Journal of Mathematics and Statistics 5 (1): 1-9, 2009
ISSN 1549-3644

The Performance of Latent Root-M based Regression

[1]Habshah Midi and [2]Lau Ung Hua
[1]Laboratory of Applied and Computational Statistics, Institute for Mathematical Research, University Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia
[2] Faculty of Science, University of Technology Mara, Sarawak, Malaysia

Abstract: Problem statement: In the presence of multicollinearity, the estimation of parameters in multiple linear regression models by means of Ordinary Least Squares (OLS) is known to suffer severe distortion. An alternative approach was to use the modified OLS which was based on the latent roots and latent vectors of the correlation matrix of the independent and dependent variables. This procedure is called the Latent Root Regression (LRR) which serves the purpose to improve the stability of the estimates for data plagued by multicollinearity. However, there was evidence that the LRR estimators were easily affected by a few atypical observations that we call outliers. It is now evident that the robust method alone cannot rectify the combined problems of multicollinearity and outliers. **Approach:** In this study, we proposed a robust procedure for the estimation of the regression parameters in the presence of multicollinearity and outliers. We called this method Latent Root-M based Regression (LRMB) because here we employed the weight of the M-estimator in the weighted correlation matrix. Numerical examples and some simulation studies were presented to illustrate the performance of the newly proposed method. **Results:** Results of the study showed that the LRMB method is more efficient than the existing methods. **Conclusion/Recommendations:** In order to get a reliable estimate, we recommend using the LRMB when both multicollinearity and outliers are present in the data.

Key words: latent root regression, M-estimator, multicollinearity, outliers

INTRODUCTION

Consider a multiple linear regression model:

$$Y = X\beta + \varepsilon \quad (1)$$

Where:
- Y = The n×1 vector of standardized dependent variables
- X = The nxk full rank matrix of standardized known constants
- Xβ = The k×1 vector of model parameters
- ε = The n×1 vector of random disturbances with ε~NID (0, σ^2)
- p = The number of independent variables
- n = The number of observations.

Using the least squares criterion, the estimator of β are found by minimizing the sum of squares residuals:

$$Q = \sum_{i=1}^{n} r_i(\beta)^2$$

where, $r_i(\beta) = Y - X\beta$. This gives the OLS estimator for β:

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

According to the Gauss- Markov Theorem, the OLS estimators, in the class of unbiased linear estimators, have minimum variance that is they are Best Linear Unbiased Estimator (BLUE). Nonetheless, the presence of multicollinearity will produce inflated standard errors that will lead to misleading parameter inferences. To remedy this problem, Hawkins[1], Gunst and Mason[2], Gunst *et al*.[3] and Lawrence and Arthur[4] have introduced a new biased estimation procedures known as Latent Root Regression (LRR) to improve the precision of the regression estimates. The major advantage of LRR is that it is not only identifies the multicollinearities present in the independent variables, but also allows the researcher to distinguish between predictive and non predictive multicollinearity, hence appropriately adjust the OLS estimates for the non

Corresponding Author: Habshah Midi Laboratory of Applied and Computational Statistics, Institute for Mathematical Research, University Putra Malaysia, 43400 Serdang, Selangor, Malaysia

predictive multicollinearities. However, this technique is inefficient if the underlying disturbances are not normal, which may arise as a result of outliers. As an alternative, we may turn to robust methods which are not sensitive to the presence of outliers[5-8]. Nevertheless, robust method alone cannot overcome the combined problem of multicollinearity and outliers. In this study, we propose a Robust Latent Root Regression to rectify these two problems simultaneously by using Latent Root Regression based on robust weighted correlation matrix.

MATERIALS AND METHODS

The Latent Root Regression (LRR): The latent root regression utilizes the latent roots and latent vectors of the correlation matrix of the dependent and independent variables, denoted as A. The latent roots, λ_j and latent vectors, γ_j of A'A are defined by:

$$\left|A'A - \lambda_j I\right| = 0 \text{ and } \left(A'A - \lambda_j I\right)\gamma_j = 0 \quad j = 0, 1, \ldots, k$$

Analysis of these latent roots and latent vectors enables one to:

- Identify near singularities in X
- Determine whether the near singularities have predictive value
- Obtain the modified least squares estimates of parameters which adjust for non-predictive near singularities

The OLS estimator in (2) can also be expressed in terms of these latent roots and latent vectors:

$$\hat{\beta} = -\eta\sum_{j=1}^{k}\alpha_j\gamma_j^{0} \tag{3}$$

Where:

$$\eta^2 = \sum_{i=1}^{n}\left(Y_i - \bar{Y}\right)^2$$

$$\alpha_j = \frac{\gamma_{0j}\lambda_j^{-1}}{\sum_{r=0}^{k}\gamma_{0r}^{2}\lambda_r^{-1}}$$

$\gamma_j^{0\prime} = \left(\gamma_{1j}, \gamma_{2j}, \ldots, \gamma_{kj}\right)$ and the residual sum of squares given by:

$$SSE = \eta^2\left(\sum_{j=1}^{k}\frac{\gamma_{0j}^{2}}{\lambda_j}\right)^{-1}$$

Gunst *et al*.[3] and Lawrence and Arthur[4] suggested small latent roots and latent vectors in which $\lambda_j \le 0.3$ and $|\gamma_{0j}| \le 0.1$ which indicates the presence of non predictive singularities. But later, they discovered that a tighter cut-off value of $\lambda_j \le 0.2$ and $|\gamma_{0j}| \le 0.1$ could improve the analysis.

Suppose now that the latent vectors $\gamma_{0,}\ \gamma_{1,\ldots,}\ \gamma_{p-1}$ correspond to non predictive near singularities. The non predictive multicollinearities are eliminated and only the predictive are retained. The above OLS estimator can be adjusted by setting $\alpha_0 = \alpha_1 = \ldots = \alpha_{p-1} = 0$. Then the modified least squares coefficients are:

$$\hat{\beta}_{LRR} = -\eta\sum_{j=1}^{k}\alpha_j\gamma_j^{0} \tag{4}$$

Where:

$$\alpha_j = \frac{\gamma_{0j}\lambda_j^{-1}}{\sum_{r=p}^{k}\gamma_{0r}^{2}\lambda_r^{-1}} \quad j = p, p+1, \ldots, k$$

with residual sum of squares, $SSE_{LRR} = \eta^2\left(\sum_{j=p}^{k}\frac{\gamma_{0j}^{2}}{\lambda_j}\right)^{-1}$.

If all of the principal components for the correlation matrix of the dependent and independent variables are predictive, then none of the α_j's equal zero, the latent root estimator and the OLS estimator will be identical.

It is well-known that the variance covariance matrix for the OLS estimator is given by $\sigma^2\left(X'X\right)^{-1}$ and its trace (sum of diagonals) represents its un weighted mean squared error:

$$MSE(\hat{\beta}) = \sigma^2 tr\left(X'X\right)^{-1} \tag{5}$$

or in terms of latent roots of X'X

$$MSE(\hat{\beta}) = \sigma^2\sum_{j=1}^{p}\ell_j^{-1} \tag{6}$$

ℓ_j are the latent root of X'X and are ordered such that $\ell_1 \le \ell_2 \le \ldots \le \ell_k$

For a near multicollinearity situation, ℓ_1 approaches 0 and (6) implies that $MSE(\hat{\beta})$ approaches infinity, that is $\hat{\beta}$ is subjected to very large variance. This inflation cause the estimation becomes less accurate and less precise, thus unstable.

Robust M-estimator: The OLS estimation method optimizes the fit of the model by minimizing the sum of the squared deviations between the actual and predicted Y values, $\sum(y-\hat{y})^2$. The OLS method can be represented as:

$$\min \sum_{i=1}^{n} e_i^2 \tag{7}$$

Huber[5] developed a group of estimators called M-estimators, which are based on the idea of replacing the squared residuals, e_i^2, with another function of the residuals, given by:

$$\min \sum_{i=1}^{n} \rho(e_i) \tag{8}$$

where, ρ is a symmetric function with a unique minimum at zero. The robust M-estimates (ROBM) are calculated using Iteratively Re Weighted Least Squares (IRLS). In IRLS, the initial fit is calculated and then a new set of weights is calculated based on the results of the initial fit. The iterations are continued until a convergence criterion is met.

Robust latent root regression: Robust latent root regression incorporates resistance in the ordinary latent root regression. This is done by imposing weight to the correlation matrix of the dependent and independent variables, A'A.

The pair wise Pearson correlation coefficient for the two variables is defined as:

$$r = \frac{\sum_{i=1}^{n} (Y_i - \overline{Y})(X_i - \overline{X})}{\sqrt{\left(\sum_{i=1}^{n} (Y_i - \overline{Y})^2\right)\left(\sum_{i=1}^{n} (X_i - \overline{X})^2\right)}} \tag{9}$$

Where:

$$\overline{Y} = \frac{\sum_{i=1}^{n} Y_i}{n} \text{ and } \overline{X} = \frac{\sum_{i=1}^{n} X_i}{n}$$

The correlation coefficient, r in (9) is based on sample means $\overline{x}$ and $\overline{y}$, respectively, which are known to be very sensitive to the presence of outliers. As an alternative, a robust location estimates which are less affected by outliers are proposed to replace $\overline{x}$ and $\overline{y}$ in (9). Following the idea of Mokhtar[9], we propose using the weighted correlation coefficient between the dependent and the independent variables. We may use the weight from the final step of any robust estimators, but in this study the weight is confined to the final step of the robust M-estimation. The pair wise correlation coefficient in Eq. 9 is modified to obtain a weighted pair wise correlation coefficient, as follows;

$$r_w = \frac{\sum_{i=1}^{n} w_i (Y_i - \overline{Y}_w)(X_i - \overline{X}_w)}{\sqrt{\left(\sum_{i=1}^{n} w_i (Y_i - \overline{Y}_w)^2\right)\left(\sum_{i=1}^{n} w_i (X_i - \overline{X}_w)^2\right)}} \tag{10}$$

Where:

$$\overline{Y}_w = \frac{\sum_{i=1}^{n} w_i Y_i}{\sum_{i=1}^{n} w_i} \text{ and } \overline{X}_w = \frac{\sum_{i=1}^{n} w_i X_i}{\sum_{i=1}^{n} w_i}$$

In this study, we have chosen the Tukey's Biweight function in the M estimation technique[7,8]. By using (10) a robust weighted correlation matrix for dependent and independent variables, which originally denoted as A can be formulated. Based on this weighted correlation matrix, the latent roots and the latent vectors are computed and the latent root routines are then incorporated to estimate the parameters of the model. We call this method the Latent Root- M based Regression (LRMB) because here we have employed the weight of the M-estimator in the weighted correlation matrix. We would expect the modified method to be more robust than the OLS, ROBM and LRR.

RESULTS

Numerical example: In order to compare the performance of the LRMB with the other existing methods such as the OLS, LRR and ROBM, two real data sets are considered. The first data set presents the Palm Oil data which is taken from the Malaysian Palm Oil Board[10]. The dependent variable is the palm oil annual export (tonnes) while the independent variables are oil palm planted area (hectares) and crude palm oil production (tonnes). By incorporating the weight obtained from the final step of the ROBM estimator, yields the robust-weighted correlation matrix with the corresponding latent roots and latent vectors which are displayed in Table 1 and 2, respectively. The presence of outlier in the data was detected by using Robust Mahalanobis Distance (RMD)[7,11] The standard error, confidence interval length and the R^2 of the four methods are shown in Table 3.

Table 1: The robust-weighted correlation matrix

Y	X_1	X_2
1.0000	0.9918	0.9923
0.9918	1.0000	0.9898
0.9923	0.9898	1.0000

Table 2: The latent roots and latent vectors of the robust-weighted correlation matrix

	1	2	3
λ_j	2.9826	0.0102	0.0072
γ_j	0.5776	0.0864	0.8117
	0.5772	-0.7464	-0.3313
	0.5773	0.6598	-0.4810

Table 3: The standard error, confidence interval length and the R^2 for oil-palm data

		β_1	β_2	R^2
OLS	Est.	7.204	-1.5403	
	S.E	3.2113	0.9187	0.5104
	t	2.2433	-1.6765	
	C.I	(0.482,13.925) [13.443]	(-3.463,0.383)[3.846]	
ROBM	Est.	1.3212	0.4346	
	S.E	0.5261	0.1505	0.9824
	t	2.5111	2.8872	
	C.I	(0.220,2 .422) [2.202]	(0.119,0.749) [(0.63]	
LRR	Est.	0.7771	0.2979	
	S.E	0.295	0.0979	0.4067
	t	2.634	3.043	
	C.I	(0.160,1.395) [1.235]	(0.093,0.503) [0.410]	
LRMB	Est.	1.1402	0.4873	
	S.E	0.2305	0.0956	0.9891
	t	4.9466	5.0973	
	C.I	(0.658,1.623)[0.965]	(0.287,0.687) [0.400]	

Table 4: The standard error, confidence interval length and the R^2 for Gujarati data

		β_1	β_2	R^2
OLS	Est.	0.9415	-0.0424	
	S.E	0.8229	0.0807	0.9635
	t	1.1442	-0.5261	
	C.I	(-1.004, 2.887) [3.891]	(-0.233,0.148) [0.381]	
ROBM	Est.	1.0132	-0.0503	
	S.E	0.8900	0.0872	0.9631
	t	1.1384	-0.5762	
	C.I	(-1.091, 3.118) [4.209]	(-0.257,0.156) [0.413]	
LRR	Est.	0.2271	0.0276	
	S.E	0.0822	0.0099	0.9595
	t	2.7613	2.7778	
	C.I	(0.033,0.422) [0.389]	(0.004,0.051) [0.047]	
LRMB	Est.	0.2177	0.0278	
	S.E	0.075	0.0095	
	t	2.9017	2.9246	0.9605
	C.I	(0.040,0.395) [0.355]	(0.005,0.050) [0.045]	

The performances of these four estimators are further examined by applying these estimators to another data set which is taken from Gujarati [12] where consumption expenditure being the dependent variable while the independent variables are income and wealth.

Table 4 exemplified the standard error, confidence interval length and the R^2 of the Gujarati's data. The confidence interval lengths for Table 3 and 4 are in square bracket.

Simulation study: A simulation study similar to that of Lawrence and Arthur[4] has been performed in order to compare the performance of the four estimators. The model used was $y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \varepsilon_i$.

The parameter values β_0, β_1 and β_2 were set equal to one. The explanatory variables x_{i1} and x_{i2} were generated as follows:

$$x_{ij} = \left(1-\rho^2\right) z_{ij} + \rho z_{ij} \quad i = 1,2,\ldots, n;\ j = 1, 2$$

where, z_{ij} are independent standard normal random variables. The value of ρ^2 were chosen as 0.0, 0.5 and 0.95 and they represent the correlation between the two independent variables. Sample sizes, n of 25 and 50 (each corresponding to small and large sample) were examined. Four error disturbances were employed as follows:

- Standard normal distribution
- Cauchy distribution with median zero and scale parameter one
- t-Student distribution with three degrees of freedom
- Contaminated normal distribution where the underlying distribution is standard normal with probability 0.85 and normal with mean zero and standard deviation five with probability 0.15

The non-normal distribution, such as the Cauchy and student-t with 3 degrees of freedom are symmetrical bell-shaped with heavy tailed distribution which prone to produce considerable amount of outliers. These distributions were generated to investigate the effect of combined problems of multicollinearity and outliers on different estimators.

All the four methods were then applied to each of the sets of the generated data. In each simulation run, there were 1000 replications. Some summary statistics such as the bias, Standard Errors (SE) and Root Mean Squared Errors (RMSE) over 1,000 runs were computed and the results are exemplified in Table 5, 7, 9 and 11. Table 6, 8, 10 and 12 show the efficiency of the estimators by observing at the MSE ratios of two estimators. Values less than one indicate that the first estimator is more efficient than the second, values equal to one imply that both estimators are equally good, while values more than one indicate that the second estimator is more efficient than the first estimator.

The values in all Table 1-12 are for sample size 25 while values for n = 50 are shown in bold.

Table 5: Bias, RMSE and SE of $\hat{\beta}_1$ and $\hat{\beta}_2$ with disturbance distribution normal (0,1)

	Values of ρ^2								
	0.0			0.5			0.95		
Method	Bias	RMSE	SE	Bias	RMSE	SE	Bias	RMSE	SE
$\hat{\beta}_1$									
OLS	-0.0157	0.2174	0.2169	-0.0088	0.3327	0.3326	-0.09730	3.1103	3.1088
	0.0017	0.1451	0.1451	-0.0026	0.2284	0.2284	-0.05070	2.0939	2.0933
ROBM	-0.0158	0.2292	0.2287	-0.0098	0.3491	0.3489	-0.13180	3.2573	3.2547
	0.0016	0.1514	0.1514	-0.0003	0.2347	0.2347	-0.03330	2.1418	2.1415
LRR	-0.0157	0.2174	0.2169	-0.0097	0.3328	0.3327	0.01240	0.2090	0.2087
	0.0017	0.1451	0.1451	-0.0023	0.2287	0.2287	0.00280	0.0808	0.0808
LRMB	-0.0158	0.2292	0.2287	-0.0102	0.3492	0.3491	-0.00260	0.4283	0.4282
	0.0016	0.1514	0.1514	-0.0003	0.2349	0.2349	0.00360	0.0889	0.0888
$\hat{\beta}_2$									
OLS	-0.0051	0.2258	0.2258	0.0091	0.3496	0.3495	0.10400	3.1226	3.1209
	0.0069	0.1484	0.1482	0.0094	0.2303	0.2301	0.05550	2.0927	2.0920
ROBM	-0.0041	0.2361	0.2361	0.0133	0.3629	0.3626	0.13940	3.2668	3.2638
	0.0054	0.1526	0.1525	0.0083	0.2361	0.2359	0.03953	2.1414	2.1410
LRR	-0.0051	0.2258	0.2258	0.0099	0.3499	0.3498	-0.00350	0.2013	0.2013
	0.0069	0.1484	0.1482	0.0091	0.2303	0.2302	0.00130	0.0836	0.0836
LRMB	-0.0041	0.2361	0.2361	0.0137	0.3631	0.3628	0.01240	0.4277	0.4275
	0.0054	0.1526	0.1525	0.0083	0.2363	0.2362	0.00180	0.0895	0.0895

Table 6: MSE ratios of 6 pair wise estimators of $\hat{\beta}_1$ and $\hat{\beta}_2$ with disturbance distribution normal (0,1)

		Values of ρ^2					
		$\hat{\beta}_1$			$\hat{\beta}_2$		
Estimator1 vs	Estimator 2	0.0	0.5	0.95	0	0.5	0.95
LRMB	OLS	1.11	1.10	0.02	1.09	1.08	0.02
		1.09	1.06	0.00	1.06	1.05	0.00
	ROBM	1.00	1.00	0.02	1.00	0.02	0.02
		1.00	1.00	0.00	1.00	1.00	0.00
	LRR	1.11	1.10	4.20	1.09	1.08	4.51
		1.09	1.06	1.21	1.06	1.05	1.15
LRR	OLS	1.00	1.00	0.00	1.00	1.00	0.00
		1.00	1.00	0.00	1.00	1.00	0.00
	ROBM	0.90	0.91	0.00	0.91	0.93	0.00
		0.92	0.95	0.00	0.95	0.95	0.00
ROBM	OLS	1.11	1.10	1.10	1.09	1.08	1.09
		1.09	1.06	1.05	1.06	1.05	1.05

DISCUSSION

Here we discuss the results that we have acquired from the previous section. The result of Table 1 suggests that the oil palm planted areas and oil production are highly correlated. The presence of two outliers in this data were detected based on RMD. By the application of LRR techniques, the vector deletion criteria in which latent vectors are deleted if $\lambda_\varphi \leq 0.2$ and $|\gamma_{0j}| \leq 0.1$, leads to the deletion of the second latent vector from the robust-weighted correlation matrix of Table 2. By this deletion, the LRMB has substantially reduced the standard error of the estimates. It can be observed from Table 3 that the OLS estimates have been strongly affected by outliers and multicollinearity. This is indicated by its largest standard error among the four estimates, smaller R^2 value and negative coefficient of $\hat{\beta}_2$. Moreover, it possesses confidence interval length which is remarkably larger than the other intervals. The performance of the ROBM and the LRR are also not encouraging since their standard errors and confidence interval lengths are still relatively large. However, the RLMB can be considered the best method because it has the smallest standard errors and confidence interval length and the highest R^2 value than the other three estimators.

Table 7: Bias, RMSE and SE of $\hat{\beta}_1$ and $\hat{\beta}_2$ with disturbance distribution Cauchy

	Values of ρ^2								
	0.0			0.5			0.95		
Method	Bias	RMSE	SE	Bias	RMSE	SE	Bias	RMSE	SE
$\hat{\beta}_1$									
OLS	2.9485	126.3610	126.3270	1.1476	47.5185	47.5047	6.2723	263.8537	263.7791
	0.1949	18.5650	18.5639	0.3225	32.0980	32.0964	5.1771	340.4129	340.3735
ROBM	0.0259	0.4512	0.4505	0.0449	0.6850	0.6835	0.3998	6.0744	6.0612
	-0.0024	0.2681	0.2681	-0.0151	0.4207	0.4204	-0.2165	3.8499	3.8439
LRR	2.9485	126.3610	126.3270	1.5514	45.5053	45.4789	0.4533	12.7839	12.7759
	0.1949	18.5650	18.5639	0.2991	32.0867	32.0853	-0.1244	5.6074	5.6060
LRMB	0.0259	0.4512	0.4505	0.0425	0.6849	0.6836	0.0075	0.2950	0.2949
	-0.0024	0.2681	0.2681	-0.0152	0.4209	0.4207	0.0053	0.1471	0.1471
$\hat{\beta}_2$									
OLS	3.9460	144.0250	143.9716	1.7290	58.1426	58.1169	-5.3485	245.9591	245.9009
	-0.2085	17.3463	17.3450	-0.6186	36.8859	36.8807	-5.5099	344.5208	344.4767
ROBM	-0.0175	0.4663	0.4660	-0.0345	0.6916	0.6907	-0.3904	6.0919	6.0793
	0.0144	0.2810	0.2806	0.0253	0.4285	0.4278	0.2229	3.8610	3.8546
LRR	3.9360	144.0250	143.9716	1.2933	58.2196	58.2052	0.5077	14.0675	14.0584
	-0.2085	17.3463	17.3450	-0.5975	36.8847	36.8800	-0.1184	5.6021	5.6008
LRMB	-0.0175	0.4663	0.4660	-0.0319	0.6899	0.6892	-0.0015	0.2847	0.2847
	0.0144	0.2810	0.2806	0.0254	0.4283	0.4275	0.0017	0.1469	0.1469

Table 8: MSE ratios of 6 pairwise estimators of $\hat{\beta}_1$ and $\hat{\beta}_2$ with disturbance distribution Cauchy

		Values of ρ^2					
		$\hat{\beta}_1$			$\hat{\beta}_2$		
Estimator1 vs	Estimator 2	0.0	0.5	0.95	0.0	0.5	0.95
LRMB	OLS	0.00	0.00	0.00	0.00	0.00	0.00
		0.00	0.00	0.00	0.00	0.00	0.00
	ROBM	1.00	1.00	0.00	1.00	1.00	0.00
		1.00	1.00	0.00	1.00	1.00	0.00
	LRR	0.00	0.00	0.00	0.00	0.00	0.00
		0.00	0.00	0.00	0.00	0.00	0.00
LRR	OLS	1.00	0.92	0.00	1.00	1.00	0.00
		1.00	1.00	0.00	1.00	1.00	0.00
	ROBM	78431.53	4413.09	4.43	95399.80	7086.45	5.33
		4795.08	5817.08	2.12	3810.67	7409.53	2.11
ROBM	OLS	0.00	0.00	0.00	0.00	0.00	0.00
		0.00	0.00	0.00	0.00	0.00	0.00

Let us now focus to the Gujerati's data. There is evidence that income and wealth for the Gujarati's data are highly correlated. This data has no outlier but has multicollinearity problem. Since this data has only multicollinearity problem, we expect that the performance of the LRMB is closed to the LRR. It is interesting to note that the results of Table 4 are consistent with the earlier findings except that the LRR and LRMB are equally good as expected because when there is no outlier and only multicollinearity exist, the LRMB become closer to LRR. We have not scrutinized the analysis of the example to the final conclusion, but a reasonable explanation up to this point is that the LRMB is not easily affected by the presence of both multicollinearity and outliers.

Next, we will discuss the simulation results obtained from the standard normal and heavy tail distributions whether they confirm the conclusion of the numerical examples.

Standard normal distribution of disturbances: Table 5 shows that for standard normal disturbances with $\rho = 0$, all four methods are virtually indistinguishable with respect to the values of the bias, SE and RMSE. The performance of the OLS and the LRR are slightly better than ROBM and LRMB for small ρ-value.

Table 9: Bias, RMSE and SE of $\hat{\beta}_1$ and $\hat{\beta}_2$ with disturbance distribution t-student (3)

	Values of ρ^2								
	0.0			0.5			0.95		
Method	Bias	RMSE	SE	Bias	RMSE	SE	Bias	RMSE	SE
$\hat{\beta}_1$									
OLS	0.01480	0.3630	0.3627	0.03150	0.5736	0.5727	0.2231	5.1562	5.1514
	-0.00960	0.2435	0.2433	-0.01530	0.3815	0.3812	-0.1524	3.4542	3.4508
ROBM	0.01030	0.2850	0.2848	0.02660	0.4618	0.4610	0.1859	4.1788	4.1746
	-0.00780	0.1876	0.1875	-0.00920	0.2919	0.2918	-0.0877	2.6684	2.6670
LRR	0.01480	0.3630	0.3627	0.03910	0.5738	0.5725	0.0031	0.1980	0.1980
	-0.00960	0.2435	0.2433	-0.01690	0.3817	0.3813	0.0028	0.1346	0.1345
LRMB	0.01030	0.2850	0.2848	0.02685	0.4620	0.4612	-0.0118	0.4162	0.4160
	-0.00780	0.1876	0.1875	-0.00970	0.2923	0.2921	0.0025	0.1091	0.1091
$\hat{\beta}_2$									
OLS	-0.00300	0.3613	0.3613	-0.01790	0.5636	0.5633	-0.2139	5.1522	5.1480
	0.00140	0.2573	0.2573	0.01540	0.3811	0.3808	0.1560	3.4463	3.4428
ROBM	-0.00890	0.2942	0.2941	-0.01340	0.4533	0.4531	-0.1710	4.1738	4.1703
	0.00150	0.1896	0.1896	0.00930	0.2925	0.2924	0.0895	2.6698	2.6683
LRR	-0.00300	0.3613	0.3613	-0.02420	0.5588	0.5582	0.0088	0.2043	0.2041
	0.00140	0.2573	0.2573	0.01690	0.3806	0.3802	0.0002	0.1340	0.1340
LRMB	-0.00890	0.2942	0.2941	-0.01360	0.4535	0.4533	0.0265	0.4271	0.4262
	0.00150	0.1896	0.1896	0.00980	0.2926	0.2925	-0.0007	0.1075	0.1075

Table 10: MSE ratios of 6 pairwise estimators of $\hat{\beta}_1$ and $\hat{\beta}_2$ with disturbance distribution t-student (3)

		Values of ρ^2					
		$\hat{\beta}_1$			$\hat{\beta}_2$		
Estimator1 vs	Estimator 2	0.0	0.5	0.95	0.0	0.5	0.95
LRMB	OLS	0.62	0.65	0.01	0.66	0.65	0.01
		0.59	0.59	0.00	0.54	0.59	0.00
	ROBM	1.00	1.00	0.01	1.00	1.00	0.01
		1.00	1.00	0.00	1.00	1.00	0.00
	LRR	0.62	0.65	4.42	0.66	0.66	4.37
		0.59	0.59	0.66	0.54	0.59	0.64
LRR	OLS	1.00	1.00	0.00	1.00	0.98	0.00
		1.00	1.00	0.00	1.00	1.00	0.00
	ROBM	1.62	1.54	0.00	1.51	1.52	0.00
		1.68	1.71	0.00	1.84	1.69	0.00
ROBM	OLS	0.62	0.65	0.66	0.66	0.65	0.66
		0.59	0.59	0.60	0.54	0.59	0.60

When the multicollinearity is high ($\rho = 0.95$) as to be expected, the LRR give the best results followed by the LRMB, OLS and ROBM. This result is supported by Table 6, where for high correlation; the LRR is more efficient than LRMB indicated by the value of the MSE ratios which is greater than one. Similarly, the MSE ratios signify that the LRR is better than the OLS and ROBM for high value of ρ. Evidently, in this situation, the OLS is better than the ROBM. The LRR estimates emerge to be conspicuously more efficient in the presence of high multicollinearity with no contamination in the model.

Heavy tails distribution of the Disturbances; Here we discuss the results of Cauchy, t with 3 degrees of freedom and contaminated normal. Let us first focus our attention to Table 7 and 8, for cauchy distribution. The results in Table 7 show that when there is no multicollinearity ($\rho = 0.0$) for this type of data with only the presence of outliers, as can be expected the performance of the ROBM is similar to that of RLMB. The OLS is as good as the LRR and their performance are less efficient than the LRMB and ROBM. For small correlation ($\rho = 0.5$), the LRMB is slightly better than the ROBM estimates and they are more efficient than the OLS and LRR. The presence of both outlier and high multicollinearity changes the situation dramatically. The biases and the RMSE of the OLS, LRR and ROBM estimates increase significantly. On the other hand, the LRMB is not affected by the outliers and multicollinearity, as shown by the biases and the RMSE which were decreasing and consistently the smallest among the four estimators. It is evident that the LRMB is the best estimator followed by the ROBM, LRR and OLS. The MSE ratios in Table 8 supported the results obtained from Table 7 where for skewed data with small and no multicollinearity, the ROBM is fairly close to LRMB and their performances are much better than the LRR and OLS. The results of Table 8 signify that the LRMB seems to perform extremely well compared to ROBM, LRR and OLS for high multicollinearity, evidenced by the values of the MSE ratios which are less than one. The LRMB and ROBM are equally efficient when ρ is zero or low indicated by the MSE ratios which are equal to one.

Table 11: Bias, RMSE and SE of $\hat{\beta}_1$ and $\hat{\beta}_2$ with disturbance distribution contaminated normal

	Values of ρ^2								
	0			0.5			0.95		
Method	Bias	RMSE	SE	Bias	RMSE	SE	Bias	RMSE	SE
$\hat{\beta}_1$									
OLS	-0.0169	0.4436	0.4433	-0.0299	0.6932	0.6925	-0.3728	6.2794	6.2683
	-0.0015	0.3281	0.3281	-0.0042	0.5041	0.5040	0.0679	4.5923	4.5918
ROBM	-0.0035	0.2879	0.2879	-0.0138	0.4350	0.4348	-0.1844	4.0224	4.0182
	0.0002	0.1867	0.1867	-0.0113	0.2839	0.2867	-0.0573	2.6140	2.6134
LRR	-0.0169	0.4436	0.4433	-0.0314	0.6896	0.6889	0.0126	0.2353	0.2350
	-0.0015	0.3281	0.3281	-0.0040	0.5039	0.5038	-0.0099	0.1632	0.1629
LRMB	-0.0035	0.2879	0.2879	-0.0149	0.4346	0.4343	-0.0012	0.4716	0.4716
	0.0002	0.1867	0.1867	-0.0112	0.2836	0.2834	-0.0055	0.1024	0.1022
$\hat{\beta}_2$									
OLS	0.0205	0.4820	0.4816	0.0473	0.6992	0.6976	0.3911	6.2654	6.2533
	-0.0087	0.3159	0.3157	-0.0238	0.4986	0.4981	-0.0872	4.5926	4.5918
ROBM	0.0127	0.3016	0.3013	0.0252	0.4482	0.4475	0.1940	4.0289	4.0242
	0.0027	0.1846	0.1846	-0.0046	0.2895	0.2895	0.0437	2.6205	2.6202
LRR	0.0205	0.4820	0.4816	0.0489	0.6957	0.6940	0.0099	0.2379	0.2377
	-0.0087	0.3159	0.3157	-0.0240	0.4970	0.4964	-0.0098	0.1632	0.1629
LRMB	0.0127	0.3016	0.3013	0.0262	0.4491	0.4483	0.0111	0.4742	0.4740
	0.0027	0.1846	0.1846	-0.0048	0.2897	0.2897	-0.0076	0.0990	0.0987

Table 12: MSE ratios of 6 pair wise estimators of $\hat{\beta}_1$ and $\hat{\beta}_2$ with disturbance distribution contaminated Normal

		Values of ρ^2					
		$\hat{\beta}_1$			$\hat{\beta}_2$		
Estimator1 vs	Estimator 2	0.0	0.5	0.95	0.0	0.5	0.95
LRMB	OLS	0.42	0.39	0.01	0.39	0.41	0.01
		0.32	0.32	0.00	0.34	0.34	0.00
	ROBM	1.00	1.00	0.01	1.00	1.00	0.01
		1.00	1.00	0.00	1.00	1.00	0.00
	LRR	0.42	0.40	4.02	0.39	0.42	3.97
		0.32	0.32	0.39	0.34	0.34	0.37
LRR	OLS	1.00	1.00	0.00	1.00	0.99	0.00
		1.00	1.00	0.00	1.00	0.99	0.00
	ROBM	2.37	2.51	0.00	2.55	2.41	0.00
		3.09	3.15	0.00	2.93	2.95	0.00
ROBM	OLS	0.42	0.39	0.41	0.39	0.41	0.41
		0.32	0.32	0.32	0.34	0.34	0.33

The results of Table 9 and 10 illustrate the summary statistics for the t distribution with 3 degrees of freedom. Like the Cauchy distribution, the performances of the ROBM and LRMB estimator are equally good for small and no multicollinearity. Similarly, the RLMB and ROBM are slightly better than the OLS and RLL in such a situation. Nevertheless, when $\rho = 0.95$ and $n = 25$, the performance of the RLL is slightly better than the RLMB. It is interesting to note that when the size of the sample is increased to 50, the RLMB is better than the RLL. These are indicated by its bias and RMSE which are smaller than the RLL in this situation. Similar results are obtained by observing the MSE ratios in Table 10. The results of Table 11 and 12 for contaminated data are consistent with the finding obtained from the t distribution.

CONCLUSION

The OLS performs poorly in the presence of outliers and multicollinearity. The ROBM is not sufficiently robust compared with LRR and LRMB when the degree of multicollinearity is high. The LRR estimator is a better choice than the other estimators in eliminating the problem of multicollinearity. However, its performance was inferior to ROBM and LRMB when contamination occurs in the data. The empirical study shows that the LRMB has improved the accuracy of the estimates in the situation when both multicollinearity and non-normal disturbances are present. The results seem to suggest that the RLMB estimator may provide a robust alternative to the LRR.

REFERENCES

1. Hawkins, D.M., 1973. On investigation of alternative regressions by principal component analysis. Applied Stat., 22: 275-286. http://www.jstor.org/pss/2346776
2. Gunst, R.F. and R.L. Mason, 1980. Regression Analysis and its Application: A data oriented approach. 1st Edn., M. Dekker, New York, ISBN: 9780824769932, pp: 424.

3. Gunst, R.F., R.L. Mason and J.T. Webster, 1976. A comparison of least squares and latent root regression estimators. Technometrics, 18: 75-83. http://www.jstor.org/pss/1267919
4. Lawrence, K.D. and J.L. Arthur, 1989. Robust Regression: Analysis and Applications.1st Edn., Marcel Dekker, New York, ISBN: 9780824781293, pp: 287.
5. Huber, P.J., 2003. Robust Statistics. 1st Edn., Wiley, New York, ISBN: 9780471650720, pp: 328.
6. Rousseeuw, P.J. and A.M. Leroy, 1987. Robust Regression and Outlier Detection, 1st Edn., Wiley, New York, USA., ISBN: 0471852333, pp: 329.
7. Rousseeuw, P.J. and A.M. Leory, 2003. Robust Regression and Outlier Detection. 1st Edn., Wiley, New York, ISBN: 978-0471488552, pp: 360.
8. Maronna, R., 2006. Robust Statistics. 1st Edn., Wiley, New York, USA., ISBN: 978-0470010921, pp: 436.
9. Mokhtar, A., 1990. On a robust correlation coefficient. Statistician, 39: 455-460. http://cat.inist.fr/?aModele=afficheN&cpsidt=5488438
10. Malaysian Palm Oil Board, 2001. Malaysian Oil Palm Statistics 2000. 21st Edn., Ministry of Primary Industries, Malaysia, pp:128.
11. Hussain, S., M.A. Mohamed, R. Holder, A. Almasri, G. Shukur, 2008. Performance evaluation based on the robust mahalanobis distance and multilevel modeling using two new strategies. Commun. Stat. Simulat. Comput., 37: 1966-1980. DOI: 10.1080/03610910802311692
12. Gujerati, D., 2002. Basic Econometrics. 4th Edn., McGraw-Hill, New York, USA., ISBN: 0071123431, pp: 1002.