Journal of Physics: Conference Series

PAPER • OPEN ACCESS

Developing a New Estimator in Linear Regression Model

Developing a New Estimator in Linear Regression Model

Adewale F. Lukman[1], Kayode Ayinde[2], Alabi Olatayo[2], Rasaq Bamidele[1], Benedicta B. Aladeitan[1] and Theophilus A. Adagunodo[3]

[1] Department of Mathematics, Landmark University, Omu-Aran, Kwara State, Nigeria.
[2]Department of Statistics, Federal University of Technology, Akure, Ondo State, Nigeria.
[3] Department of Physics, Covenant University, Ota, Ogun State, Nigeria

Corresponding email: adewale.folaranmi@lmu.edu.ng, kayinde@lautech.edu.ng, a.olatayoolusegunvictor@gmail.com, bamidele.rasak@lmu.edu.ng, aladeitan.benedicta@lmu.edu.ng, theophilus.adagunodo@covenantuniversity.edu.ng

Abstract. The most popularly used method of estimating the parameters in a linear regression model is the Ordinary Least Squares (OLS) Estimator. This estimator is considered best when certain assumptions are satisfied. However, when any of these assumptions fail there is a need for other estimators. In this article, we proposed a new estimator that can handle jointly the violations of three of these assumptions which include multicollinearity, outliers, and an auto-correlated error term. Three estimators were combined to form Generalized-Ridge-Lad estimator. We compared the performance of the new estimator with some of the existing estimators in terms of their mean square error. The proposed estimator (GLSRIDGELAD) perform consistently better than other estimators when the three problems exist.

Keywords: Multicollinearity, Autocorrelated error, Outliers, Generalized ridge, LAD-estimator

1. Introduction

The parameters of the regression model are most popularly estimated using Ordinary Least Squares (OLS) estimator. The OLS estimates are far from their actual values for instance in a model where the observations occur at successive time points [1]. OLS estimator produced inaccurate estimates due to the following violations: multicollinearity, autocorrelated error terms, outliers or their combined presence. Different authors have focused on the individual effects of each of these problems. Some authors have considered the combined impact of these problems. For instance, attention has been paid to the combined effect of multicollinearity and autocorrelated error terms. These include [2-7]. Both [2] and [3] derived a ridge estimator based on GLS and show that it performs well. [3] obtained an expression for the mean squared error (MSE) of ridge estimators in the presence of AR (1) errors. This expression shows that collinearity and autocorrelation interact with inflating the MSE of OLS and ridge estimators. [1] provided generalized least squares (GLS) based adaptive ridge estimators for regression problems in which the independent variables are collinear rs, and the errors are autocorrelated and compared with the conventional OLS estimator. The GLS based methods are best especially when the independent variables are related and also serially correlated. [7] proposed the

Two Stages Ridge Estimator for the multiple linear models that suffer from both problems of autocorrelation AR (1) and multicollinearity.

Also, the following authors have considered the combined effect of both problems of multicollinearity and outlier. These include [8-17]. However, it is surprising that the effect of these three problems has drawn little attention. [15] and a few other authors recognized that the issue of collinearity might come associated with that of autocorrelation and outlier. Several authors as stated earlier have considered the use of Ridge regression within the context of the Generalized Least Squares to handle both collinearity and autocorrelation. The aim of this article is to proposed generalized least squares (GLS) based adaptive ridge estimators based on least absolute deviation (LAD) estimator for regression problems in which the independent variables are collinear, the errors are auto-correlated, and there is an outlier in the y-axis.

2. Existing and proposed Estimators

The Linear Regression Model is defined as:

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

where, y is the n×1 vector of response variable, X is n×p matrix of explanatory variables. β is the p×1 vector of regression parameters and ε represents the error terms with zero mean and constant variance σ^2. Assume that the response variable y and the explanatory variables in Equation (1) are standardized. Let Λ and T be the matrices of eigen-values and eigen-vectors of $X'X$ respectively such that $T'X'XT = \Lambda$ = diagonal ($\lambda_1, \lambda_2, \ldots \lambda_p$), where λ_i denotes i-th eigen-values of $X'X$ and $T'T = TT'$ $= I_p$. Equation (1) can be written in canonical form as

$$y = Z\alpha + \varepsilon \tag{2}$$

where, $Z = XT$ such that $Z'Z = \Lambda$ and $\alpha = T'\beta$.

The OLS estimator of α is :

$$\hat{\alpha}_i = (Z'Z)^{-1} Z'Y = \Lambda^{-1} Z'Y \tag{3}$$

The mean square error of OLS is defined as

$$MSE(\hat{\alpha}_{OLS}) = \sigma^2 \sum_{i=1}^{p} \lambda_i^{-1} \tag{4}$$

To overcome multicollinearity, [18] added ridge parameter (*k*) to the diagonal elements of the correlation matrix $Z'Z$ in equation (3) to form ordinary ridge estimator. This is given as:

$$\hat{\alpha}_{ORE} = (I - k_{OLS}(\Lambda + k_{OLS})^{-1})\hat{\alpha}_i \tag{5}$$

where $\hat{k}_{OLS} = \frac{p\hat{\sigma}^2_{OLS}}{\sum_{i=1}^{p} \hat{\alpha}_{i,OLS}{}^2}$ as suggested by [19] such that $k \geq 0$. The mean square error of ordinary ridge estimator is

$$MSE(\hat{\alpha}_{ORE}) = \sigma^2_{OLS} \sum_{i=1}^{p} \lambda_i (\lambda_i + k_{OLS})^{-2} + \sum_{i=1}^{p} \frac{(k_{OLS}\alpha_i)^2}{(\lambda_i + k_{OLS})^2} \tag{6}$$

Ridge regression and OLS estimators are sensitive to outliers. [20] introduced LAD estimator as alternative method to OLS when there is outlier in the y-direction. This technique minimize the sum of the absolute deviations of the response variable form its fitted values i.e., $\min_{\beta} \sum |y_i - m(x_i, \beta)|$. This estimator is statistically more efficient than the least squares estimator and ridge estimator when the error terms come from heavy-tailed distributions such as non-normal stable distribution, the Laplace distribution or contaminated normal distributions or the outlier is in the y-direction. The MSE of LAD estimator is defined as:

$$MSE(\hat{\alpha}_{LAD}) = \sum_{i=1}^{p} \Omega_{ii,LAD} \tag{7}$$

[21] suggested ridge estimator based on LAD (RIDGELAD) as an alternative method to the Ordinary Ridge Estimator defined in equation (5). It is defined as follows:

$$\hat{\alpha}_{\mathrm{LRE}} = (I - k_{LAD}(\Lambda + k_{LAD})^{-1})\hat{\alpha}_{LAD} \tag{8}$$

$$\hat{k}_{LAD} = \frac{p\hat{\sigma}^2_{LAD}}{\sum_{i=1}^{p} \hat{\alpha}_{i,LAD}{}^2}$$

where p denotes the number of estimated parameters. The modified MSE formula for the Lad ridge estimator suggested by [21] is presented in Equation (9)

$$MSE(\hat{\alpha}_{LRE}) = \sum_{i=1}^{p} \lambda_i^2 (\lambda_i + k_{LAD})^{-2} \Omega_{ii} + \sum_{i=1}^{p} \frac{(k_{LAD}\alpha_i)}{(\lambda_i + k_{LAD})^2}^2 \tag{9}$$

k_{LAD} is the robust version of the k parameter and $\Omega = Cov(\hat{\alpha}_{LRE})$.

The error term in equation (1) and (2) are assumed to follow AR (1) scheme, namely,

$$\varepsilon_t = \rho\varepsilon_{t-1} + \mathrm{u_t} \tag{10}$$

where $|\rho| < 1$, $\mathrm{E(u_t)} = 0$, $\mathrm{E(u_t^2)} = \sigma_u^2$ and $\mathrm{E(u_t u_s)}=0$, t≠s. Consequently, $\varepsilon \sim N(0,\sigma^2 V)$ where V is a positive definite and symmetric matrix given by

$$V=\begin{bmatrix} 1 & \rho & \rho^2 & \dots & \rho^{n-1} \\ \rho & 1 & \rho & \cdots & \rho^{n-2} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ \rho^{n-1} & \rho^{n-2} & \rho^{n-3} & \dots & 1 \end{bmatrix}$$

Consequently, there exists an invertible matrix Q such that $V^{-1} = Q'Q$ where Q is the n×n matrix defined as:

$$Q=\begin{bmatrix} \sqrt{1-\rho^2} & 0 & 0 & \dots & 0 \\ -\rho & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix}$$

Pre-multiplying model (1) by Q, the transformed model is obtained as

IOP Conf. Series: Journal of Physics: Conf. Series **1299** (2019) 012128 doi:10.1088/1742-6596/1299/1/012128

$$y^* = X^*\beta + \varepsilon^* \tag{11}$$

where $y^* = \begin{bmatrix} \sqrt{1-\rho^2}\, y_1 \\ y_2 - \rho y_1 \\ y_3 - \rho y_2 \\ \vdots \\ y_n - \rho y_{n-1} \end{bmatrix}$ $X^* = \begin{bmatrix} \sqrt{1-\rho^2} & \sqrt{1-\rho^2}\, x_{11} & \sqrt{1-\rho^2}\, x_{12} & \dots & \sqrt{1-\rho^2}\, x_{1k} \\ 1-\rho & x_{21} - \rho x_{11} & x_{22} - \rho x_{12} & \cdots & x_{2k} - \rho x_{1k} \\ \vdots & \vdots & \vdots \; \dots & & \vdots \\ 1-\rho & x_{n1} - \rho x_{(n-1)1} & x_{n2} - \rho x_{(n-1)2} & \dots & x_{nk} - \rho x_{(n-1)k} \end{bmatrix}$

such that $E(\varepsilon^*) = 0$ and $E(\varepsilon^* \varepsilon^{*'}) = \sigma_\varepsilon^2 I_n$. The generalized least squares estimator of β is given as:

$$\hat{\beta}_{GLS} = \left(X^{'} V^{-1} X\right)^{-1} X^{'} V^{-1} y \tag{12}$$

[22] suggested estimating ρ in equation (10) as follows:

$$\hat{\rho} = \frac{\sum_{t=2}^{n} \hat{\varepsilon}_t \hat{\varepsilon}_{t-1}}{\sum_{t=1}^{n} \hat{\varepsilon}_t^2} \tag{13}$$

where $\hat{\varepsilon}_t = y_t - \hat{\beta}_1 X_1 - \hat{\beta}_2 X_2 - \dots - \hat{\beta}_p X_p$ are the OLS residuals. Model (11) is re-written in canonical form as:

$$y^* = Z^* \alpha + \varepsilon^* \tag{14}$$

where $Z = X^* T$ and $\hat{\alpha}_{GLS} = T' \hat{\beta}_{GLS}$

The GLS estimator of α is defined as:

$$\hat{\alpha}_{GLS} = \Lambda^{-1} Z' y^* = T' \hat{\beta}_{GLS} \tag{15}$$

where $Z = X^* T$. The mean square error is defined as follows:

$$MSE(\hat{\alpha}_{GLS}) = \sum_{i=1}^{p} \Omega_{ii,GLS} \tag{16}$$

GLS estimator produces large variance when the independent variables are collinear. To reduce multicollinearity, ridge regression estimator is applied to the transformed model in Equation (14). The generalized ridge estimator based on OLS (GLSRIDGE) of α is defined as:

$$\hat{\alpha}_{GLS}(k_1) = (I + k_1 \Lambda^{-1})^{-1} \hat{\alpha}_{GLS} \tag{17}$$

where $k_1 = \dfrac{p\hat{\sigma}^2_{GLS}}{\sum_{i=1}^{p}\hat{\alpha}^2_{i,GLS}}$

The mean square error is defined as follows:

$$MSE(\hat{\alpha}_{GLSRIDGE}) = \sum_{i=1}^{p}\lambda_i^2(\lambda_i + k_1)^{-2}\Omega_{ii,GLS} + \sum_{i=1}^{p}\frac{(k_1\alpha_{i,GLS})^2}{(\lambda_i + k_1)^2} \tag{18}$$

where $\Omega = Cov(\hat{\alpha}_{GLS})$

When LAD estimator is applied to model (14) as alternative to OLS, the resulting estimator is GLS estimator based on LAD (GLSLAD) of α is defined as:

$$\hat{\alpha}_{GLSLAD} = T\hat{\beta}_{GLSLAD} \tag{19}$$

The MSE is defined as:

$$MSE(\hat{\alpha}_{GLSLAD}) = \sum_{i=1}^{p}\Omega_{ii,GLSLAD} \tag{20}$$

The proposed estimator which is referred to as Generalized Ridge estimator based on LAD (GLSRIDGELAD) of α is defined as:

$$\hat{\alpha}_{GLSRIDGELAD}(k_2) = (I + k_2\Lambda^{-1})^{-1}\hat{\alpha}_{GLSLAD} \tag{21}$$

where $k_2 = \dfrac{p\hat{\sigma}^2_{GLSLAD}}{\sum_{i=1}^{p}\hat{\alpha}^2_{i,GLSLAD}}$

$$MSE(\hat{\alpha}_{GLSRIDGELAD}) = \sum_{i=1}^{p}\lambda_i^2(\lambda_i + k_2)^{-2}\Omega_{ii,GLS} + \sum_{i=1}^{p}\frac{(k_1\alpha_{i,GLS})^2}{(\lambda_i + k_2)^2} \tag{22}$$

3. Simulation study and Numerical Example

Following [1], we carried out a simulation study to compare the performance of the estimators when the error terms and the explanatory variables are correlated coupled with the presence of outliers in the y-direction. Different specifications of the regression model are considered by varying the following factors: n, ρ, γ and σ. n is the sample sizes selected arbitrarily to be 20, 50 and 100, ρ is the level of correlation between any two error terms selected arbitrarily as 0.9 and 0.999 respectively, γ is the level of multicollinearity selected arbitrarily as 0.9, 0.99 and 0.999 respectively. p is taken to be three (3) for the experiment. The Q matrix used in this study were obtained from the selected values of ρ= 0.9 and 0.999. Following [13] and others, to obtain different levels of collinearity the following equation was used to generate the explanatory variables:

$$X_{ti} = (1-\gamma^2)^{\frac{1}{2}}Z_{ti} + \gamma Z_{tp} \quad t=1,\ldots,n.\ i=1,2,3 \tag{23}$$

where $Z_{ti} \sim N(0,1)$.

The dependent variable is obtained by the following equation:

$$y_t = \beta_1 X_{t1} + \beta_2 X_{t2} + \beta_3 X_{t3} + \varepsilon_t \tag{24}$$

where $\varepsilon_t = \rho\varepsilon_{t-1} + u_t$ such that $\varepsilon_t \sim (0, \sigma^2\Omega)$ and $u_t \sim N(0,1)$. β_0 is chosen to be zero, β_1=0.8, β_2=0.1, β_3=0.6 and σ are 1, 5 and 10.

Outliers in the y-direction was introduced into the model by considering two cases. These include a case of one outlier and five outliers. We changed the tenth observation to $y_{10}^* = \mathrm{Max}(y_t) + 20\sigma$ when there is one outlier. For five outlier case, the second, fifth, tenth, twelfth and fifteenth observation are altered as follows: $y_2^* = y_2 - 20\sigma$, $y_5^* = y_5 + 10\sigma$, y_{10}^* =Max(y_t) +20σ, $y_{12}^* = y_{12} + 50\sigma$ and $y_{15}^* = y_{15} - 50\sigma$.

Each model specification was replicated 5000 times. The different estimators were compared using the measures of average mean square error (AMSE). For any estimator $\hat{\beta}$, AMSE is defined as follows:

$$AMSE(\hat{\beta}) = \frac{1}{5000}\sum_{i=1}^{p}\sum_{j=1}^{5000}\left(\hat{\beta}_{ij} - \beta_i\right)^2 \tag{25}$$

where $\hat{\beta}_{ij}$ is i[th] estimated regression coefficient replicated j times . β_i are the true regression coefficients. An estimator is considered best when it possess the smallest AMSE.

The simulated results for the different estimators are presented in Tables 2 to 4. From these Tables, it was observed that as sample size increase the MSE decrease for all estimators. As the degree of multicollinearity (γ) and level of correlation among error terms (ρ) increases, MSE estimates increases. All estimators are sensitive to number of observations (n), correlation between regressors (γ) and correlation between errors (ρ). RIDGE and GLSRIDGELAD estimators perform consistently better than other estimators when there is multicollinaerity and autocorrelated error only. GLSRIDGELAD, RIDGELAD and GLSLAD perform consistently better than other estimators when there is multicollinearity, autocorrelated error and outlier(s) in the y-direction. However, the GLSRIDGELAD estimator does reasonably well in most cases.

The simulated results are illustrated by considering the dataset discussed by [7]. It comprises three explanatory variables. The matrix $X'X$ has eigenvalues 63942300, 85405.87595 and 8815.62635. The condition number is 7253.29108 and it is very large. This implies the data suffers from multicollinearity. The Durbin-Watson statistics, $d = \frac{(\sum_{t=2}^{n}\varepsilon_t - \varepsilon_{t-1})^2}{\sum_{t=1}^{n}\varepsilon_t^2}$ is used to detect the presence of autocorrelated error. The Durbin-Watson value obtained from the data set is 0.832613 with p-value (p-value = 0.000100437). This result indicated the presence of autocorrelated error. Outlier in the y-direction is often detected using studentized residual. Observation whose absolute value of its studentized residual value is greater than two (2) indicates presence of outlier. Using the data set, the value of studentized residual of case number 30 and 31 is given as 2.11610 and 2.38107 respectively. This implies there are two outlying values in the y-direction. It was observed from the diagnostic checks that the dataset considered in this study suffers the following problems. Multicollinearity, outlier and autocorrelated error. The results for all the estimators and their mean square error are provided in Table 1.

Table 1: Regression Coefficients and MSE values of the Proposed and Existing estimators

Coef.	Estimators							
	OLS	**LAD**	**RIDGE**	**RIDGELAD**	**GLS**	**GLSLAD**	**GLSRIDGE**	**GLSRIDGELAD**
$\hat{\boldsymbol{\beta}}_1$	0.2079	0.5451	0.2076	0.2074	0.2078	0.2905	0.2062	0.2067
$\hat{\boldsymbol{\beta}}_2$	0.9205	0.6315	0.9201	0.9197	0.9081	0.8808	0.8957	0.89651
$\hat{\boldsymbol{\beta}}_3$	-0.1340	-0.1242	-0.1345	-0.1349	-0.1615	-0.1137	-0.1525	-0.15247
$\hat{\boldsymbol{k}}$	0.000	0.000	0.0367	0.0666	0.000	0.000	0.0269	0.0393
MSE	0.1120	0.1336	0.0887	0.1273	0.0734	**0.0326**[2]	**0.0594**[3]	**0.03023**[1]

In term of minimum MSE, the first three preferred estimators are GLSRIDGELAD, GLSLAD and GLSRIDGE. The result agrees with the simulated results. However, the best estimator is GLSRIDGELAD when the three problems are present.

Table 2: Estimated AMSE values of the OLSE, LAD, RIDGE, GLS, GLSLAD, GLSRIDGE and GLSRIDGELAD when there is no outlier

N	σ	ρ	γ	OLS	LAD	RIDGE	RIDGELAD	GLS	GLSLAD	GLSRIDGE	GLSRIDGELAD
20	1	0.8	0.9	0.00051[3]	0.00389	0.00010[1]	0.00179	0.00187	0.00080	0.00115	0.00030[2]
			0.99	0.00330[3]	0.03105	0.00019[1]	0.01339	0.01962	0.00727	0.01197	0.00241[2]
			0.999	0.02752[3]	0.28948	0.00058[1]	0.12419	0.20359	0.07276	0.12451	0.02478[2]
		0.9	0.9	0.00099[3]	0.01035	0.00020[1]	0.00390	0.00309	0.00138	0.00181	0.00046[2]
			0.99	0.00685[3]	0.08380	0.00049[1]	0.03018	0.03372	0.01337	0.01990	0.00438[2]
			0.999	0.06439[3]	0.78222	0.00226[1]	0.28111	0.35401	0.13538	0.20932	0.04539[1]
50	1	0.8	0.9	0.00069	0.00178	0.00019	0.00087	0.00024	0.00014[3]	0.00013[2]	0.00005[1]
			0.99	0.00700	0.02372	0.00170	0.01204	0.00230	0.00110[2]	0.00124[3]	0.00034[1]
			0.999	0.07250	0.26290	0.01753	0.13494[3]	0.02529	0.01201[2]	0.01362	0.00384[1]
		0.9	0.9	0.00079	0.00314	0.00007[1]	0.00060	0.00038	0.00020	0.00019[3]	0.00005[1]
			0.99	0.00920	0.04493	0.00028[1]	0.01099[3]	0.00401	0.00190	0.00199	0.00049[2]
			0.999	0.10259	0.51314	0.00382[1]	0.13378	0.04326	0.020713[3]	0.02151	0.00563[2]
100	1	0.8	0.9	0.00023	0.00011	0.00008	0.00002[1]	0.00016	0.00013	0.00006[2]	0.00006[2]
			0.99	0.00289	0.00132	0.00056[3]	0.00031[1]	0.00183	0.00183	0.00042[2]	0.00078
			0.999	0.03175	0.01503	0.00574[3]	0.00391[1]	0.01715	0.02102	0.00492[2]	0.00934
		0.9	0.9	0.00039	0.00048	0.00007[3]	0.00009	0.00020	0.00017	0.00005[1]	0.00006[2]
			0.99	0.00464	0.00617	0.00052[2]	0.00133	0.00243	0.00260	0.00053[1]	0.00101[3]
			0.999	0.05054	0.06775	0.00561[1]	0.01615	0.02680	0.03025	0.00658[2]	0.01225[3]
20	5	0.8	0.9	0.01298[3]	0.08232	0.00151[1]	0.03834	0.04136	0.01681	0.02571	0.00589[2]
			0.99	0.08263[3]	0.73375	0.00314[1]	0.32512	0.47141	0.17521	0.29183	0.06134[2]
			0.999	0.68809[3]	7.10794	0.01156[1]	3.08484	5.02479	1.80182	3.09073	0.62930[2]
		0.9	0.9	0.02477[3]	0.23555	0.00333[1]	0.09083	0.07125	0.03116	0.04286	0.01076[2]
			0.99	0.17117[3]	2.02769	0.00915[1]	0.74638	0.81931	0.32600	0.49025	0.11207[2]
			0.999	1.60982[3]	19.34939	0.04906[1]	7.01495	8.76671	3.35929	5.20893	1.14839[2]
50	5	0.8	0.9	0.01737	0.04945	0.00437	0.02495	0.00632	0.00276[2]	0.00330[3]	0.00081[1]
			0.99	0.17493	0.60023	0.04207	0.30977	0.06498	0.03032[2]	0.03478[3]	0.00978[1]
			0.999	1.81254	6.58177	0.43820	3.39829	0.66310	0.31569[2]	0.35526[3]	0.10283[1]
		0.9	0.9	0.01997	0.08440	0.00152[2]	0.01868	0.01096	0.00476[3]	0.00524	0.00117[1]
			0.99	0.23019	1.12307	0.00928[1]	0.28583	0.11061	0.05177[3]	0.05470	0.01423[2]
			0.999	2.56464	12.79975	0.10404[1]	3.37575	1.12077	0.53675	0.55598[3]	0.14923[2]
100	5	0.8	0.9	0.00595	0.00282	0.00117[3]	0.00070[1]	0.00279	0.00376	0.00072[2]	0.00171
			0.99	0.07230	0.03531	0.01332[3]	0.00960[1]	0.03775	0.04824	0.01110[2]	0.02218
			0.999	0.79379	0.38573	0.14339[3]	0.10625[1]	0.42374	0.53470	0.12915[2]	0.24518
		0.9	0.9	0.00980	0.01204	0.00117[2]	0.00262	0.00437	0.00561	0.00093[1]	0.00233[3]
			0.99	0.11619	0.15240	0.01332[1]	0.03700	0.05903	0.06987	0.01476[2]	0.02926[3]
			0.999	1.26352	1.68470	0.14312[1]	0.41847	0.66443	0.77131	0.17315[2]	0.32126[3]
20	10	0.8	0.9	0.05193[3]	0.32306	0.00583[1]	0.15154	0.16398	0.06700	0.10279	0.02398[2]
			0.99	0.33056[3]	2.91525	0.01210[1]	1.29698	1.87753	0.69908	1.16528	0.24716[2]
			0.999	2.75238[3]	28.36878	0.04502[1]	12.3307	20.06825	7.20020	12.35356	2.52330[2]
		0.9	0.9	0.09906[3]	0.93180	0.01300[1]	0.36117	0.28331	0.12432	0.17169	0.04373[2]
			0.99	0.68470[3]	8.07852	0.03540[1]	2.98260	3.26690	1.30132	1.95902	0.45072[2]
			0.999	6.43930[3]	77.29604	0.19274[1]	28.05456	35.02659	13.42617	20.82540	4.60165[2]
50	10	0.8	0.9	0.06947	0.20157	0.01761[3]	0.10252	0.02675	0.01187	0.01385[2]	0.00353[1]
			0.99	0.69972	2.40591	0.16851	1.24473	0.26504	0.12414[2]	0.14151[3]	0.04026[1]
			0.999	7.25017	26.33317	1.75321	13.60691	2.66925	1.27200[2]	1.42888[3]	0.41508[1]
		0.9	0.9	0.07989	0.34184	0.00619[2]	0.07751	0.04581	0.02023[3]	0.02180	0.00507[1]
			0.99	0.92077	4.49356	0.03843[1]	1.14979	0.44893	0.21062[3]	0.22174	0.05833[2]
			0.999	10.25856	51.18608	0.42064[1]	13.51971	4.50417	2.15797[3]	2.23350	0.60155[2]
100	10	0.8	0.9	0.02379	0.01229	0.00457[3]	0.00324[2]	0.01160	0.01622	0.00321[1]	0.00756
			0.99	0.28920	0.14366	0.053263[3]	0.03985[1]	0.15125	0.19550	0.04563[2]	0.09088
			0.999	3.17517	1.54932	0.57386[3]	0.42973[1]	1.69384	2.14478	0.52073[2]	0.98750
		0.9	0.9	0.03921	0.04914	0.004675[2]	0.01128	0.01816	0.02414	0.00420[1]	0.01020[3]
			0.99	0.46475	0.60991	0.05366[1]	0.15018	0.23649	0.28317	0.06066[2]	0.11971[3]
			0.999	5.05409	6.73566	0.57419[1]	1.68212	3.09415	3.09415	0.69786[2]	1.29330[3]

*****Superscript 1, 2 and 3 are the rank of the first three preferred estimators**

Table 3: Estimated AMSE values of the OLSE, LAD, RIDGEOLS, GLSOLS, GLSLAD, GLSRIDGE and GLSRIDGELAD with one outlier

N	σ	ρ	γ	OLS	LAD	RIDGE	RIDGELAD	GLS	GLSLAD	GLSRIDGE	GLSRIDGELAD
20	1	0.8	0.9	0.01779	0.00304	0.00571	0.00015^{1}	0.01198	0.00057^{3}	0.00140	0.00018^{2}
			0.99	0.20923	0.03437	0.04792	0.00088^{2}	0.14204	0.00136^{3}	0.01568	0.00023^{1}
			0.999	2.41632	0.40647	0.54256	0.00890^{3}	1.80527	0.00596^{2}	0.21058	0.00135^{1}
		0.9	0.9	0.07523	0.00931	0.02397	0.00176^{2}	0.04900	0.00074	0.00588^{3}	0.00019^{1}
			0.99	0.90460	0.06219	0.22744	0.02832^{3}	0.66378	0.00183^{2}	0.08666	0.00022^{1}
			0.999	10.40151	0.54953	2.59589	0.32747^{3}	8.39948	0.00836^{2}	1.16777	0.00095^{1}
50	1	0.8	0.9	0.00785	0.00223	0.00181	0.00006^{1}	0.01043	0.00015^{3}	0.00223	0.00006^{1}
			0.99	0.08817	0.03012	0.02045	0.00183^{3}	0.11960	0.00116^{2}	0.02789	0.00016^{1}
			0.999	0.94702	0.33564	0.22193	0.02433^{3}	1.26769	0.01269^{2}	0.30172	0.00276^{1}
		0.9	0.9	0.02627	0.00314	0.00431	0.00005^{1}	0.04190	0.00020^{3}	0.00745	0.00012^{2}
			0.99	0.30954	0.04493	0.05194	0.00042^{2}	0.48057	0.00190^{3}	0.09260	0.00008^{1}
			0.999	3.38719	0.51314	0.58018	0.00654^{2}	5.12660	0.02071^{3}	1.00601	0.00337^{1}
100	1	0.8	0.9	0.00072	0.00009	0.00009	0.00002^{2}	0.00046	0.00014	0.00004^{3}	0.00001^{1}
			0.99	0.00735	0.00088	0.00044	0.00024^{3}	0.00378	0.00184	0.00010^{1}	0.00011^{2}
			0.999	0.07514	0.01012	0.00417	0.00277^{3}	0.03642	0.02103	0.00101^{1}	0.00146^{2}
		0.9	0.9	0.00144	0.00049	0.00011	0.00001^{1}	0.00081	0.00024	0.00006^{3}	0.00002^{2}
			0.99	0.01415	0.00618	0.00058	0.00006^{1}	0.00690	0.00353	0.00019^{2}	0.00034^{3}
			0.999	0.14255	0.06776	0.00549	0.00136^{1}	0.06677	0.04040	0.00190^{2}	0.00433^{3}
20	5	0.8	0.9	0.45369	0.06312	0.13920	0.00099^{2}	0.27009	0.00569^{3}	0.02445	0.00023^{1}
			0.99	5.32989	0.85243	1.21062	0.01615^{2}	3.60881	0.01954^{2}	0.37742	0.00075^{1}
			0.999	61.53380	10.18723	13.79974	0.20123^{3}	46.23112	0.11969^{2}	5.34414	0.02544^{1}
		0.9	0.9	1.27508	0.14494	0.44085	0.01840^{3}	0.79640	0.00733^{2}	0.08544	0.00021^{1}
			0.99	1.89298	0.19623	0.59243	0.03338^{3}	1.17111	0.00853^{2}	0.12671	0.00022^{1}
			0.999	261.91163	13.49084	65.32125	8.15395^{3}	212.01923	0.17423^{2}	29.34078	0.01417^{1}
50	5	0.8	0.9	0.18939	0.06137	0.04431	0.00346^{3}	0.26125	0.00290^{2}	0.06061	0.00004^{1}
			0.99	2.14399	0.75903	0.50208	0.05553^{3}	2.90994	0.03194^{2}	0.69823	0.00385^{1}
			0.999	23.11618	8.39295	5.45528	0.65809^{3}	30.83021	0.33267^{2}	7.42679	0.07045^{1}
		0.9	0.9	0.65140	0.08440	0.10937	0.00066^{2}	1.06306	0.00476^{3}	0.19938	0.00010^{1}
			0.99	7.66461	1.12307	1.29498	0.01379^{2}	11.9167	0.051766^{3}	2.33394	0.00204^{1}
			0.999	83.84903	12.79975	14.40057	0.18010^{2}	126.7893	0.53675^{3}	25.02799	0.08658^{1}
100	5	0.8	0.9	0.01789	0.00215	0.00163	0.00024^{3}	0.00848	0.00376	0.00023^{2}	0.00009^{1}
			0.99	0.18302	0.02525	0.01176	0.00641^{3}	0.08717	0.04824	0.00258^{1}	0.00347^{2}
			0.999	1.87351	0.26889	0.10950	0.07256^{3}	0.88988	0.53470	0.02841^{1}	0.04072^{2}
		0.9	0.9	0.03649	0.01204	0.00264	0.00012^{1}	0.01675	0.00788	0.00042^{3}	0.00026^{2}
			0.99	0.35818	0.15240	0.01626	0.00296^{1}	0.16512	0.09451	0.00478^{2}	0.00981^{3}
			0.999	3.60637	1.68470	0.14423	0.03967^{1}	1.66044	1.02965	0.05091^{2}	0.11348^{3}
20	10	0.8	0.9	1.82024	0.24732	0.55580	0.00337^{2}	1.06752	0.01975^{3}	0.09354	0.00029^{1}
			0.99	21.38449	3.40782	4.85269	0.06218^{2}	14.47896	0.07239^{3}	1.50466	0.00186^{1}
			0.999	246.86354	40.76323	55.35433	0.79426^{3}	185.63410	0.46567^{2}	21.43969	0.09863^{1}
		0.9	0.9	7.59386	0.76796	2.37141	0.12893^{3}	4.67106	0.03045^{2}	0.49924	0.00024^{1}
			0.99	91.28668	5.81949	22.87442	2.76463^{3}	66.99277	0.10855^{2}	8.53103	0.00085^{1}
			0.999	1049.69749	53.84119	261.76171	32.59891^{3}	850.2329	0.68109^{2}	117.62064	0.05301^{1}
50	10	0.8	0.9	0.75615	0.24953	0.17755	0.01531^{3}	1.04970	0.01247^{2}	0.24615	0.00006^{1}
			0.99	8.55666	3.04058	2.00600	0.22723^{3}	11.61850	0.13065^{2}	2.79762	0.01592^{1}
			0.999	92.24939	33.57427	21.78599	2.65554^{3}	122.9938	1.33991^{2}	29.67026	0.28306^{1}
		0.9	0.9	2.60295	0.34184	0.43868	0.00296^{3}	4.26119	0.02023^{2}	0.80483	0.00006^{1}
			0.99	30.62154	4.49356	5.17857	0.05735^{1}	47.6195	0.21062^{3}	9.34596	0.00895^{1}
			0.999	334.98191	51.18608	57.55071	0.72929^{1}	506.47307	2.15797^{3}	100.05161	0.34814^{1}
100	10	0.8	0.9	0.07185	0.00958	0.00655	0.00112^{3}	0.03364	0.01623	0.00091^{2}	0.00056^{1}
			0.99	0.73398	0.10404	0.04759	0.02624^{3}	0.34740	0.19550	0.01084^{1}	0.01460^{2}
			0.999	7.50998	1.08494	0.43990	0.29214^{3}	3.55870	2.14478	0.11577^{1}	0.16505^{2}
		0.9	0.9	0.14641	0.04914	0.01077	0.00067^{1}	0.06648	0.03350	0.00155^{2}	0.00140^{1}
			0.99	1.43672	0.60991	0.06609	0.01285^{1}	0.65879	0.38265	0.01952^{2}	0.04023^{3}
			0.999	14.46334	6.73566	0.58038	0.16135^{1}	6.64579	4.12986	0.20610^{2}	0.45674^{3}

*****Superscript 1, 2 and 3 are the rank of the first three preferred estimators**

Table 4: Estimated AMSE values of the OLSE, LAD, RIDGEOLS, GLSOLS, GLSLAD, GLSRIDGE and GLSRIDGELAD when there are five outliers

N	σ	ρ	γ	OLS	LAD	RIDGE	RIDGELAD	GLS	GLSLAD	GLSRIDGE	GLSRIDGELAD
20	1	0.8	0.9	0.63436	0.00075^{3}	0.15284	0.00016^{1}	0.05093	0.00206	0.00306	0.00021^{2}
			0.99	6.80071	0.00407^{3}	1.53154	0.00005^{1}	0.32023	0.01749	0.06749	0.00092^{2}
			0.999	70.92410	0.04399^{3}	15.44985	0.012815^{1}	2.64012	0.16838	0.81270	0.031846^{2}
		0.9	0.9	2.39297	0.00269	0.59134	0.00014^{1}	0.17638	0.00156^{3}	0.00810	0.000203^{2}
			0.99	25.63094	0.01986	5.91383	0.00038^{2}	1.00865	0.01170^{3}	0.18546	0.00026^{1}
			0.999	267.20028	0.22302	59.61078	0.08209^{2}	7.91794	0.10744^{3}	2.33382	0.00391^{1}
50	1	0.8	0.9	0.08130	0.00147	0.01477	0.00006^{1}	0.02026	0.00011^{2}	0.00079	0.00018^{3}
			0.99	0.90276	0.01936	0.16376	0.00027^{2}	0.23318	0.00036^{3}	0.01479	0.00016^{1}
			0.999	9.62578	0.21344	1.74423	0.00453^{3}	2.48854	0.00336^{2}	0.16320	0.00005^{1}
		0.9	0.9	0.34214	0.00314	0.05540	0.00016^{2}	0.08668	0.00011^{1}	0.00223	0.00020^{3}
			0.99	3.85984	0.04493	0.62737	0.00017^{1}	1.00225	0.00090^{3}	0.04153	0.00017^{1}
			0.999	41.41610	0.51314	6.73928	0.00112^{2}	10.78817	0.01011^{3}	0.45985	0.000056^{1}
100	1	0.8	0.9	0.01346	0.00010	0.00214	0.00005^{1}	0.00127	0.00014	0.00005^{1}	0.00009^{3}
			0.99	0.15292	0.00104^{3}	0.02321	0.00022^{2}	0.01787	0.00184	0.00204	0.00006^{1}
			0.999	1.63370	0.01221^{3}	0.24252	0.00415^{2}	0.21497	0.02103	0.02438	0.00125^{1}
		0.9	0.9	0.03715	0.00038	0.00611	0.00005^{1}	0.00360	0.00024	0.00010^{2}	0.00011^{3}
			0.99	0.42425	0.00483	0.06680	0.00060^{2}	0.04448	0.00353^{3}	0.00432	0.00011^{1}
			0.999	4.54058	0.05310	0.69839	0.01021^{2}	0.53056	0.04040^{3}	0.05135	0.00371^{1}
20	5	0.8	0.9	15.87503	0.01587^{3}	3.81752	0.00012^{1}	1.34387	0.04025	0.08829	0.00033^{2}
			0.99	170.31007	0.11158^{3}	38.31054	0.00146^{1}	8.26255	0.40372	1.71741	0.01227^{2}
			0.999	1776.65129	1.15642^{3}	386.6941	0.32807^{1}	67.44524	4.10437	20.43272	0.77086^{2}
		0.9	0.9	59.84514	0.06575	14.78027	0.00006^{1}	4.54082	0.02948^{3}	0.22423	0.00021^{2}
			0.99	641.20916	0.52569	147.8824	0.01438^{2}	25.64322	0.26752^{3}	4.69176	0.00072^{1}
			0.999	6685.60663	5.71488	1491.025	2.07199^{2}	200.1808	2.61232^{3}	58.53612	0.076948^{1}
50	5	0.8	0.9	2.03052	0.04125	0.36987	0.00002^{1}	0.51694	0.00061^{3}	0.02507	0.00019^{2}
			0.99	22.58658	0.49188	4.10418	0.00796^{3}	5.74456	0.00791^{2}	0.38022	0.00015^{1}
			0.999	240.93000	5.35069	43.67906	0.11740^{3}	61.0968	0.08658^{2}	4.07310	0.00069^{1}
		0.9	0.9	8.55261	0.08440	1.38981	0.00020^{2}	2.20715	0.00204^{3}	0.06734	0.00019^{1}
			0.99	96.52932	1.12307	15.71046	0.00174^{2}	24.96075	0.02469^{3}	1.06302	0.00014^{1}
			0.999	1035.97028	12.7997	168.6452	0.02676^{2}	267.8493	0.26475^{3}	11.51676	0.00194^{1}
100	5	0.8	0.9	0.33703	0.00274	0.05376	0.00003^{1}	0.04527	0.00375	0.00271^{3}	0.00008^{2}
			0.99	3.82554	0.03104^{3}	0.58281	0.00536^{2}	0.47326	0.04824	0.05597	0.00138^{1}
			0.999	40.85846	0.32666^{3}	6.07538	0.10570^{2}	5.42971	0.53470	0.62655	0.03233^{1}
		0.9	0.9	0.92743	0.01201	0.15236	0.00006^{1}	0.11278	0.00788	0.00526^{3}	0.00007^{2}
			0.99	10.58869	0.13051	1.66286	0.01625^{2}	1.15749	0.09454^{3}	0.11508	0.00329^{1}
			0.999	113.34444	1.36099	17.37630	0.25967^{2}	13.38389	1.02965^{3}	1.30123	0.09554^{1}
20	10	0.8	0.9	63.50987	0.06332^{3}	15.26841	0.00008^{1}	5.41255	0.15666	0.35987	0.00052^{2}
			0.99	681.43111	0.45267^{3}	0.007010^{1}	33.19448	1.59956	6.88668	6.88668	0.045000^{2}
			0.999	7108.8925	4.65558^{3}	1547.069	1.31704^{1}	270.6499	16.3664	81.80015	3.07141^{2}
		0.9	0.9	239.41824	0.26356	59.11463	0.00005^{1}	18.24009	0.11447^{3}	0.90933	0.00023^{2}
			0.99	2565.4025	2.11880	591.5828	0.06109^{2}	102.8926	1.05904^{3}	18.80589	0.00164^{1}
			0.999	26748.547	22.9308	5964.954	8.29841^{2}	802.6935	10.4140^{3}	234.3130	0.29849^{1}
50	10	0.8	0.9	8.12171	0.16856	1.48036	0.00021^{2}	2.07865	0.00268^{3}	0.10365	0.00020^{1}
			0.99	90.35214	1.97291	16.42180	0.03298^{3}	22.96090	0.03249^{2}	1.52761	0.00012^{1}
			0.999	963.83210	21.41164	174.74851	0.47214^{3}	243.9707	0.34914^{2}	16.29613	0.00341^{1}
		0.9	0.9	34.20998	0.34184	5.56209	0.00027^{2}	8.84983	0.00902^{3}	0.27568	0.00020^{1}
			0.99	386.13401	4.49356	62.85545	0.00635^{2}	99.79685	0.10125^{3}	4.26485	0.00011^{1}
			0.999	4144.16607	51.18608	674.66285	0.10683^{2}	1070.474	1.06644^{3}	46.07768	0.00913^{1}
100	10	0.8	0.9	1.34723	0.012172	0.21494	0.00005^{2}	0.18891	0.01623	0.01215^{3}	0.00004^{1}

IOP Conf. Series: Journal of Physics: Conf. Series **1299** (2019) 012128 doi:10.1088/1742-6596/1299/1/012128

			0.99	15.29611	0.12803^{3}	2.32918	0.02208^{2}	1.90877	0.19550	0.22652	0.00603^{1}
			0.999	163.38186	1.31886^{3}	24.27809	0.42427^{2}	21.76002	2.14478	2.51262	0.13037^{1}
		0.9	0.9	3.70819	0.05036	0.60924	0.00045^{2}	0.46354	0.03350	0.023060^{3}	0.00003^{1}
			0.99	42.33938	0.52832	6.64478	0.06628^{2}	4.65495	0.38265^{3}	0.46399	0.01416^{1}
			0.999	453.22718	5.46218	69.42943	1.04137^{2}	53.60931	4.12986^{3}	5.21157	0.38401^{1}

*****Superscript 1, 2 and 3 are the rank of the first three preferred estimators**

4. Conclusions

In this study, generalized ridge estimator based on LAD (GLSRIDGELAD) is introduced as alternative study to OLS, GLS, LAD, RIDGE, RIDGELAD, GLSLAD and GLSRIDGE estimators in order to handle the following problems in linear regression model: multicollinearity, autocorrelated error and outlier in the y-direction deal. The simulated results showed that GLSRIDGELAD is generally preferred to other estimators especially when the three problems are present. However, when there is problem of multicollinearity and autocorrelated error with no outlier RIDGEOLS is occasionally preferred to it. Furthermore, the result of the numerical example agrees with the simulated results.

References

[1] Firinguetti, L. L. (1989). A simulation study of ridge regression estimation with autocorrelated errors. Commun. Statist. B 18(2):673–702.

[2] Swamy, P.A.V.B. and Rappaport, P.Y. (1978). Relative Efficiencies of Some Simple Rayes Estimators of Coefficients in a Dynamic Equation With Serially Correlated Errors - II. Journal of Econometrics, 7, 245-258.

[3] Hsu, J.J., (1980). Multicollinearity, Autocorrelation and Ridge Regression. Unpublished M.,Sc. Thesis, Faculty of Commerce and Business Administration, University of British Columbia, Vancouver, British Columbia.

[4] Gosling, B. J. and Puterman, M. L (1985). Ridge estimation in regression problems with autocorrelated errors: a monte carlo study. Commun. Statist.-simula. Computa, 14(3), 577-613.

[5] Hussein, Y. and Zari, M. (2012). Generalized Two Stage Ridge Regression Estimator TR for Multicollinearity and Autocorrelated Errors. Canadian Journal on Science and Engineering Mathematics, 3(3), 79-85.

[6] Alkhamisi, M. A. (2010). Ridge Estimation in Linear Models with Autocorrelated Errors. Communications in Statistics—Theory and Methods, 39: 2630–2644.

[7] Eledum, H. and Zari, M. (2012). Relaxation Method for Two Stages Ridge Regression Estimator. International Journal of Pure and Applied Mathematics, 85(4), 653-667.

[8] Holland, P. W. (1973). Weighted ridge regression: Combining ridge and robust regression methods. NBER Working Paper Series, 11, 1-19.

[9] Askin, G. R., and Montgomery, D. C. (1980). Augmented robust estimators. Techonometrics, 22, 333-341.

[10] Midi, H. and Zahari, M. (2007). A Simulation Study on Ridge Regression Estimators in the Presence of Outliers and Multicollinearity. Jurnal Teknologi. 47(C): 59-74.

[11] Hatice, S. and Alpu, O. (2010). Ridge Regression Based on Some Robust Estimators. Journal of Modern Applied Statistical Methods: 9(2), 495-501.

[12] Maronna, R.A. (2011). Robust Ridge Regression for High-Dimensional Data. Technometrics, 53(1): 44-53.

[13] Lukman. A. F, Osowole, O.I. and Ayinde, K. (2015). Two Stage Robust Ridge Method in a Linear Regression Model. Journal of Modern Applied Statistical Methods, 14(2), 53-67.

[14] Ertas, H., Kaciranlar, S. and Guler, H. (2015). Robust Liu-type Estimator for Regression Based on M-estimator. Communications in Statistics - Simulation and Computation.

[15] Ertas, H., Toker, S. and Kacıranlar, S. (2015). Robust two parameter ridge Mestimator for linear regression, Journal of Applied Statistics.

[16] Norouzirad, M., Arashi, M. and Ahmed, S. E. (2017). Improved robust ridge M-estimation. Journal of statistical computation and simulation. DOI: 10.1080/00949655.2017.1371174

[17] Ertaş, H. (2017). A Modified Ridge M-Estimator for Linear Regression Model with Multicollinearity and Outliers. Accepted paper by Communications in Statistics - Simulation and Computation. DOI: 10.1080/03610918.2017.1310231.

[18] Hoerl, A.E. and Kennard, R.W. (1970). Ridge regression: biased estimation for non-orthogonal problems. Technometrics, 12, 55-67.

[19] Hoerl, A. E., Kennard, R. W. and Baldwin, K. F. (1975). Ridge regression: Some simulation. Communications in Statistics, 4 (2), 105–123.

[20] Dielman, T.E. (1984). Least absolute value estimation in regression models: An annotated bibliography. Communications in Statistics - Theory and Methods. 4, 51 3-541.

[21] Lukman, A. F., Arowolo, O.T and Ayinde, K. (2014). Some Robust Ridge Regression for Handling Multicollinearity and Outlier. International Journal of Sciences: Basic and Applied Research, 192-202.

[22] Prais, S. J. and Winsten, C. B. (1954). Trend Estimators and Serial Correlation. Cowless Commision Discussion Paper N0383, Chicago.