STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
Stat., Optim. Inf. Comput., Vol. **13**, May 2025, pp 1819–1831.
Published online in International Academic Press (www.IAPress.org)

Robust Lasso Estimator for the Liu-Type Regression Model and its Applications

Tarek Omara [1,2]

[1] *Department of Statistics, Mathematics and Insurance,Faculty of Commerce, Kafrelsheikh University,Egypt*
[2] *Department of Mathematics, Faculty of Sciences - Islamic University of Madinah, Madinah,KSA*

Abstract In this paper, we propose a new estimator for Liu-Type regression model, called the LAD-Lasso-Liu estimator, which addresses the issues of multicollinearity, outliers and it performs the variable selection. By combining the LAD Lasso and Liu-Type estimators, our proposed estimator achieves double shrinkage for the parameters and at the same time it has the robust properties. We thoroughly discuss the properties of the new estimator and conduct a simulation study to demonstrate its superiority over the LAD, S, MM, Liu-type, Lasso, and LAD-Lasso estimators. We used the Median(MSE) as a criteria to compare between the estimators at a different factors. The simulation results showed that the proposed estimator has superiority over the other estimators especially when the correlation coefficient between the explanatory variables increases and when the error variance decreases. In addition, the proposed estimator has better correct selection of the number of zeros coefficients than other penalized estimators. To demonstrate the work of the estimator presented under real data, we apply the proposed estimator to prostate cancer data. Our results for the empirical data indicate that the proposed estimator outperforms the other estimators and can provide accurate results in challenging scenarios with multicollinearity and outliers.

Keywords Multicollinearity, outliers, lasso estimator, Liu-type estimator, LAD estimator, LAD-Lasso-Liu estimator

AMS 2020 subject classifications 62J07

DOI: 10.19139/soic-2310-5070-2106

1. Introduction

The ordinary least squares (OLS) estimator is considered the best unbiased estimator for linear regression models, but it is susceptible to outliers [1]. Robust estimators, such as the Least Absolute Deviation (LAD) estimator, S-estimator, M-estimator and MM-estimator have been introduced to address this issue. The LAD estimator is a popular robust estimator that minimizes the sum of the absolute values of the residuals and is resistant to heavy-tailed errors that arise due to the presence of outliers. The LAD is a special case of M-estimator [2], and has been used in various studies to deal with outliers, e.g., [3, 4, 5, 6, 7]. Since the LAD estimator minimizes the sum of absolute residuals, large deviations do not have a disproportionately large effect (as opposed to the squared deviations in OLS). As a result, extreme values (vertical outliers) do not unduly influence the overall estimate. By focusing on the median rather than the mean, the LAD estimator ensures that the majority of the data drives the regression, rather than being overly sensitive to extreme values. While the LAD estimator is robust to vertical outliers, it is less effective at handling leverage points, which are outliers in the predictor space (i.e., extreme values in the independent variables). [8] Leverage points can still influence the LAD estimator significantly because the estimator depends on the positions of these points in the x-space.
However, LAD estimator cannot handle multicollinearity in the linear regression model. The shrinkage estimator

*Correspondence to: Tarek Omara (Email: drtarek.em @gmail.com)

ISSN 2310-5070 (online) ISSN 2311-004X (print)

is important methods which dependent to add more information to matrix X^TX to remove the ill condition which appears because of the presence of the multicollinearity. [9] introduced the Ridge (RE) estimator to deal with multicollinearity, which obtained by minimizing the sum of squared errors with a penalty on the L_2 norm of the coefficients. The RE estimator is biased and depends on one biased parameter known as Ridge parameter. The good selection for the Ridge parameter is important point in the way for overcome the problem of multicollinearity since if we don't select it large enough, it is leads to the persistence of the problem of multicollinearity [10]. In the same way, [11] introduce the Liu estimator which mixes between the RE estimator and Stein estimator. [12]introduced the Liu-Type estimator, which improves the RE estimator by addressing high multicollinearity. The Liu-Type estimator has two parameters work to gather to parlay to overcome the multicollinearity and at the same time improve the fitting for the estimators [13].
In the regression model, the most important problem is how to select the most significant predictors, because if we omitted the relevant predictors, it may lead to biased estimates and less accurate predictions [1]. To improve the interpretability of the model, the Linear Absolute Shrinkage and Selection Operator (Lasso) estimator was introduced by [14], which shrinks some coefficients and sets other coefficients equal exactly zero. The Lasso estimator is obtained by minimizing the sum of squared errors with a penalty on the L1 norm of the coefficients, which makes it susceptible to outliers. Many literatures illustrated new estimators by combine between two of the LAD and Lasso and Liu-Type estimators. In this way, The LAD-Lasso estimator, introduced by [15] which combines the LAD estimator and the Lasso estimator, providing the features of the Lasso estimator and LAD estimator simultaneously. The LAD-Lasso estimator has been used in several studies, e.g., [16, 33, 18, 19, 20]. To deal with the problem of multicollinearity and variable selection at the same time, [21] introduced the Liu estimator with a penalty on the L1 norm. In this study, we propose new biased estimator called LAD-Lasso-Liu Estimator. This estimator has the features for the three penalized estimators, LAD, Lasso and Liu-type estimators.
This study is organized as follows: section 1 provides an introduction. Section 2 reviews previous literature. Section 3 discusses the proposed estimator and its properties. We choose a tuning parameter in section 4 and present the simulation study in section 5. The application to real data is presented in section 6. In Section 7, future applications of the proposed estimator are suggested, and in Section 8, the limitations and assumptions are addressed. Section 9 outlines the conclusions.

2. The previous literature of the idea:

There is many literature that dealt with lasso, Liu, LAD estimators, which will be reviewed in this section.

2.1. The previous literature of the idea

Consider the following linear regression model

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

Where $Y = (y_1, y_2, \cdots, y_n)^T$ is is the responses variable, $X = (X_1, X_2, \cdots, X_n)^T$ is the represents design matrix of rank P where,$X_1^T = (X_{i_1}, X_{i_2}, \cdots, X_{i_p})$,$i = 1, 2, \cdots, n$, $\beta = (\beta_1, \beta_2, \cdots, \beta_p)^T$ is the represents the vector of unknown parameters and ε represents the error vector, which has median zero and at a same time its continuous and positive density function. The canonical form for the linear regression model is defined as

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

Where $Z = XQ$, $\alpha = Q^T\beta$ and $\Lambda = Z^TZ = diag(\lambda_1, \lambda_2, \cdots, \lambda_p)$ where $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p > 0$ are eigenvalues of X^TX and Q represents orthogonal matrix that has eigenvectors columns.

$$\widehat{\beta}_{OLS} = argmin_\beta \sum_{i=1}^{n}\left(y_i - \sum_{j=1}^{p} x_{ij}\beta_j\right)^2 = \left(X^TX\right)^{-1}X^TY \tag{3}$$

This estimator is affected by multicollinearity which causes ill condition in X^TX matrix. That means the elements of $\left(X^TX\right)^{-1}$ become large and the condition number $CI = \sqrt{\lambda_1/\lambda_P}$ become high [12]. The one of bad effects of the multicollinearity is the confidence intervals for the individual parameters may become wider and may the signs of produce parameter become wrong [22]. There are several techniques available to overcome the problem of ill-conditioned matrices. The Ridge estimator is a one effective method to remove of ill condition by adding more information to the matrix X^TX. The ridge estimator (RE) is given by

$$\widehat{\beta}_{RE} = argmin_\beta \sum_{i=1}^{n}\left(y_i - \sum_{j=1}^{p} x_{ij}\beta_j\right)^2 + k_1\sum_{j=1}^{p}\beta_j^2 = \left(X^TX + k_2 I\right)^{-1} X^TY \tag{4}$$

where $k_2 > 0$ is ridge parameter. This estimator is biased but at the same time it has mean squared errors less than OLS estimator. However, this estimator needs to increase the Ridge parameter k_1 sufficiently to overcome the multicollinearity. So, [12] introduced the Liu-Type estimator that has two shrinkage parameters, if one of them decreases it is compensated by the other. The Liu-Type estimator is given by

$$\widehat{\beta}_{Liu-Type} = argmin_\beta\left(\sum_{i=1}^{n}\left(y_i - \sum_{j=1}^{p} x_{ij}\beta_j\right)^2 + \sum_{j=1}^{p}\left(k_{1_j}^{1/2}\beta_j - \frac{d_j\widehat{\beta}_j}{k_{1_j}^{1/2}\beta_j}\right)\right) = \left(X^TX + k_1 I\right)^{-1}\left(X^TY - d\beta\right) \tag{5}$$

where $k_1 > 0$, $-\infty < d < \infty$ and $\widehat{\beta}$ is any estimator of β. [23] augmenting the equation $\left(-d/k_1^{1/2}\right)\widehat{\beta} = k_1 1/2\beta + \mu^T$ to linear regression model in $Eq.(1)$ to get the Liu-Type regression model

$$Y^* = X^*\alpha + \varepsilon^* \tag{6}$$

where $Y^*_{(n+p)\times 1} = \begin{bmatrix} Y \\ \left(-d/k_1^{1/2}\right)\widehat{\beta}_{OLS}\end{bmatrix}$, $X^*_{(n+p)\times p} = \begin{bmatrix} X \\ nk_1^{1/2} I\end{bmatrix}$, $\varepsilon^*_{(n+p)\times 1} = \begin{bmatrix}\varepsilon \\ \mu\end{bmatrix}$
On the other hand, to select the most significant predictors for linear regression model, [14] introduced Lasso estimator that has shrinkage some coefficients and variable selection simultaneously. The Lasso estimator is

$$\widehat{\beta}_{Lasso} = argmin_\beta\left(\sum_{i=1}^{n}\left(y_i - \sum_{j=1}^{p} x_{ij}\beta_j\right)^2 + n\sum_{j=1}^{p} k_{2_j}\left|\beta_j\right|\right) \tag{7}$$

where $k_2 > 0$ For dealt with the multicollinearity and selection variable problems simultaneously, [21] introduces the Liu–lasso estimator by using the L1 norm for the Liu estimator as

$$\widehat{\beta}_{Liu-Type-Lasso} = argmin_\beta\left(\sum_{i=1}^{n}\left(y_i - \sum_{j=1}^{p} x_{ij}\beta_j\right)^2 + \sum_{j=1}^{p}\left(k_{1_j}^{1/2}\beta_j - \frac{d_j\widehat{\beta}_j}{k_{1_j}^{1/2}\beta_j}\right) + n\sum_{j=1}^{p} k_{2_j}\left|\beta_j\right|\right) \tag{8}$$

2.2. *LAD and LAD- lasso estimator for linear regression model*

The RE, Lasso and Liu-Type estimators are shrinking the parameters to avoid the ill condition, but they are affected by outliers. It is known that outliers have a direct effect on the nature of errors, possibly making the variance of errors non-specific [15]. The robust estimators are the appropriate solution of the outliers, that replaces the sum squares error by the robust function that it data in the presence of outliers. The (LAD) estimator is robust estimator that has many features to overcome the outlier. The (LAD) estimator is given by

$$\widehat{\beta}_{LAD} = argmin_\beta \sum_{i=1}^{n}\left|y_i - \sum_{j=1}^{p} x_{ij}\beta_j\right| \tag{9}$$

For dealt with the outliers and selection variable problems simultaneously, [15] introduced LAD-Lasso estimator that minimizes the sum of the absolute values of error with penalty the L1 norm. The LAD-Lasso estimator is given by

$$\widehat{\beta}_{LAD-Lasso} = argmin_{\beta} \sum_{i=1}^{n} \left| y_i - \sum_{j=1}^{p} x_{ij}\beta_j \right| + n\sum_{j=1}^{p} k_{2_j} \left|\beta_j\right| \tag{10}$$

Define $(y_i^{\circ}, x_i^{\circ})$, $i = 1, 2, \cdots, n, n+1, \cdots, n+p$
where$(y_i^{\circ}, x_i^{\circ}) = \left\{ \begin{matrix} (y_i, x_i), i = 1, 2, \cdots, n \\ \left(0, n\hat{k}_2 e_i\right), j = 1, 2, \cdots, p \end{matrix} \right\}$, $e_j = \left\{ \begin{matrix} 1 & withj^{th}term \\ 0 & els \end{matrix} \right\}$.
Then, we can rewrite $Eq.(10)$ as

$$\widehat{\beta}_{LAD-Lasso} = argmin_{\beta} \sum_{i=1}^{n+p} \left| y_i^{\circ} - \sum_{j=1}^{p} x_{ij}^{\circ}\beta_j \right| \tag{11}$$

3. The proposed estimator and its properties:

For the Liu-Type regression model in$Eq.(6)$, we get (LAD-Lasso) estimator for Liu-Type regression by minimizing the sum the absolute value for error with penalty on the L1 norm. Since the LAD-Lasso estimator function includes both LAD (robust to outliers) and Lasso (inducing sparsity), making the estimator suitable for datasets with noisy observations and many irrelevant predictors. In addition, LAD-Lasso-Liu estimator is designed to balance the fit of the model (minimizing absolute deviations) with regularization (shrinking the coefficients toward zero). This balance is key to handling multicollinearity and improving prediction accuracy in high-dimensional settings. In fact, this estimator has features of LAD-Lasso estimator and Liu-Type estimator at the same time. We illustrate the LAD-Lasso estimator for Liu-Type regression as

$$\widehat{\beta}_{LAD-Lasso-Liu} = argmin_{\beta} \left(\sum_{i=1}^{n+p} \left| y_i^{*} - \sum_{j=1}^{p} x_{ij}^{*}\beta_j \right| + n\sum_{j=1}^{p} k_{2_j} \left|\beta_j\right| \right) \tag{12}$$

Define $(y_i^{\circ\circ}, x_i^{\circ\circ})$, $i = 1, 2, \cdots, n, n+1, \cdots, n+p, \cdots, n+2p$

$$\text{where } (y_i^{\circ\circ}, x_i^{\circ\circ}) = \left\{ \begin{matrix} (y_i, x_i) & for & i = 1, 2, \cdots, n \\ -d_i\hat{k}_{1_i}^{-1/2} e_{i-n}\hat{\beta}_i, \hat{k}_{1_i}^{1/2} I & for & i = n+1, n+2, \cdots, n+p \\ o, n\hat{k}_{2_i} e_i & for & i = n+p+1, n+p+2, \cdots, n+2p \end{matrix} \right\}$$
$$= \left\{ \begin{matrix} (y_i^{*}, x_i^{*}) & for & i = 1, 2, \cdots, n, n+1, \cdots, n+p \\ 0, n\hat{k}_{2_i} e_i & for & i = n+p+1, n+p+2, \cdots, n+2p \end{matrix} \right\}, e_j = \left\{ \begin{matrix} 1 & withj^{th}term \\ 0 & els \end{matrix} \right\}.$$

Then, we can rewrite $Eq.(12)$ as

$$\widehat{\beta}_{LAD-Lasso-Liu} = argmin_{\beta} \sum_{i=1}^{n+2p} \left| y_i^{\circ\circ} - \sum_{j=1}^{p} x_{ij}^{\circ\circ}\beta_j \right| \tag{13}$$

The new estimator in Eq. (12) achieved two shrinks for the parameters, the first shrink through Liu-Type estimator and second shrink through Lasso estimator. To study the asymptotic properties for the new estimator in $Eq.(12)$, we need to use the assumptions (A,B) for [15] . For this assumptions the error ε^{*} is a continuous and positive defined $(P.d)$ and has $(P.d)$ variance. These conditions about the error are necessary for deriving consistent estimators and ensuring that there is no perfect multicollinearity among the explanatory variables. We add the other assumption $\frac{d}{\sqrt{n}} \to d_0$. This assumption ensures that the effect of regularization stabilizes as the sample size grows. We assumption $\hat{\beta} = \beta_0$. This is a standard assumption in asymptotic analysis, ensuring that the coefficients do not change as the sample size increases. At the end, we assumption $k_{1_i} = \left(k_{1_1}, k_{1_2}, \cdots, k_{1_p}\right)$ where $max\left(k_{1_i}\right) = 0\left(1\right)$ and $k_{2_i} = \left(k_{2_1}, k_{2_2}, \cdots, k_{2_p}\right)$ where $max\left(k_{2_i}\right) = 0\left(1\right)$. These conditions are necessary to control the amount of regularization imposed on the estimates. We defended $\beta = \left(\beta_a^T, \beta_b^T\right)^T$ where $\beta_a = \left(\beta_1, \beta_2, \cdots, \beta_{p_0}\right)^T$ as nonzero coefficients and $\beta_b = \left(\beta_{p_0+1}, \beta_{p_0+2}, \cdots, \beta_p\right)^T$ as zero coefficients. Let$\hat{\beta}_{LAD-Lasso-Liu} = \left(\beta_a^T, \beta_b^T\right)^T$ is objective function for the LAD-Lasso-Liu estimator. Furthermore, let $\hat{x}_i = \left(x_{ia}^T, x_{ib}^T\right)^T$ where $x_{ia} = \left(x_{i1}, x_{i2}, \cdots, x_{ip_0}^T\right)^T$ and $x_{ib} = \left(x_{ip_0+1}, x_{ip_0+2}, \cdots, x_{ip}\right)^T$. . In addition, let $a_n = max\left(k_j : 1 \le j \le p_0\right), b_n = max\left(k_j : p_0 \le j \le p\right)$. To facilitate access to the asymptotic properties of the new estimator we will use the following lemma.

Lemma (1) : [31] as $Q_a = \lim_{n\to\infty} \frac{\left(x_a^T x_a\right)}{n}$, this ensures that the matrix has full rank, which is necessary for obtaining unique and consistent estimates, $\sqrt{n}\left(\hat{\beta} - \beta_0\right) \sim N\left(\frac{Q_a^{-1}\left(x_a^T x_a\right)Q_a^{-1}}{2f(0)^2}\right)$ where $f(t)$ is the density of errors ε_i.

Thearem (1) : Form the model (6), suppose that $\left(y_i^*, x_i^*\right)$, $i = 1, 2, \cdots, n$ are iid and identically distributed. Under A and B for [15], $\frac{d}{\sqrt{n}} \to d_0$, $\hat{\beta} = \beta_0$ and $k_{1_i} = \left(k_{1_1}, k_{1_2}, \cdots, k_{1_p}\right)$ where $max\left(k_{1_i}\right) = 0\left(1\right)$ and $k_{2_i} = \left(k_{2_1}, k_{2_2}, \cdots, k_{2_p}\right)$ where $max\left(k_{2_i}\right) = 0\left(1\right)$, then:

(i) If $\sqrt{n}a_n \to 0$ and $\sqrt{n}b_n \to \infty$, then the LAD-Lasso-Liu estimator is $\sqrt{n} - consistant$ and satisfy $p\left(\hat{\beta}_b = 0\right) \to 1$.

(ii) $\sqrt{n}\left(\beta_{LAD-\hat{L}asso-Liu} - \beta\right) \sim N\left(0, \frac{\Sigma_0^{-1}}{4f(0)^{-2}}\right)$ where $\Sigma_0^{-1} = Q_d^{*-1^T}\left(X_{a_{*T}} X_{a_*}\right) Q_d^{*-1}$ and $Q_d^* = \lim_{n\to\infty} \frac{\left(X_a^T X_a + K_1 I\right)}{n}$

Proof :

Since, $Y^* = \begin{bmatrix} Y \\ -d/k_1^{-1/2}\hat{\beta} \end{bmatrix}, X^* = \begin{bmatrix} X \\ nk_1^{1/2} I \end{bmatrix}$, then we can deal with the LAD-Lasso estimation problem for the Liu-Type regression model similarly with the LAD-Lasso estimation problem for the linear regression model. Directly from $Lemma\,(1)$ for [15], we Blade to prove the Theorem.

This result shows that the LAD-Lasso-Liu estimator is asymptotically normal with a variance that depends on the design matrix and the regularization parameters.

The LAD-Lasso-Liu estimator introduces two types of shrinkage: one from the Liu-Type estimator and another from the Lasso penalty. These two shrinkage mechanisms complement each other by stabilizing the estimates Liu shrinkage and promoting sparsity (Lasso shrinkage) . The Liu shrinkage reduces variance by stabilizing estimates in the presence of multicollinearity but introduces bias. The Lasso shrinkage promotes sparsity by shrinking some coefficients to zero, reducing the model complexity, but also introduces bias. The trade-off is between bias Lasso shrinkage (shrinkage increases bias) and variance (shrinkage decreases variance). A well-tuned regularization parameter will strike a balance, reducing the overall error of the estimator.

4. Choose tuning parameters:

The new estimator has three tuning parameters$(k_1, k_2 andd)$. To choose tuning parameters$(dandk_1)$, we will develop the tuning parameters for [12] as

$$\hat{k}_1 = \frac{\lambda_1 - 81\lambda_p}{99} : CN > 9 \tag{14}$$

where $CN = \sqrt{\lambda_p/\lambda_1}$ is condition number of the matrix $\left(X^T X\right)$.

We select the d tuning parameter as

$$\hat{d} = \frac{\sum_{j=1}^{p}\left(\left(\sigma \hat{_{RE_j}} - \hat{k}_{1_j}\alpha_{RE_j}^2\right) / \left(\lambda_j + \hat{k_{1_j}}\right)^2\right)}{\sum_{j=1}^{p}\left(\left(\lambda_j \sigma \hat{_{RE_j}} - \hat{k}_{1_j}\alpha_{RE_j}^2\right) / \lambda_j\left(\lambda_j + \hat{k_{1_j}}\right)^2\right)} \tag{15}$$

where $\hat{\sigma}_{RE_j}^2$ is the variance error for ridge estimator. Since, the Cross-Validation (CV) depended on minimizing the production error, so it strongly candidates to choose the tuning parameter. The general cross-validation (GCV), which introduced by [24], is the important way that is used to choose the Ridge parameter. [25] developed the statistic (GCV), to choose the tuning parameters for Liu-Type estimator. There are many studies used S-fold cross validation to choose tuning parameter. The S-fold cross validation divided the dataset into S subset S-fold with indices $E_j, 1 \leq j \leq S$. Such that, if $i \in E_j$ then (x_i, y_i) is a validation set . Then we choose the tuning parameter by minimizing

$$CV_\theta = n\sum_{1\leq j\leq k}^{S}\sum_{i\in j}\left(y_i - \sum_{j=1}^{p} x_{ij}\beta_{j_{(\theta)}}^{(S)}\right)^2 \tag{16}$$

Where $\beta_{j_{(\theta)}}^{(S)})$ is jth elements for unpenalized estimator.

In fact, the classification of the S-fold cross validation is non-robust cross-validation, so [26, 27, 28] used robust S-fold cross-validation in many ways to choose the tuning parameter for penalty regression. We use a type of robust S-fold cross-validation which depended on LAD function to select the tuning parameters. We defined it as

$$LAD - CV_{k_2} = n\sum_{1\leq j\leq k}^{S}\sum_{i\in j}\left|y_i - \sum_{j=1}^{p} x_{ij}\beta_{j_{LAD(k_2)}}^{(S)}\right| \tag{17}$$

where $\beta_{j_{LAD(k_2)}}^{(S)}$ is jth elements for LAD estimator. To select the tuning parameter, we set S=4 and then choose (k_2) that minimizes the $LAD - CV\,(k_2)$.

5. The simulation study:

In this section, we will test the performances of the new estimator in comparison with a group of other estimators. The LAD, Lasso and LAD-Lasso estimators are candidate in this direction. We use median mean square error $MedianMSE$ criterion to check it out. We defined it as

$$MedianMSE = Median\left(\frac{\left(\beta_0 - \hat{\beta}_f\right)^T\left(\beta_0 - \hat{\beta}_f\right)}{300}\right)^T \tag{18}$$

where $\hat{\beta}_f$ is the target estimator, and the number of iteration equal 300. The data is generated by normal distribution with mean 2 and variance ν_a, where

$\nu_a = (\rho_{ij})$ and ρ_{aij} represent the correlation between any two explanatory variables regression coefficient x_i, x_i with $i = 1, 2, \cdots, n$ and $i \neq j$. The correlation coefficient ρ was chosen 0.70 and 0.95. In adding, the dependent variable generated by the linear regression model

$$y_i = \sum_{j=1}^{p} x_{ij}\beta_j + \sigma\varepsilon_i \quad i = 1, 2, \cdots, n \tag{19}$$

where errors ε are generated with heavy-tailed distributions $\varepsilon \sim t\,(3)$ and $laplace\,(0,1)$ and two value for $\sigma\,(10, 0.5, 1)$ are used. We set $\beta_0 = (1, 2, 0, 0, 0, 1, 0, 2.5, 0, 0.5)$ for p=10, we has five significant regression variables. The $\hat{\beta}$ estimator is selected as OLS estimator. The number of observation was chose as n=50 and 100. We use(14), (15) and (16) to choose the tuning parameters. We used the MATLAB program to create a code which allow to calculate the Median Mean Squared Error (Median MSE), and compute the "Correct" and "Incorrect" values (i.e., the average number of zero and nonzero coefficients) for 300 simulated datasets as presented in Table $(1, 2)$. We use the following steps to estimate procedure for LAD-Lasso-Liu estimator:

Step 1: We generated the data set (y_i, x_i). We create the design matrix X with multicollinearity structure based on ρ.
Step 2: We set at $\beta_0 = (1, 2, 0, 0, 0, 1, 0, 2.5, 0, 0.5)$ and generated the error terms.
Step 3: Use the data set to choose the tuning parameters d, k_1, we used the Eq. 14,15 and 18.
Step 4: Building the augmented data set (y_i^*, x_i^*).
Step 5: Estimate the LAD estimator by minimizing $\sum_{i=1}^{n} \left| y_i - x_i^T \beta_i \right|$and use this estimator to compute k_2.
Step 6: We get the LAD-lasso estimator for the augmented data set $(y_i^\circ, x_i^\circ)\,, i = 1, 2, \cdots, n, n+p$.
Step 7: We get the LAD-lasso-Liu estimators for the augmented data set$(y_i^{\circ\circ}, x_i^{\circ\circ})\,, i = 1, 2, \cdots, n, n+p$.

For each estimator, the simulation procedure:
1. Generate 300 datasets.
2. For each dataset, estimate β using the defined estimators.
3. Compute the MSE for each simulation run.
4. Compute the median of these MSE values across the 300 runs.
5. Count the correct (zero) and incorrect (non-zero) selections of coefficients.

In the previous Table $(1, 2)$, simulation results have been summarized in columns" Correct." and " Incorrect." which represents the average number of zero coefficients and nonzero coefficients for 300 simulated datasets. In addition this is "Median MSE" column which represents the median mean square error in the same way as[14, 15, 16, 29, 30].
In table $(1, 2)$, we aim to make comparisons between a set of robust and penalized estimators. We use the LAD estimator, S-estimator and MM-estimator as robust un-penalized estimator and LAD-Lasso and LAD-Lasso-Liu as robust penalized estimator and Lasso and Liu-type estimator as penalized estimator.
The results in table $(1, 2)$ show that, when ρ increases, the LAD-Lasso and LAD-Lasso-Liu estimators works better than other estimators. All estimators perform better when the number of observations increases and the value of σ decreases. In addition, the LAD-Lasso-Liu is the best estimator when ρ increase and the value of σ decreases that means the LAD-LASSO-Liu is dealing well with the problems of outliers and multicollinearity. There is slight improvement for the estimators when $\varepsilon \sim t\,(3)$.
From Table $(1, 2)$, when comparing MM-estimator and S-estimator, with the LAD estimator, we observe that the LAD estimator has an advantage over the robust unpenalized estimators under comparison. Additionally, it is noted that the values of MediamMSE for all estimators converge at higher observation levels. The Liu-Type estimator performs poorly in handling outliers, but it achieves good results in dealing with strong correlations between variables.
On the other hand, the results show that LAD-Lasso-Liu and Lasso estimator have under fitting effect when the number of observations decrease. The LAD-Lasso-Liu estimator has better ability to select the variable than other penalized estimators like Lasso estimator, since the LAD-Lasso-Liu estimator has better correct selection of the number of zeros coefficients. From figures $(1, 2)$, we compare the estimated values of the LAD, S-estimator, MM-estimator, Lasso, Liu-type, LAD-Lasso and LAD-Lasso-Liu estimators using a box plot for two cases:$\varepsilon \sim Laplace\,(0,1)$ and $\varepsilon \sim t\,(3)$. The results in Figure $(1, 2)$ were consistent with those in Table $(1, 2)$.

Table (1): Summarize simulation result $(\rho = 0.70)$**.**

				$\varepsilon \sim Laplace\,(0,1)$			$\varepsilon \sim t\,(3)$		
σ	ρ	n	Method	Correct	Incorrect.	MedianMSE	Correct	Incorrect.	MedianMSE
10	0.70	50	LAD	0.000	0.000	3.214	0.000	0.000	3.314
			S-estimator	0.000	0.000	3.656	0.000	0.000	3.594
			MM-estimator	0.000	0.000	3.508	0.000	0.000	3.569
			Lasso	4.521	0.089	4.658	4.535	0.072	4.654
			Liu-type	0.000	0.000	4.193	0.000	0.000	4.202
			LAD-Lasso	4.321	0.012	3.023	4.421	0.015	3.584
			LAD-Lasso-Liu	4.621	0.011	2.158	4.731	0.011	2.325
0.5	0.70	50	LAD	0.000	0.000	3.115	0.000	0.000	3.245
			S-estimator	0.000	0.000	3.599	0.000	0.000	3.517
			MM-estimator	0.000	0.000	3.501	0.000	0.000	3.511
			Lasso	4.698	0.058	4.125	4.651	0.053	4.024
			Liu-type	0.000	0.000	4.034	0.000	0.000	4.127
			LAD-Lasso	4.392	0.025	2.954	4.402	0.022	3.056
			LAD-Lasso-Liu	4.689	0.009	2.035	4.663	0.008	2.125
10	0.70	100	LAD	0.000	0.000	3.024	0.00	0.000	2.987
			S-estimator	0.000	0.000	3.326	0.000	0.000	3.495
			MM-estimator	0.000	0.000	3.297	0.000	0.000	3.447
			Lasso	4.724	0.021	3.534	4.715	0.020	3.354
			Liu-type	0.000	0.000	3.986	0.000	0.000	4.094
			LAD-Lasso	3.608	0.013	1.962	3.599	0.012	1.985
			LAD-Lasso-Liu	4.721	0.007	1.542	4.687	0.006	1.412
0.5	0.70	100	LAD	0.000	0.000	2.958	0.000	0.000	2.865
			S-estimator	0.000	0.000	3.021	0.000	0.000	3.325
			MM-estimator	0.000	0.000	2.998	0.000	0.000	3.302
			Lasso	4.794	0.025	3.456	4.767	0.030	2.568
			Liu-type	0.000	0.000	3.845	0.000	0.000	3.973
			LAD-Lasso	3.861	0.032	1.895	3.991	0.028	1.962
			LAD-Lasso-Liu	4.798	0.008	1.421	4.554	0.004	1.385

6. Application to real data:

In this section, we apply real data to check the performance of the proposed estimator. We use the prostate cancer data by as [32]. This data contains 102 observations. For this data, we use prostate-specific antigen as a dependent variable Y which effected by eight independent variables (X,s). We use the logarithm for all variables. The independent variables are the cancer volumeX_1 , capsule penetration X_2, Gleason score X_3, amount of benign prostatic hyperplasia X_4, seminal vesicle invasion X_5, prostate weight X_6, age X_7 and the percentage of Gleason score 4 or 5 X_8. We test the normality for the dependent variable dependent on the Kolmogorov - Smirnov test. The results of test indicated that the dependent variable does not follow a normal distribution. In addition, we use the VIF to determine the level of multicollinearity. The value of VIF ranges between 4.5 and 7.8 which refers to the presence of moderate level of multicollinearity. Moreover, we use the MedianMSE criterion to compare between the estimators.

We summarize the results in table 3 which referred to the value of coefficients and MedianMSE. The results show that, the lasso, LAD-Lasso and LAD-Lasso-Liu estimators excluded X_3 for the model. In addition, the lasso and LAD-Lasso estimators excluded X_2 and X_7 from the model. The LAD-Lasso-Liu estimator has less zero coefficients than the other estimators which indicate of the direction of the estimator to achieve the accuracy of the forecast.

According to the results of MedianMSE, the LAD-Lasso-Liu estimator has the best performance in comparison to

Table (2): Summarize simulation result ρ=0.95.

				$\varepsilon \sim Laplace\,(0,1)$			$\varepsilon \sim t\,(3)$		
σ	ρ	n	Method	Correct	Incorrect.	MedianMSE	Correct	Incorrect.	MedianMSE
10	0.95	50	LAD	0.000	0.000	4.321	0.000	0.000	4.389
			S-estimator	0.000	0.000	4.415	0.000	0.000	4.325
			MM-estimator	0.000	0.000	4.208	0.000	0.000	4.182
			Lasso	4.154	0.051	2.254	4.228	0.044	2.401
			Liu-type	0.000	0.000	3794	0.000	0.000	3.456
			LAD-Lasso	4.362	0.013	2.153	4.388	0.011	2.307
			LAD-Lasso-Liu	4.632	0.010	1.222	4.668	0.008	1.241
0.5	0.95	50	LAD	0.000	0.000	4.021	0.00	0.000	3.542
			S-estimator	0.000	0.000	4.382	0.000	0.000	4.193
			MM-estimator	0.000	0.000	4.125	0.000	0.000	4.002
			Lasso	4.206	0.038	2.054	4.296	0.033	3.972
			Liu-type	0.000	0.000	4.025	0.000	0.000	4.389
			LAD-Lasso	4.389	0.021	1.857	4.407	0.018	3.021
			LAD-Lasso-Liu	4.694	0.011	1.125	4.728	0.009	0.897
10	0.95	100	LAD	0.000	0.000	3.995	0.00	0.000	3.368
			S-estimator	0.000	0.000	3.927	0.000	0.000	3.797
			MM-estimator	0.000	0.000	3.901	0.000	0.000	3.384
			Lasso	4.331	0.022	2.124	4.389	0.017	4.157
			Liu-type	0.000	0.000	3.768	0.000	0.000	3.658
			LAD-Lasso	3.584	0.031	1.867	3.689	0.028	1.635
			LAD-Lasso-Liu	4.764	0.005	0.968	4.865	0.004	0.857
0.5	0.95	100	LAD	0.000	0.000	3.587	0.00	0.000	3.202
			S-estimator	0.000	0.000	3.681	0.000	0.000	3.395
			MM-estimator	0.000	0.000	3.668	0.000	0.000	3.027
			Lasso	4.408	0.019	2.019	4.447	0.011	2.132
			Liu-type	0.000	0.000	3.482	0.000	0.000	3.339
			LAD-Lasso	3.878	0.019	1.765	3.905	0.012	1.784
			LAD-Lasso-Liu	4.861	0.004	0.686	4.921	0.003	0.694

other estimators. Table 4 presents a comparison of the proposed estimator with a set of penalized and non-penalized estimators, as well as robust estimators, at different levels of tuning parameters d, k_1 and k_2. These comparisons have shown that MedianMSE increases as k_1 and k_2. increase and as d decreases for all estimators. Additionally, the LAD-Lasso-Liu estimator has a relative advantage over all other estimators at the different levels of tuning parameters d, k_1 and k_2.

7. Future Examples of the Proposed Estimator:

7.1. Housing Data:

The dependent variable is the average house rent, which typically exhibits the presence of outliers. The independent variables include the number of rooms in the house, the distance to the city center, and the crime rate in the area. Additional variables can be introduced that are interrelated to ensure the presence of linear coupling. The goal of the estimator is to address the challenges posed by linear coupling and outliers, while simultaneously achieving an optimal selection of non-zero variables.

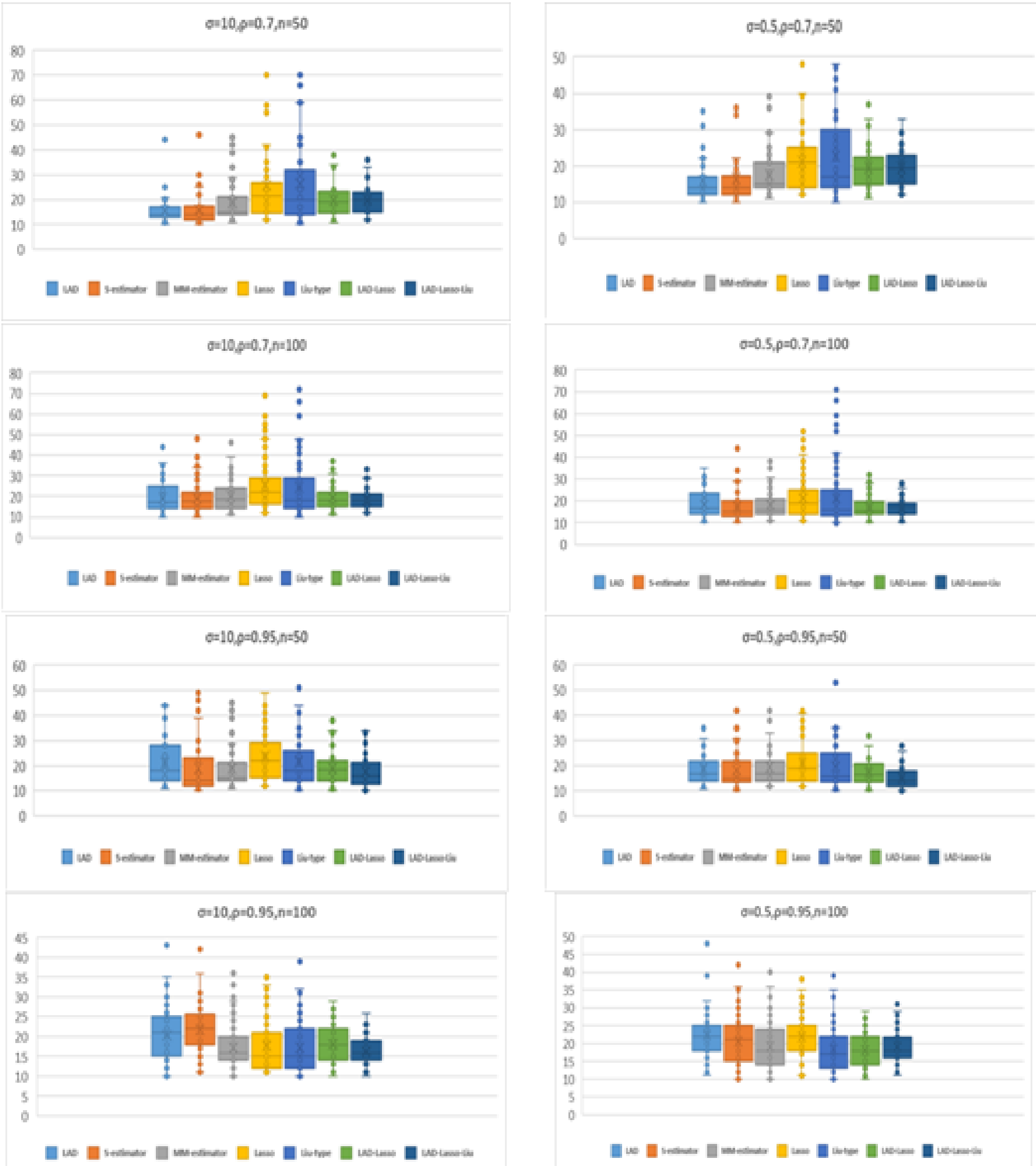

Figure 1. Comparing the estimated values of the LAD, S-estimator, MM-estimator, Lasso, Liu-type, LAD-Lasso and LAD-Lasso-Liu estimators using a box plot $\varepsilon \sim laplace\,(0.1)$

7.2. *Fuel Efficiency in Cars:*

The dependent variable is the distance a car travels per liter of gasoline, with some outlier readings present. The independent variables include the size of the car, engine power, car weight, number of engine failures, and a set of variables measuring the quality of the car engine. These variables contribute to linear coupling, and the goal is to select an optimal subset of non-zero.

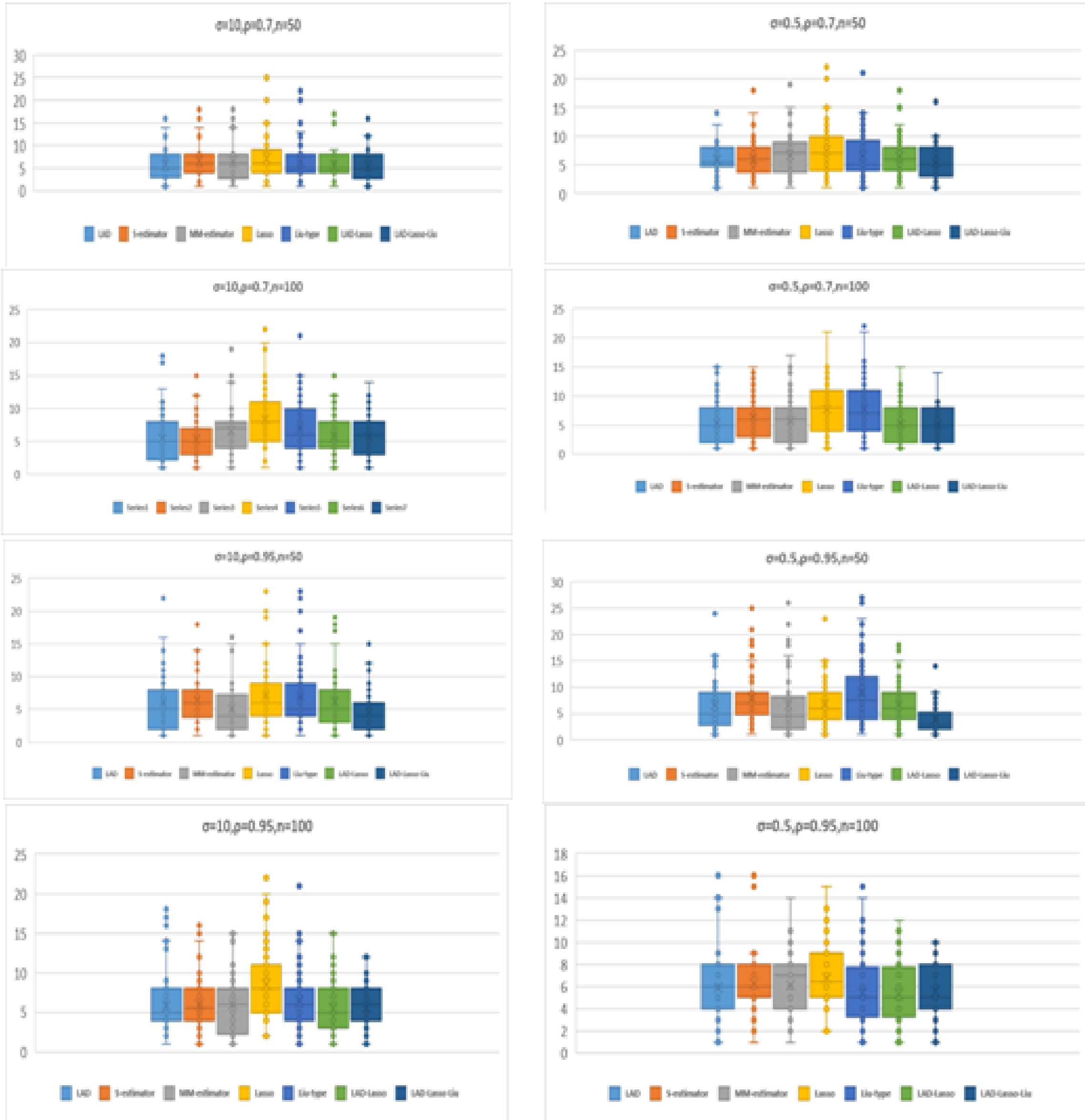

Figure 2. Comparing the estimated values of the LAD, S-estimator, MM-estimator, Lasso, Liu-type, LAD-Lasso and LAD-Lasso-Liu estimators using a box plot $\varepsilon \sim t\,(3)$

8. limitations and assumptions of the proposed estimator:

Although the LAD (Least Absolute Deviation) component provides robustness to vertical outliers, it may lack resilience against leverage points, which are outliers within the predictor variables. Consequently, these leverage points can exert a significant influence on the estimation process. In addition, the Lasso penalty introduces bias by shrinking some coefficient estimates toward zero. While this bias is advantageous for variable selection, it may also result in biased estimates, particularly in scenarios with strong multicollinearity or when the true coefficients are not sparse. Moreover, the performance of the LAD Lasso estimator is contingent upon the appropriate selection

Table (3): Coefficients and MedianMSE for the estimators to the dependent on the prostate cancer data.

Estimators	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	MedianMSE	No. of zero
LAD	5.3021	-0.252	-0.321	0.4485	1.625	1.772	-3.113	0.399	0.0485	0
S-estimator	4.003	0.235	0.485	0.4983	1.997	2.364	-3.895	0.487	0.0547	0
MM-estimator	4.231	0.244	0.427	0.558	1.704	1.895	-3.365	0.303	0.0507	0
Lasso	4.0631	0.000	0.000	0.5241	1.372	1.854	0.000	0.195	0.0614	3
Liu-type	4.002	0.164	0.582	0.469	1.092	1.472	-1.053	0.258	0.0214	0
LAD-Lasso	3.994	0.000	0.000	0.415	1.194	1.666	0.000	0.117	0.0351	3
LAD-Lasso-Liu	3.254	-0.124	0.000	0.384	1.038	1.502	-0.927	0.102	0.0012	1

Table (4): Coefficients and MedianMSE for the estimators at different levels of k_1, k_2 and d

	d=0.1				d=2			
Estimators	$k_1 = 0.001$		$k_1 = 0.05$		$k_1 = 0.001$		$k_1 = 0.05$	
	$k_2 = 0.01$	$k_2 = 0.1$	$k_2 = 0.01$	$k_2 = 0.1$	$k_2 = 0.01$	$k_2 = 0.1$	$k_2 = 0.01$	$k_2 = 0.1$
LAD	0.0921	0.0978	0.1394	0.1995	0.0881	0.0912	0.0954	0.1925
S-estimator	0.1387	0. 1645	0.1871	0.2015	0.2142	0.1215	0.1925	0.1858
MM-estimator	0.1025	0.1421	0.1672	0.1887	0.0975	0.1148	0.1486	0.1722
Lasso	0.1852	0.1965	0.2134	0.2571	0.1372	0.1785	0.2004	0.0214
Liu-type	0.0857	0.0942	0.1224	0.1895	0.0732	0.0685	0.0534	0.0501
LAD-Lasso	0.0942	0.1021	0.1125	0.1574	0.0824	0.0555	0.0596	0.0624
LAD-Lasso-Liu	0.0082	0.0091	0.0099	0.0117	0.0071	0.0084	0.0088	0.0101

of the regularization parameter. This parameter selection is typically challenging and necessitates the use of cross-validation or alternative model selection techniques. Improper selection of regularization parameter can lead to issues such as overfitting or underfitting the model.

Furthermore, the Liu-type regression model introduces an additional two biasing parameter which are designed to address multicollinearity. When combined with the LAD Lasso estimator, this parameter introduces another layer of complexity. The choice of the Liu-type parameter can significantly impact the estimates, and identifying the optimal value is often nontrivial. The simultaneous application of the LAD and Lasso techniques can also complicate the interpretation of the resulting model. This is especially true when considering the interplay between the Lasso penalty, which shrinks coefficients, and the Liu-type regression biasing parameter. The inherent bias-variance tradeoff in these methods can obscure the true relationship between the predictors and the response variable. Although the LAD method does not require the assumption of normally distributed errors, it may still underperform in situations where the error distribution exhibits heavy tails or when there is considerable heteroscedasticity (non-constant variance of errors). Finally, combining LAD with Lasso for Liu-type regression model adds complexity due to the need to balance parameter choices, which influence bias, variance, and model interpretation.

9. Conclusion:

In this study, we introduced the LAD-Lasso estimator for Liu-Type regression defined as LAD-LASSO-Liu-Type estimator to overcome the multicollinearity and outliers in the same time and also, it has ability to select the variable. This estimator has the properties of the three estimators, LAD, Lasso and Liu-Type estimators, so it has robust and penalized capabilities. The simulation study illustrated that, in many factors, the LAD-LASSO-Liu-Type estimator has superiority over LAD, Lasso, LAD-Lasso estimator according to the median mean square error (MedianMSE) criterion. The results of the application in real data for prostate cancer patients are identical with the results of the simulation study.

REFERENCES

1. M. Norouzirad ,S. Hossain , M. Arashi , *Shrinkage and penalized estimators in weighted least absolute deviations regression models*, Journal of Statistical Computation and Simulation, vol.88 no.8,pp.1557-1575, 2018.
2. O. Arslan , *Weighted LAD-LASSO method for robust parameter estimation and variable selection in regression*, Computational Statistics and Data Analysis, vol.56, no.6, 2012.
3. P. Phillips, *A Shortcut to LAD Estimator Asymptotics*, Economic theory, vol.7, no.3,pp.450-463,2013.
4. D. Yadolah , *LAD Regression for Detecting Outliers in Response and Explanatory Variables*, Journal of multivariate analysis, vol.61,pp.144-158 ,1997.
5. G. Avi , S. Jeffrey,S. Bhaskar , *Robust weighted LAD regression*, Computational Statistics and Data Analysis, vol.15,pp.3124-3140,2006.
6. L. Jinyu, L. Wei, H. Shuyuan, *Empirical likelihood for LAD estimators in infinite variance ARMA models*, Journal of Statistics Computation and Simulation, vol.81, no.2,pp.212-219, 2011.
7. T. Omara, *Weight LAD and Weight LAD Ridge Estimator for Seemingly unrelated regression Models*, Advances and Applications in Statistics, vol.52, no.6,pp.212-219, 2018.
8. F. Hampel, E. Ronchetti, P. Rousseeuw, W. Stahel, *Robust Statistics: The Approach Based on Influence Functions*, New York: Wiley, 1986.
9. A. Hoerl, R. Kennard, *Ridge Regression: Biased Estimation for Non-Orthogonal Problems*, Technometrics, vol.12, no.1,pp.55-67, 1970a.
10. T.Omara, *Modifying Two-Parameter Ridge Liu Estimator Based on Ridge Estimation*, Pakistan Journal of Statistics and Operation, vol. 15, no. 4,PP 881-890, 1990.
11. K. Liu , *New Class of Biased Estimate in Linear Regression*, Communications in Statistics—Theory and Methods, Vol.22, pp. 393-402, 1993.
12. Liu K. *Using Liu-type estimator to combat collinearity* , Communications in Statistics - Theory and Methods, vol.32, no.5, 2003.
13. T. Omara, *Almost unbiased modified ridge-type estimator: an application to tourism sector data in Egypt*, Heliyon, vol.8,no.9e10684, 2022.
14. R. Tibshirani , *Regression shrinkage and selection via the Lasso*, Journal of the Royal Statistical Society, Series B,pp.267-288.
15. H. Wang, G. Li and G. Jiang , *Robust regression shrinkage and consistent variable selection through the LAD-Lasso*, Journal of business and economic statistics, vol.25, 2007.
16. J. Kang , *weighted least absolute deviation Lasso estimator*, Communication of Korean statistical society, vol.18, no.6, pp.733-739., 2011.
17. W. Liu, *The L1 penalized LAD estimator for high dimensional linear regression*, Journal of Multivariate Analysis, vol. 120, pp.135-151, 2013.
18. L. Qiang and W.Liming , *Robust change point detection method via adaptive LAD-LASSO*, Stat Papers, vol. 61, no.1,PP.109-121, 2017.
19. Y. Hu and L. Ning, *WLAD-LASSO method for robust estimation and variable selection in partially linear models* Communications in Statistics - Theory and Methods, vol. 47, no.20, pp.4958-4976.,2018.
20. S. Yue,F. Zhiguo and Y. Cedric, *A descent method for least absolute deviation lasso Problems*, Optimization Letters, vol. 13, pp.543–559., 2017.
21. M. Genc , *A new double-regularized regression using Liu and lasso regularization*, Computational Statistics, vol. 37, no.2, pp.159-227, 2022.
22. M. Roozbeh, *Robust ridge estimator in restricted semi parametric regression models*, Journal of Multivariate Analysis, vol.147,pp.127-144, 2016.
23. S.Sakallioglu and S. Kaçiranlar, *A new biased estimator based on ridge estimation*, Statistic Paper, Vol 49,pp.669–689., 2008.
24. G. Golub , G. Wahba and M.Heath , *Generalized Cross Validation as a Method for Choosing a Good Ridge Parameter*, Technimetrics, vol.21, no.2, pp.215–223, 1979.
25. O. Fela , *Evaluation estimation performances of Liu Type and Two parameter using Monte Carlo experiments*, Annual Conference, Seville, 2018.
26. B. Efron, T. Hastie, I. Johnston, R. Tibshirani *Least angle regression*, The Annals of Statistics, vol.32, no.2, 2004.
27. E. Chi, D. Scott, *Robust parametric classification and variable selection by a minimum distance criterion*, Journal of Computational and Graphical Statistics, vol. 23, pp.111-128., 2014.
28. M. Ana, B. Graciela, C. Gonzalo, *Penalized robust estimators in logistic regression with applications to sparse models*, arXiv: 1911.00554 [stat.ME], vol.1 , 2019.
29. A. Olcay, *weighted LAD-LASSO method for robust parameter estimation and variable selection regression*, computational statistics and data analysis, Vol.56, 1952.
30. S.Hssan , *weighted Lasso sub sampling for high dimensional regression*, Electronic journal of applied statistical analysis, vol.12,no.1 , 2019.
31. A. Giloni, J. Simonoff and B. Sengupta, *Robust weighted LAD regression*, Computational Statistics and Data Analysis, vol.50, no.11, 2006a.
32. T. Stamey and J. Kabalin , *Prostate specific antigen in the diagnosis and treatment of adenocarcinoma of the prostate*, Official journal of the American Urological Association , vol.141, pp.1070-1075, 1989.
33. W. Lie , *The L1 penalized LAD estimator for high dimensional linear regression*, Journal of Multivariate Analysis , vol.120, pp.135-151, 2023.