# Bayesian Reciprocal Lasso for Right Censored Data with Application

By

Fatiam Muhammed Kahdim    Dr. Ahmad Naeem Flaih

## Abstract

This paper discusses the Bayesian reciprocal lasso (rlasso) regularization method as variable selection procedure that produced the more interpretability model with minimum set off predictor variables in right censored limited response variable. The reciprocal lasso introduced the reciprocal of L1-rorm in the penalty function of the penalized parameter estimates minimization problem. Reciprocal lasso is recently developed as regularization methods that produce a parsimonious regression model. We utilized the scale mixture of double Pareto (SMDP) and the scale mixture of truncated normal (SMTN) that discussed by Mallick et al. (2020) with a modification for (SMTN) in the hierarchical prior model. We used the (SMDP) and the modified (SMTN) in the right censored regression model with real data analysis. The results show that the employed two scale mixture types outperform other common regularization methods.

Keywords: reciprocal lasso, SMTN, scale mixture of double Pareto, Hierarchical prior model, Gibbs sampler.

1- Introduction

It is well known that the typical linear regression model is the usual tool of representing the $\boldsymbol{n}$ observations $(x_1, y_1)$, … ,$(x_1, y_n)$ selected randomly and independently from a particular population. The response (outcome) variable $y_i$; $i = 1, 2, \ldots, n$, is a function of the $p$ predictor variables $x_{i1}, x_{i2}, \ldots, x_{ip}$ , such that:

$$y_i = \beta_0 + \sum_{i=1}^{n}\sum_{j=1}^{p}\beta_j x_{ij} + u_i \ , \qquad i = 1,2,3, \ldots, n \ldots. (2.1)$$

the matrix form of model (2.1) can be defines as follows,

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{u}, \quad \ldots (2.2)$$

Where

$$\boldsymbol{y}_{\boldsymbol{n}\times\boldsymbol{1}} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, \boldsymbol{X}_{\boldsymbol{n}\times(\boldsymbol{p}+\boldsymbol{1})} = \begin{pmatrix} 1 & x_{11} & \cdots & x_{1p} \\ 1 & x_{21} & \cdots & x_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_{n1} & \cdots & x_{np} \end{pmatrix},$$

$$\boldsymbol{\beta}_{(\boldsymbol{p}+\boldsymbol{1})\times\boldsymbol{1}} = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_p \end{pmatrix}, \boldsymbol{u}_{\boldsymbol{n}\times\boldsymbol{1}} = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}.$$

$\boldsymbol{\beta}_j$ are the unknown parameters that we try to estimate and $\boldsymbol{u}$ is the random error. Ordinary least squares method results are called BLUE if the error term satisfy; $E(u) = 0$**,** $Var(u) = \sigma^2 I$, $u \sim N(O, \sigma^2 I)$**,** and *Cov* $(u_i, u_j) = 0$, where i ≠ j. (Chatterjee and Hadi 2006, AlNasser (2014)**.** Following the above assumption of error term and taking the expected of (2.1)**,** we have

$$E(y_i|x_i) = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_p x_{ip} = x_i^T \hat{\beta}$$

Then, the error is as follows,

$$\hat{u} = y_i - x_i^T \hat{\beta}$$

The following minimization problem represents the solution of **OLS**:

$$argmin \sum_{i=1}^{n} (u)^2 = argmin \sum_{i=1}^{n} (y_i - x_i^T \hat{\beta})^2$$

So, based on the above minimization problem , the ordinary least squares estimator defined by:

$$\hat{\beta} = (X^T X)^{-1} X^T y$$

The **OLS** estimator in (2.5) is unbiased with smallest variance**.**

See Chatterjee and Hadi (2013) and AlNasser (2014)**.**

The analysis of limited dependent (response) variable is widely observed in many applications, where there is a boundary or limit on the response variable which means there are some of the values of $\boldsymbol{y}$ reach this limit or boundary. Limited dependent variable leads us to the censored sample which its observations are $(\mathbf{y_1}, \mathbf{y_2}, \dots, \mathbf{y_n})$ resulting from a latent variable $(\mathbf{y}^*)$ based on some structural function form. An awareness of this type of dependent variable is very important, because adopting the inappropriate statistical tool will yields unsatisfied regression model. Hence, censored is only for the value of the dependent variable. In general there are three types of censoring value (from below (left), from above (right), interval). In this paper we are concerning in right censoring data. In the analysis of regression model, the number of covariates included in regression model brings the researcher to develop the mechanism of variable selection procedure. So, the variable selection procedure treated with the regression form specification. Tibshirani (1996) developed new regularization method named lasso which gives sparse solution for the linear regression coefficients. Lasso adds penalty function that include L1-norm function which controlled by the shrinkage parameter. The parameter estimates for some predictor variables reach the zero value and the solution regards as sparse solution.

The above mentioned regularization methods are frequents methods. But Park and Casella (2008) was the first work that introduced the Bayesian analysis for the regularization method based on lasso linear regression that developed the posterior distribution through new scale mixture for the prior distribution.

Mallick and Yi (2014) developed new scale mixture that mixed uniform distribution with particular gamma distribution $(2, \lambda)$ as the prior representation of the Laplace distribution. Therefore, based on the proposed scale mixture a new lasso solution has developed for the linear regression model, as well as, new hierarchical prior model and new Gibbs sampler algorithm have proposed. The new proposed model examined by simulation study and the results outperforms the new method over some exists regularization methods. Alhamzawi (2016) proposed new Bayesian elastic net in Tobit quantile regression model where the proposed method is sparsity. In this paper gamma priors was employed to develop the hierarchical prior model. New Gibbs sampler algorithm introduced for the MCMC algorithm. Simulation study have conducted to examine the proposed model in terms of variable selection procedure, also the proposed method have applied on real data and the results shows outperforms of the proposed model comparing with some penalized method. Hilali (2019) proposed a transformation for the scale mixture of double exponential prior distribution that developed by Mallick and Yi (2014). This new representation of the prior distribution employed into new hierarchical prior model and new Gibbs sampler algorithm. Bayesian adaptive lasso Tobit regression has used based on the new transformation. Variable selection procedure has examined under this proposed model with new posterior distribution. The results of simulation and real data analysis are comparable with some exists regularization methods. Flaih et al. (2020) proposed using scale mixture that mixed Rayleigh with normal distribution in lasso and adaptive lasso regression. Moreover, the proposed scale mixture employed in deriving new hierarchical prior model as well as new Gibbs sampler algorithm. The results of simulation real data analysis showed the outperforms of the proposed posterior distribution in part of variable selection and the efficiency of the proposed estimator.

The reciprocal lasso (rlasso) proposed by Song (2014) introduced the following penalty function:

$$P(\beta,\lambda) = \lambda \sum_{j=1}^{p} \frac{1}{|\beta_j|} I(\beta_j \neq 0) \cdots\cdots\cdots\cdots (1)$$

where $\lambda \geq 0$ is shrinkage parameter penalty function gives sparse solutions with infinity penalties , in contract of lasso that gives spares solution with  nearly zero penalty funds. The function (1) is decreasing in the interval $(0, \infty)$ , Discontinuous at zero.

Mallick et al. (2020) introduced the reciprocal Bayesian lasso by employing scale mixture of double pareto with truncated normal distribution. The liner reciprocal Bayesian lasso estimator is defined as follow

$$h(\beta) = \ argmin\ RSS + \lambda \sum_{j=1}^{p} \frac{1}{|\beta_j|} I(\beta_j \neq 0) \cdots\cdots\cdots\cdots (2)$$

Alhamzawi and Mallick (2021) introduced the Bayesian reciprocal lasso quantile regression by defined the following estimator:

$$Q(\beta) = argmin \sum_{i=1}^{n} \rho(y_i - x_i^T \beta) + \lambda \sum_{j=1}^{p} \frac{1}{|\beta_j|} I(\beta_j \neq 0)$$

Where $\rho(.)$ is the loss function.

Song (2014) the first work that concerned the reciprocal lasso estimators that have the oracle property. This work was discussed the Bayesian variable selection procedure for ultra-high dimensional linear regression through the strategy of split-and-merge. The estimators are consistent and have asymptotic properties that give better results than the elastic net and lasso methods. Song and Liang (2015), and Song (2018) discussed the reciprocal L1-norm Bayesian variable selection in lasso methods. Based on the type of the observations of response variable, the regression models are formulated. In many applications the observations of response variables are in some known ranges. So, the structural form of the regression observations ( censored) are based on unobserved latent variable.

The Censored model from the below limit (left constraint) is defined as follows: (Maddala, 1993)

$$y_i = \begin{cases} y_i^* & if \quad y_i^* \geq c \\ c & if \quad y_i^* < c \end{cases}$$

Where $y_i^* = x_i' \beta + u_i$ , $y_i^*$ is the latent variable (unobservable variable). The vector $(y_1, y_2, \ldots, y_n)$ is the censured sample and $\beta$ is the $(1 \times k)$ vector of unknown coefficients, $x_i$ is the vector of known observations (constants), $u_i$ is the error term, $u_i \sim N(0, \sigma^2)$, and $c$ is the known constant (censored point). Setting $c = 0$ in the above mode structure yields Tobit regression model.

The Censored model from the upper limit (upper constraint) is defined as follows:

$$y_i = \begin{cases} \mathrm{y}_\mathrm{i}^* & if \quad \mathrm{y}_\mathrm{i}^* < c \\ c & if \quad \mathrm{y}_\mathrm{i}^* \geq c \end{cases} \quad \ldots \ldots (3)$$

We can equivalently write (3) as follows:

$$y_i = \min(y_i^*, c) \ \text{ or } \ y_i = \delta_i y_i^* + (1 - \delta_i)c,$$

Where $\delta_i = \mathrm{I}_{(\mathrm{y}_\mathrm{i}^* < c)}$ is the censoring indicator.

## 2- Bayesian Hierarchical Prior models

Referring to the formula (1), the structure model (3), and with some modification for the above proposition and based on the work of Park and Casella (2008), we propose the following hierarchical prior model:

$$y_i = \begin{cases} x_i'\beta + u_i & if \quad x_i'\beta + u_i < c \\ c & if \quad x_i'\beta + u_i \geq c \end{cases} , \qquad \ldots (4)$$

$$y_i^* | x_i'\beta, \sigma^2 \sim N(x_i'\beta, \sigma^2 I_n); \quad i = 1,2, \ldots . n$$

$$y^* = X_i'\beta + e_i,$$

$$\beta | \sigma^2, \tau \sim \prod_{j=1}^{p} N(0, \sigma^2 \, \tau^2) \, ,$$

$$\tau_1^2, \ldots, \tau_p^2 \sim \prod_{j=1}^{p} \frac{\delta^2}{2} e^{-\delta^2 \tau_j^2 / 2} d \, \tau_j^2, \ \ \tau_1^2, \ldots, \tau_p^2 > 0,$$

$$\delta^2 | \eta \sim Gamma \ (k, \eta),$$

$$\eta | \lambda \sim Inverse \ Gamma \ (2, \lambda),$$

$$\sigma^2 \sim \pi(\sigma^2) \propto \frac{1}{\sigma^2}$$

## 3. The Gibbs Sampler for proposed Model

Now we can implement the hierarchical model (4) with a Gibbs sampler algorithm. The Gibbs sampler algorithm is a Markov Chain Monte Carlo (MCMC) algorithm that generates samples from the conditional distribution of a specific parameter given all other parameters. The hierarchical model (4) constructed in such a way that we can formulate the full conditional distributions which provides easy simulation. Now we can write the full joint density as follows:

$$f(y^* \mid \beta, \sigma^2)\, \pi(\sigma^2) \prod_{j=1}^{p} \pi(\beta_j \mid \tau_j^2, \sigma^2)\, \pi\left(\tau_j^2\right) \pi(\delta^2)\, \pi(\eta) =$$

$$\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n e^{-\frac{1}{2\sigma^2}(y^*-X\beta)'(y^*-X\beta)}$$

$$\left(\frac{1}{\sigma^2}\right) \prod_{j=1}^{p} \frac{1}{(2\sigma^2\tau_j^2)^{1/2}} e^{-\frac{\beta_j^2}{2\sigma^2\tau_j^2}} \frac{\delta^2}{2} e^{-\delta^2\tau_j^2/2} \frac{1}{\eta^k} e^{-\frac{\delta^2}{\eta}} \frac{\lambda^2}{\Gamma 2} (\eta)^{-2-1} e^{-\frac{\lambda}{\eta}} \; .$$

…(5)

Based on the hierarchical model (4) and the full joint density (5) it is easy to sample $y^*, \beta, \sigma^2, \tau^2, \delta^2, \eta, \lambda$. The full conditional posterior distributions are as follows:

**1.** The full conditional distribution of $y^*$ is :

$$y_i^* / y_i, \beta \sim N_n(X_i'\beta, \sigma^2 I_n) \qquad \ldots (6)$$

**2.** The full conditional distribution of $\beta_j$ is :

$$\pi(\beta/y_i^*, X, \tau^2, \sigma^2) \propto \pi(y_i^*/X, \beta, \sigma^2).\pi(\beta/\tau^2)$$

$$\propto e^{-\frac{1}{2\sigma^2}(y^*-X\beta)'(y^*-X\beta)}.e^{-\frac{1}{2\sigma^2}\beta' D_\tau^{-1}\beta}$$

Where $D_\tau = diag(\tau_1^2, \ldots, \tau_p^2)$,

$$= \exp\left\{-\frac{1}{2\sigma^2}[(\beta'(X'X)\beta - 2y^*X\beta + y^{*\prime}y^*) + \beta' D_\tau^{-1}\beta]\right\}$$

$$= \exp\left\{-\frac{1}{2\sigma^2}[\beta'(X'X + D_\tau^{-1})\beta - 2y^*X\beta + y^{*\prime}y^*)]\right\}$$

Now let $C = X'X + D_\tau^{-1}$, then we have

$$= \exp\left\{-\frac{1}{2\sigma^2}[\beta' C\beta - 2y^*X\beta + y^{*\prime}y^*)]\right\}$$

$$= \exp\left\{-\frac{1}{2\sigma^2}(\beta - C^{-1}X'y^*)'C(\beta - C^{-1}X'y^*)\right\} \quad \ldots(7)$$

Which is the multivariate normal with mean $C^{-1}X'y^*$and variance $\sigma^2 C^{-1}$.

**3.** The full conditional posterior distribution of $\sigma^2$ is:

$$\pi(\sigma^2/y_i^*, X, \beta)$$

$$\propto \pi(y_i^*/X, \beta, \sigma^2).\pi(\beta/\sigma^2).\pi(\sigma^2)$$

$$\propto -\frac{1}{(\sigma^2)^{\frac{n}{2}}} e^{-\frac{1}{2\sigma^2}(y^*-X\beta)'(y^*-X\beta)}.\frac{1}{(\sigma^2\tau^2)^{\frac{p}{2}}} e^{-\frac{1}{2\sigma^2}\beta' D_\tau^{-1}\beta}.\frac{1}{\sigma^2}$$

$$= (\sigma^2)^{\frac{n}{2}+\frac{p}{2}-1} \exp\left\{-\frac{1}{2\sigma^2}(y^* - X\beta)'(y^* - X\beta) + \beta' D_\tau^{-1}\beta\right\}$$

$$= (\sigma^2)^{\frac{n+p}{2}-1} \exp\left\{-\frac{1}{2\sigma^2}[(y^* - X\beta)'(y^* - X\beta) + \beta' D_\tau^{-1}\beta]\right\} \quad \ldots(8)$$

which is the invers gamma with shape parameter $\frac{n+p}{2} - 1$ and scale parameter $(y^* - X\beta)'(y^* - X\beta)/2 + \beta' D_\tau^{-1}\beta/2$.

**4.** The full conditional posterior distribution of $\tau^2$is:

$$\pi\left(\tau_j^2/\delta^2\,,\eta\right) \propto \pi\left(\beta/\tau_j^2\right).\pi(\delta^2/\eta)$$

$$\propto \left(\frac{1}{\tau_j^2}\right)^{\frac{1}{2}} e^{-\frac{1}{2\sigma^2}\frac{\beta_j^2}{\tau_j^2}} e^{-\frac{\delta^2}{\tau_j^2}}$$

$$\propto \left(\tau_j^2\right)^{-\frac{1}{2}} \exp\left[-\frac{1}{2}\left(\frac{\beta_j^2/\sigma^2}{\tau_j^2} + \delta^2\tau_j^2\right)\right] \quad \cdots\cdots (9)$$

The last formula can be treated by using the invers Gaussian distribution and its invers form. Suppose that the invers Gaussian is:

$$f(X;a,b) = \left(\frac{b}{2\pi X^3}\right)^{\frac{1}{2}} \exp\left[\frac{-b(X-a)^2}{2a^2X}\right]$$

The invers of $f(X;\,.)$ is $f'(\,.)$ defined by

$$f'(y\,;a,b) = \left(\frac{b}{2\pi y}\right)^{\frac{1}{2}} \exp\left[\frac{-b(1-ay)^2}{2a^2y}\right]$$

Where $= X^{-1}$ , then a formula (8) Can be rewrite as the reciprocal inverse Gaussian distribution as follows:

$$\propto \left(\frac{1}{\tau_j^2}\right)^{\frac{-3}{2}} \exp\left[-\frac{1}{2}\left(\frac{\beta_j^2}{\sigma^2\,\tau_j^2} + \frac{\delta^2}{1/\tau_j^2}\right)\right]$$

$$\propto \left(\frac{1}{\tau_j^2}\right)^{\frac{-3}{2}} \exp\left[-\frac{\beta_j^2\left(\left(1/\tau_j^2\right)-\sqrt{\delta^2\sigma^2/\beta^2}\right)^2}{2\sigma^2\left(1/\tau_j^2\right)}\right] \ldots (9)$$

So, we can say that $\left(\frac{1}{\tau_j^2}\right) \sim invers\ Gaussain\ with\ mean\ \sqrt{\frac{\delta^2\,\sigma^2}{\beta_j^2}}$ and shape parameter $\delta^2 = b$.

**5.** By following Park and Casella (2008), we assigned the gamma prior for $\delta^2$. Then full conditional posterior distribution of $\delta^2$ is defined as in follows:

$$(\delta^2)^{k-1} e^{-\frac{\delta^2}{\eta}} (\prod_{j=1}^{p} \frac{\delta^2}{2} e^{-\delta^2 \tau_j^2/2})$$

$$=(\delta^2)^{p+k-1} \exp[-\delta^2 (\tfrac{1}{2}\textstyle\sum_{j=1}^{p} \tau_j^2 + \tfrac{1}{\eta})] \quad \ldots(10)$$

This is also a gamma distribution with shape parameter $p + k$ and rate parameter $\frac{1}{2}\sum_{j=1}^{p} \tau_j^2$ .

**6.** The full conditional posterior distribution of $\eta$ is defined as follows:

$$\pi(\eta/\delta^2, \lambda) \propto \pi(\delta^2/\eta).\pi(\eta|\lambda)$$

$$\propto (\delta^2)^{k-1} e^{-\frac{\delta^2}{\eta}} \frac{\lambda^2}{\Gamma 2} \eta^{-2-1} e^{-\lambda/\eta}$$

$$\propto \eta^{-2-1} e^{-\frac{1}{\eta}(\delta^2+\lambda)} \qquad \ldots(11)$$

Recall the invers gamma distribution, consequently we can conclude that $\eta$ is distributed according to inverse gamma with shape parameter (2) and scale parameter $(\delta^2 + \lambda)$.

The following parameters and variables have sampled based on Gibbs sampling algorithm:

**1-** Sampling $\boldsymbol{y}^*$: we generate the latent variable $y^*$ from truncated normal distribution with mean $(\boldsymbol{x_i}^T \boldsymbol{\beta})$ and variance $(\boldsymbol{\sigma^2 I_n})$.

**2-** Sampling $\boldsymbol{\beta}$: we generate $\boldsymbol{\beta}$ from normal distribution $\boldsymbol{C}^{-1}\boldsymbol{X}'\boldsymbol{y}^{*}$and variance $\boldsymbol{\sigma}^{2}\boldsymbol{C}^{-1}$.

**3-** **S**ampling $\boldsymbol{\sigma}^{2}$: we generate $\boldsymbol{\sigma}^{2}$ from invers gamma with shape parameter $\frac{\boldsymbol{n}+\boldsymbol{p}}{\mathbf{2}} - \mathbf{1}$ and scale parameter $(\boldsymbol{y}^{*} - \boldsymbol{X\beta})'(\boldsymbol{y}^{*} - \boldsymbol{X\beta})/\mathbf{2} + \boldsymbol{\beta}'\boldsymbol{D}_{\boldsymbol{\tau}}^{-1}\boldsymbol{\beta}/\mathbf{2}$.

**4-** Sampling $\boldsymbol{\tau}^{2}$ **:** we generate $\boldsymbol{\tau}^{2}$ from inverse Gaussian with mean $\sqrt{\frac{\boldsymbol{\delta}^{2}\,\boldsymbol{\sigma}^{2}}{\boldsymbol{\beta}_{j}^{2}}}$ and shape parameter $\boldsymbol{\delta}^{2}$.

**5-** Sampling $\delta^{2}$ **:** we generate $\boldsymbol{\delta}^{2}$ from a gamma distribution with shape parameter $\boldsymbol{p} + \boldsymbol{k}$ and rate parameter $\frac{\mathbf{1}}{\mathbf{2}}\sum_{\boldsymbol{j}=\mathbf{1}}^{\boldsymbol{p}} \boldsymbol{\tau}_{\boldsymbol{j}}^{2}$ .

**6-** Sampling $\boldsymbol{\eta}$: we generate $\boldsymbol{\eta}$ inverse gamma with shape parameter **(2)** and scale parameter $(\boldsymbol{\delta}^{2} + \boldsymbol{\lambda})$**.**

## 4. Extension on proposed Models

In this section we employed the proposition and the hierarchical model that developed by Mallick et al. (2020) in the Bayesian reciprocal Laplace right censored regression model. Scale Mixture of Double Pareto (SMDP) formulation proposed by Mallick et al. (2020) which is state that if the prior distribution of $\beta$ is β~Double inverse pareto (η, 1) and η~ inverse gamma (2, λ), then β follows inverse Laplace (λ).

$$y_i = \begin{cases} x_i'\beta + u_i & if \quad x_i'\beta + u_i < c \\ c & if \quad x_i'\beta + u_i \geq c \end{cases} ,$$

$$y_i^{*}|x_i'\beta,\sigma^2 \sim N(x_i'\beta, \sigma^2 I_n); \quad i = 1,2, \ldots . n$$

$$\beta|\eta \sim \prod_{j=1}^{p} \frac{1}{uniform\left(-\frac{1}{\eta_j}, \frac{1}{\eta_j}\right)}$$

$$\eta|\lambda\sim\prod_{j=1}^{p}Gamma(2,\lambda)$$

$$\sigma^2\sim\pi(\sigma^2) \tag{12}$$

Connection with Bayesian lasso and reciprocal lasso the full conditional posterior distribution for the parameters in hierarchical prior model (5) of the (SMDP) Bayesian reciprocal Laplace right censored regression model are as follows Mallick et al. (2020):

$$y_i^*/y_i,\beta\sim N_n(X_i'\beta,\sigma^2 I_n)$$

$$\beta|y^*,X,u,\lambda,\sigma^2\sim N_p\left(\hat{\beta}_{mle},\sigma^2\left(\acute{X}X\right)^{-1}\right)\prod_{j=1}^{p}I\left\{|\beta_j|>\frac{1}{\sigma^2 u_i}\right\},$$

$$u|y^*,X,\beta,\lambda,\sigma^2\sim\prod_{j=1}^{p}exp(\lambda)\,I\left\{u_j>\frac{1}{\sigma^2|\beta_j|}\right\},$$

$$\sigma^2,y^*,X,\beta,u,\lambda\sim Inv.Gamma\left(\frac{n-1}{p},\frac{1}{2}(y^*-X\beta)'(y^*-X\beta)\right)$$

$$\lambda|\beta\sim Gamma(a+2p,b+\textstyle\sum_{j=1}^{p}\frac{1}{|\beta_j|}). \tag{13}$$

## 5- Real Data Analysis

In this section we summarized the described of the real data that we collected from the central laboratory and Rafidain's Valley in the province of Babylon. To cope with the objectives of this paper we focused on limited dependent response variable (right censored). After employing the simulation method to show the preference of our proposed method in estimating parameters and selecting variables compared to a group of previous methods. We will test the behavior of our method with real data, which also focuses on a medical phenomenon that includes the response variable that represents the normal blood sugar level within the range (80-180) for 55 patients. In this study, we focus on the normal limits of blood sugar, so the censored point is 180 and when the values are above the censored point then it will be set to 180. The independent variables represented are as follows:

X1: the patient's weight (in kilograms).

X2: the patient's age.

X3: the number of meals for the patient per day

X4: Are there genetic factors?

X5: Is the patient under psychological pressure?

X6: Does the patient have pancreatic disease?

X7: Does the patient have covid19?

X8: the patient's monthly income

X9: The number of hours of exercise per day

For comparison purpose we employed the two proposed regularization methods (SMTN-reciprocal Lasso R.C. regression and SMDP-reciprocal Lasso R.C. regression) with two other methods (R.C. regression model and Bayesian lasso R.C. regression) by using the median mean absolute deviation (MMAD) and the mean absolute error (MAE) criterion. These criteria are used to assess the prediction accuracy of the different models.

Table (1). Values of MSE and MAE with its Standard errors

| | **R.C. regression model** | **Bayesian lasso R.C. regression** | **SMDP-reciprocal Lasso R.C. regression** | **SMTN-reciprocal Lasso R.C. regression** |
|---|---|---|---|---|
| MSE | 0.852(0.493) | 0.847(0.506) | 0.848(0.495) | 0.482(0.328) |
| MAE | 0.848(0.507) | 0.757(0.416) | 0.712(0.392) | 0.573(0.377) |

Figure (1) shows the values of the MSE and MAE, where the proposed models give the less values comparing with the other two methods. This result indicates the high prediction accuracy for the proposed models.

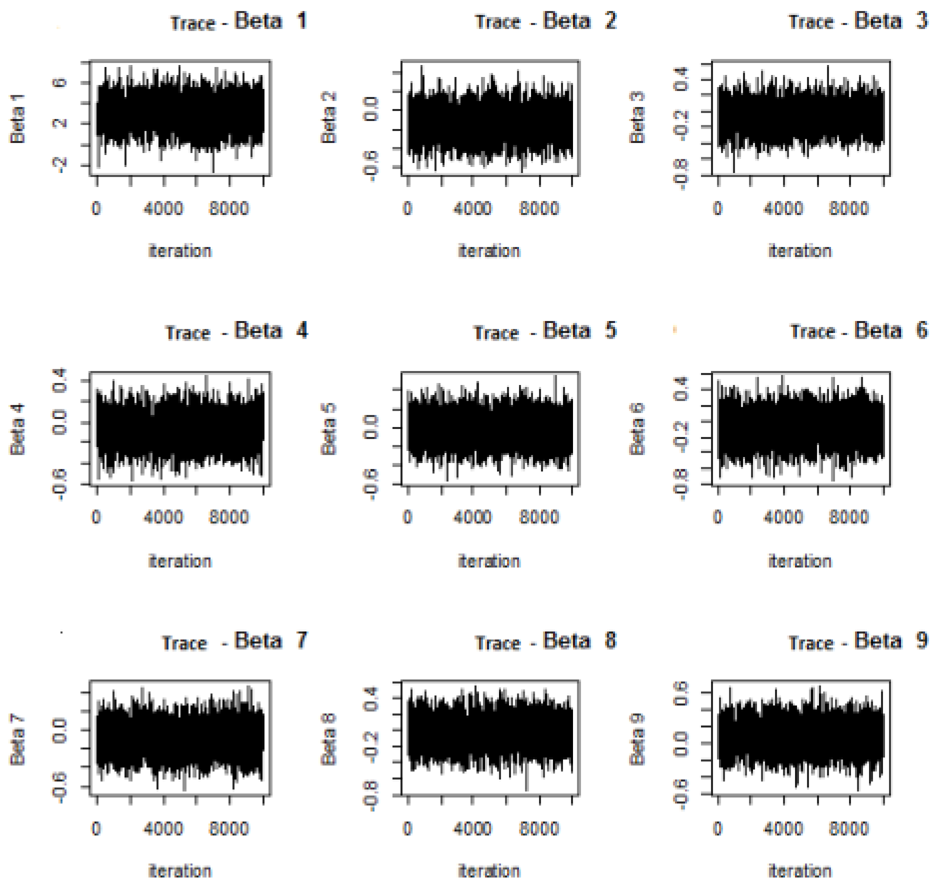


Figure (13). Trace Plot of Real Parameters

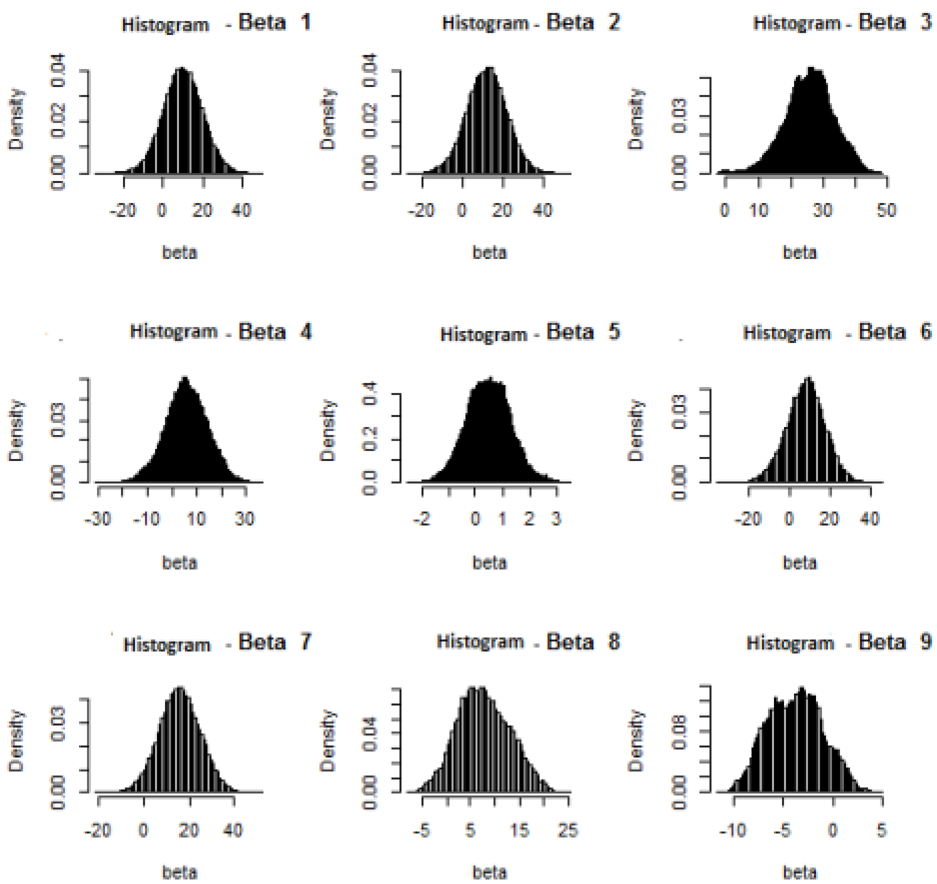


Figure (1) Histograms of real parameters

The above figure (1) shows the trace plots for the parameter estimates with the predictor variable observations of the (SMDP) model, which indicates the stationary of the proposed MCMC algorithm. Also, figure (1) illustrated that the proposed regularization methods gives parameter estimates follows the normal distribution under the (SMDP) model.

## 5- Conclusions

We employed the scale mixtures to examine the performance of the introduced Bayesian reciprocal lasso in right censored regression model according to the suggested hierarchical model. In addition to that we focused on the comparison of the quality of the coefficient estimates and variable selection problem with real data. Therefore, we used to criterion to test the performance of coefficient estimation methods; the median mean absolute deviation **(MMAD)** and standard Error **(S.E)**. The real data analysis shows that the proposed models give comparable results and outperform the other methods. Also, from figure (1) and figure (2) we can conclude that the results of the MCMC algorithm are stable in the convergence and the parameter estimates are follows the normal distributions which support the theoretical facts in this paper .

# References

Alhamzawi, R., & Mallick, H. (2020). Bayesian reciprocal LASSO quantile regression. Communications in Statistics-Simulation and Computation, 1-16.

AlNasser, Hassan (2014). On Ridge Regression and Least Absolute Shrinkage and Selection Operator. B.Sc., University of Victoria.

Chatterjee, S. & Hadi, A.S (2013). Regression Analysis by example. Wiley series in probability and statistics, Fifth edition.

Hilali, H., K., A. (2019). Bayesian adaptive Lasso Tobit regression with a practical application. MSc. Thesis. Statistics Department College of Administration and Economics University of Al-Qadisiyah.

Flaih, A.N, Alshaybawee, and Al husseini F.H, (2020). sparsity via new Bayesian Lasso . Periodicals of Engineering and Natural Sciences , vol.8 , is suc 1, 345-359.

Maddala,G.S.(1993). Limited-Dependent and Qualitative Variables in Econometrics. Cambridge University Press.

Mallick, H., & Yi, N. (2014). A new Bayesian lasso. Statistics and its interface, 7(4), 571-582.

Mallick, H., R. Alhamzawi, and V. Svetnik (2020). The reciprocal Bayesian lasso. arXiv preprint arXiv:2001.08327.

Park, T., & Casella, G. (2008). The bayesian lasso. Journal of the American Statistical Association, 103(482), 681-686. Choi, W. W., Weisenburger, D. D., Greiner, T. C., Piris, M.

Song,Qifan.(2014).Variable selection for ultra-high dimensional data. Phd. Dissertation,Texas A&M University.

Song, Qifan. (2018). An overview of reciprocal L1-regularization for high dimensional regression data. Wiley Interdisciplinary Reviews: Computational Statistics, 10(1):e1416, 2018.

Song, Qifan and Liang,Faming. (2015). High-dimensional variable selection with reciprocal L1- regularization. Journal of the American Statistical Association, 110(512):1607{1620, 2015.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society: Series B (Methodological), 267-288.