Least Absolute Value Critrion in the Multiple Linear Regression Model. Proceedings of the American Institute for Decision Sciences. Toronto. 1984; 791-793.
6. Pfaffenberger, R. C. and T. E. Dielman. A Comparison of Robust Ridge Estimators. Proceedings of the American Statistical Association Business and Economic Statistics Section. Las Vegas. 1985; 631-635.
7. Nikolai Krivulin. An analysis of the Least Median of Squares regression problem. Proc. 10th Symp. In Computational Statistics, Switzerland 1992; 1: 471-476.
8. Desire l. Massart and Leonard Kaufman. A robust method for outlier and model error detection in regression and calibration. Analytica chimica Acta. 1986;187: 171-179.
9. Rousseeuw, P.J. Least median of squares regression. J. Am. Stat. Assoc. 1984; 79: 871-880.
10. Rousseeuw, P.J. Multivariate estimation with high breakdown point. Math. Stat. Appli. 1985; 283-297.
11. Hampel, F. R. Beyond location parameters: Robust concepts and methods proceedings of the 40th session of the ISI. 1975; 46 (1) 375–391.
12. Hoerl, A.E. and Kennard, R.W. Ridge regression: iterative estimation of the biasing parameter. Commun. Statist. Theory Meth. 1976; 5 (1): 77-88.
13. Marquardt, D. W. and Snee, R. D. Ridge regression in practice. Amer. Statist. 1975; (29) 3-19.
14. Hoerl, A. E. and K. F. Baldwin. Ridge Regression: Some Simulations. Communications in Statistics. 1975; (4): 104-123.
15. Rousseeuw P.J., and A.M. Leroy. Robust Regression and Outlier Detection. Wiley-Interscience, New York, 1987.
16. Venables, W.N. and Ripley, B.D. Modern Applied Statistics with S-PLUS. 4th ed. Springer: New York, 1999; pp. 167-176.
17. Martin, R. D. Robust Statistics with the S-Plus Robust Library and Financial Applications. Vols 1 and. Insightful Corp Presentation, New York, 2002; 17-18.
18. Pfaffenberger, R.C. and Dielman, T.E. A comparison of regression estimators when both multicollinearity and outliers are present. Edit by: Arthur, Marcel Dekker, Inc. New York and Basel. 1990.
19. Hald, Statistical Theory with Engineering Applications, Wiley, New York. 1952; p. 647.
20. Siti Meriam Zahari, Mohammad said Zainol and Muhammad Iqbal al-Banna bin Ismail Weighted Ridge MM-Estimator in Robust Ridge Regression with Multicollinearity. Mathematical Models and Methods in Modern Science 2012; ISBN:978-1-61804-106-7.

10/24/2014