Model selection in quantile regression using inverse Laplace prior density

Maytham Saeed Kadhim and Prof. Dr. Rahim Al-Hamzawi

Email: mohammedsaeedkadhim1999@gmail.com

Department of Statistics, College of Administration and Economics, University of Al-Qadisiyah, Iraq.

Abstract

The problem of identifying the active covariates within a linear regression model has received much attention over the recent years. Very recently, Bayesian model selection methods employing nonlocal priors have received considerable attention. One of these methods in linear regression is the simplified shotgun stochastic search with screening (Shin et al., 2018). In fact, Shin et al., (2018) proposed two simplified shotgun stochastic search with screening algorithms. The first one is based on the product inverse-moment (piMoM) prior density (Johnson and Rossell, 2012) and the second one is based on the product exponential moment (peMoM) prior density (Rossell et al., 2013) **.** *In this paper , a new idea has been proposed through using simplified shotgun stochastic search with screening via using Bayesian approach in quantile regression model. Two simulated examples show the our proposed method perform well compared other methods in the same filed.*

Key Words *: Model Selection , Shotgun Stochastic Search, inverse-moment, exponential moment*

Introduction

1 – Introduction

Since it is introduction in Koenker and Bassett (1978), quantile regression (QR) models have been studied in-depth. It is insensitive to outliers which are unusual values in the data. QR is able to accommodate non-normal errors, which are common in a lot of real applications (Benoit et al., 2013). The θth quantile (Q) of a specific distribution is interpreted as the value such that there is $100\theta\%$ of mass on its left side. Compared to the conditional expectation, quantiles (Qs) are more robust to outliers.Model selection is important for sparse high dimensional data analysis in many fields of modern science such us economics, genetics, genomics, tomography and tumor classifications. A great body of work exists on model selection in the literature from both frequentist and Bayesian standpoints, such as the least absolute shrinkage and selection operator (LASSO, Tibshirani, 1996), smoothly clipped absolute deviation (SCAD, Fan and Li, 2001), the adaptive LASSO (Zou, 2006), the elastic net (Zou and Hastie, 2005), the adaptive elastic net (Zou and Zhang, 2009), the Bayesian LASSO (Hans, 2009; Park and Casella, 2008), the Bayesian adaptive LASSO (Alhamzawi et al., 2012) and the Bayesian elastic net (Li et al., 2010).However, the performance of these methods is usually discounted as the dimensionality grows fast. To overcome this problem, Hans et al. (2007) proposed a Bayesian method for variable selection, with a simple and efficient shotgun stochastic search (SSS) algorithm to explore subsets of covariates that are in the same neighbourhood. Fan and Lv (2008) proposed a sure independent screening (SIS) method to select active set of covariates in ultrahigh-dimensional linear models by considering only those covariates which have a large correlation with the residuals of the current model. Recently, Shin et al. (2018) proposed a Simplified Shotgun Stochastic Search with Screening (S5) algorithm to explore the enormous model space and reduces the computing time by using the idea of SIS. Specifically, Shin et al. (2018) proposed S5 algorithms based on the product exponential moment prior densities (Rossell et al., 2013) and the product inverse-moment prior densities (Johnson and Rossell, 2012) for the regression coefficients. By using simulation studies and real data analysis they show that their algorithm is effective in model selection and able to accelerate the computation speed under a variety of scenarios .Motivated by their empirical finding, we extend the S5 algorithm to quantile regression using independent inverse Laplace prior densities for the regression coefficients. Over the current decade, model selection in quantile regression has receivedconsiderable attention (for example see, Alhamzawi and Yu, 2014; Belloni et al.,2011; Bradic et al., 2011; Chen et al., 2013; Lamarche, 2010; Li et al., 2010; Zheng et al., 2013). However, variable selection in quantile regressionby using S5 algorithm (or in short, S5-QR) has not been roposed, yet. Instead of using the product exponential moment prior densities and the product inverse moment prior densities, we use the inverse Laplace prior densities for the regression coefficients. Under this prior, the Bayesian posterior mode estimate is equivalent.to reciprocal Lasso estimate (Mallick et al., 2019), which is not proposed yet in quantile regression. As demonstrated later by simulations, SS-QR provides more accurate estimates and better prediction accuracy than other existing methods in quantile regression. The rest of our paper is divided as follows . The quantile regression with reciprocal LASSO penalty, Posterior inference and ModelSetup have been showed in section two. In section three, we well introduced Simplified Shotgun Stochastic Search with Screening in QR . Simulation scenarios has been introduced in section four . In Section five, we conclude our paper by conclusion and recommendations .

2.Methods

2.1 Quantile Regression with reciprocal LASSO penalty

In a linear quantile regression setup, we have the following model:

$$\boldsymbol{y} = \boldsymbol{XB} + \boldsymbol{\epsilon}, \qquad \boldsymbol{\theta} \in (\mathbf{0}, \mathbf{1}), \qquad (1)$$

where y is the $n \times 1$ vector of centered responses, $X = (x1; \cdots; xn)$ is the $n \times p$ matrix of standardized regressors, β is the $p \times 1$ vector of quantile coefficients to be estimated, and ϵ is the $n \times 1$ vector of errors whose distribution is restricted to have the θth quantile equal to zero. The regression coefficient vector β and the error term ϵ should be indexed by θ, i.e. β_θ and ϵ_θ For sake of simplicity, however, we will omit θ in the rest of this paper. The unknown parameter vector β is estimated by minimizing (Koenker and Bassett, 1978).

$$\boldsymbol{min_\beta \sum_{i=1}^{n} \rho\theta\ (yi - \acute{x}_i\beta\ )}, \quad (2)$$

where $\rho\theta(w) = w\{\theta - I(w < 0)\}$ and $I(.)$ is the indicator function. The prediction accuracy of the ckeck function (1) can often be improved by selecting an active subsetof covariates. In this paper, to improve the prediction accuracy we consider thereciprocal LASSO quantile regression $(rLASSO - QR)$ which has not been proposed yet, that results from the following regularization problem:

$$\boldsymbol{min_\beta \sum_{i=1}^{n} \rho\theta\,(yi - \acute{x}_i\,\beta) + \lambda \sum_{j=1}^{p} \frac{1}{|\beta j|} I\,\{\beta i\ \neq 0\}} \quad (3)$$

where $I(.)$ denotes an indicator function and $\lambda > \mathbf{0}$ is the tuning parameter that controls the degree of penalization. The $\boldsymbol{rLASSO}$ penalty $\lambda \sum_j^p \frac{1}{|\beta j|} I\{\beta j \neq 0\}$ (Song and Liang, 2015), is decreasing in$(0\,,\infty)$, discontinuous at 0, and converge to ∞ when the regression parameters approach zero. It shares the same oracle property and same rate of estimator error with other LASSO-type penalties. Compared to traditional penalization functions (e.g., Lasso and SCAD) that are give nearly zero coefficients nearly zero penalties, the $\boldsymbol{rLASSO}$ penalty give nearly zero coefficients infinity penalties, which makes it very attractive for variable selection. In this paper, rather than minimizing (5), we solve the problem by constructing $\boldsymbol{S5 - QR}$ algorithm via a Gibbs sampler which involves constructing a Markov chain having the joint posterior for $\boldsymbol{\beta}$ as its stationary distribution.

2.2: Posterior inference

Since quantile regression does not equipped with a parametric likelihood, to proceed a Bayesian analysis we model the errors by the asymmetric Laplace distribution (ALD, Alhamzawi et al., 2012; Chen et al., 2013; Kozumi and Kobayashi, 2011; Yu and Moyeed, 2001). The density function of an ALD is

$$\boldsymbol{f(y\backslash\mu,\theta) = \theta(1-\theta)\exp\{-\rho\theta(y-\mu)\}} \qquad (4)$$

where, μ is a location parameter. In our model setup, the conditional distribution for the observations is

$$\boldsymbol{f(y\backslash\beta,\theta) = \theta^n\,(1-\ \theta^n)\exp\left\{\sum_{i=1}^{n}\rho\theta\,(yi-\ \acute{x}_i)\beta\right\}}. \qquad (5)$$

Maximizing the joint likelihood function over β is equivalent to minimizing the usual quantile check function $\sum_{i=1}^{n}\rho\theta(\,yi-\,\acute{x}_i\,\beta\,)$.However, direct use of this likelihood is rather unsuitable for posterior computation because the posterior distribution of β does not have a closed form. In this context, Kozumi and Kobayashi (2011) show

that ALD can be written as a location-scale mixture representation, i.e. :

$$\boldsymbol{\epsilon_i = (1-2\,\theta)wi + \sqrt{2\,w_i\,z_i}} \qquad (6)$$

where wi and zi are mutually independent, $wi\,\sim\,Exp\big(\theta(1\,-\,\theta)\big)$ and $zi\,\sim\,N(0;\,1)$. Marginally, the error distribution ϵ_i maintains its ALD form. However, conditional on the latent variable $w_i\epsilon_i$follows a normal distribution. Thus, posterior inference can be suitably carried out using Gibbs sampler. Following Mallick et al. (2019), a Bayesian solution for the minimization problem in (2) can be obtained by placing appropriate priors on the regression coefficients that will mimic the effects of the $rLASSO$ penalty. As apparent from (2), this choice of prior would be an independent inverse Laplace density on each of the coefficients

$$\pi\,(\beta) = \prod_{j=1}^{p} \frac{\lambda}{2\beta_j^2} \exp\left\{- \frac{\lambda}{|\,\beta_j\,|}\right\} I\,\{\beta_j \;\; \neq \;\; 0\}. \tag{7}$$

Hence, Gibbs sampling algorithm for $the\, rLASSO - QR$ is constructed by sampling $\boldsymbol{\beta}$ and $\boldsymbol{w} = (\boldsymbol{w}1; \cdots; \boldsymbol{w}n\;\; \grave{)}$ from their full conditional distributions. However, because no point mass at zero is assigned in this regularization problem, the samples of the regression parameters for the inactive set of covaiates would not be exactly zero. To overcome this problem, we propose an efficient Simplified Shotgun Stochastic Search with Screening in Quantile Regression $(S5 - QR)$ to explore the enormous model space.

2.3ModelSetup

To fix the terminology, let $k = \{k_1, \dots\dots.., k_{|k|}\}$ denote a model, where $1 \leq k_1 <, \dots.., < k_{|k|} \leq p$,, with $\beta_{kJ} \neq 0\, for\, J = 1, \dots.., |k|$ and all other components of $\boldsymbol{\beta}$ are 0. Let X_k and $\beta = \{\beta_k, \dots.., \beta_{k,|k|},\}$ are the design matrix and the regression coefficients of the model **k** only including the predictors with $\beta_{kJ} \neq 0$. Let $\boldsymbol{t}$ denote the true model and the cardinality of model $\boldsymbol{t}$ is denoted by$t = |t|$. Under each model $\boldsymbol{k}$, the sampling density for the observations is

$$\boldsymbol{y_n} \mid \boldsymbol{\beta_k} \sim \boldsymbol{N}\,(\boldsymbol{X_k}\,\boldsymbol{\beta_k} + (\mathbf{1} - \mathbf{2}\boldsymbol{\theta})\,\boldsymbol{w}\,, \boldsymbol{W_n}), \tag{8}$$

where $W_n = dig\,(2w_n, \dots\dots\dots, 2w_n)$.Given a model $\boldsymbol{k}$, the inverse Laplace prior on the

regression coefficients is defined as :

$$\boldsymbol{\pi}\,(\boldsymbol{\beta_k}|\boldsymbol{\lambda}, \boldsymbol{K}) = \prod_{j=1}^{|\boldsymbol{k}|} \frac{\boldsymbol{\lambda}}{\mathbf{2}\boldsymbol{\beta}_{\boldsymbol{k.j}}^{\mathbf{2}}} \mathbf{exp}\left\{- \frac{\boldsymbol{\lambda}}{|\boldsymbol{\beta}_{\boldsymbol{k},\boldsymbol{j}}|}\right\} \boldsymbol{I}\,\{\boldsymbol{\beta}_{\boldsymbol{k,j}} \neq \mathbf{0}\} \tag{9}$$

$$\neq \mathbf{0}\}$$

Following Shin et al. (2018), we put a uniform prior on the model space of the form $\pi\,(k) \propto I\,(|k| \leq q_n)\; with\; q_n < n\,,\; where\; I(.)$ denotes the indicator function. The basic idea in calculating the posterior probabilities of each model is to get the marginal likelihood of the observations $m_k\,(y)$under model $\boldsymbol{k}$ by integrating out the model parameters. Under model $\boldsymbol{k}$, the marginal likelihood of the observations $m_k\,(y)$ can be obtained by integrating out $\beta\boldsymbol{k}$, resulting in

$$\boldsymbol{m_k\,(y) = \,(4\pi w_i)\, Q_k^* \exp\{-R_k^*/\,4\}}, \qquad (10)$$

where

$$\boldsymbol{R_k^* = \acute{u}\left(I_n - WX_K\left(\acute{X}_k\, WX_k\right)\acute{X}_k\, W\right) u},$$

$$\boldsymbol{u = y - \,(1 - 2\theta\,)\omega}\,, \qquad (11)$$

$$\boldsymbol{Q_k^* \int \prod_{j=1}^{|k|} \quad \frac{\lambda}{2\beta_{k,j}^2} \exp\left\{-\,\beta_{k\,-} - \,\widehat{\beta}_k\;{}^{\prime}\;\; \Sigma_k^{*-1}\left(\beta_k\acute{-}\,\widehat{\beta}_k\,\right)/\,4 - \sum_{j=1}^{|k|} \frac{\lambda}{(\;\;|\beta_{k,j})}\right\}} \qquad (12)$$

$$\boldsymbol{\widehat{\beta}_k = (\,\acute{X}_k W X_k\,)^{-1}\,\acute{X}_k\,W_u\,, \Sigma_k^* = (\acute{X}_k\,WX_k\,)^{-1}} \qquad (13)$$

To estimate $\widehat{\beta}_k$, we assume that the size of $\widehat{\beta}_k$ is p and $\lambda \sim \lambda^{c-1}\exp(-d\lambda)$.. We follow the Gibbs sampler of Alhamzawi and Mallick (2020). This Gibbs sampler is described with some modifications in Algorithm 1.

3. Simplified Shotgun Stochastic Search with Screening in QR

Shin et al. (2018) proposed a simplified shotgun stochastic search with screening (S5) algorithm in an attempt to reduces the computing time of the SSS algorithm without losing the capacity to search the interesting region in the model space. They introduced \temperature parameter" to explore a broader spectrum of models. The Simplified Shotgun Stochastic Search with Screening (S5) algorithm (Shin et al., 2018)

Let $t_1 > t_2 > \cdots > 0$ is a set of temperature schedule and nbd $(k) = \{\Lambda^+, \Lambda^-, \Lambda^0\}, where\ \Lambda^+ = \{k \cup \{j\}: j \in k^c\}, \Lambda^- = \{k\ \{j\}: j \in k\}, and\ \Lambda^0 = [\{kJay\}] \cup \{j\}: J \in k^c, j \in k\}. This\ method\ is\ described\ in\ Algorithm\ (2).$

Algorithm 1 MCMC sampling for the Bayesian reciprocal

LASSO quantile regression (Alhamzawi and Mallick , 2020)

Input : (y, X)

Initialize : (β_k, w, u, λ)

For $t = 1, \dots \dots, (t_{max} + t_{burn-in})do$

1. Sample $w^{-1} \mid . \sim \prod_{i=1}^{n} inverse\ Gaussian\ \left(\frac{1}{2}, \frac{1}{|y_i - \acute{x}_i\, \beta_k}, \frac{1}{2}\right)$
2. Sample $u \mid . \sim \prod_{j=1}^{p} Exponential\ (\lambda) I\ \left\{u_j > \frac{1}{|\beta_j|}\right\}$
3. Sample $\beta_k \mid .$ from a truncated multivariate normal proportional to

   $N_p((\acute{X}\quad W^{-1} X)\, \acute{X}\, W^{-1}\, (y - \theta\, w), 2\, (\acute{X}\, W^{-1}\, \acute{X})\,)\prod_{j=1}^{p} I\ \left\{|\beta_j| > \frac{1}{u_j}\right\}$
4. Sample $\lambda \mid . \sim Gamma\ (\ c + 2p\ , d + \sum_{j=1}^{p} \frac{1}{|\beta_j\ |}\ )$

end for

Algorithm 2 Simplified Shotgun Stochastic Search with Screening (S5)

Input (y, X)

Initialize : $k^{(1,1)}$

Select: a set of predictors X^*corresponding to the initial model $K^{(1,1)}$

Select: a subset of predictors from X^* after the first screening step $S^{(1,1)}{}_k$

For J = 1 ,….., L do

For I = 1,……,(J – 1) do

1. Compute all $\pi\ (k \backslash y)$for all $k\ \in n\ bd_{scr}(k^{i,j})$
2. Update $k^+$from Λ^+_{scr} with probability proportional to $\pi\ (k\ , y\ )^{1\backslash 4}$

3. Update $k^{-}$from Λ^{-} with probability proportional to $\pi\ (k \backslash y\ )^{1\backslash ti}$

4. Update $k^{(i+1,j)}$from $\{k^{+}, k^{-}\}$with probability proportional to

$$\{\pi\ (k^{+}\ \ y\ )^{1\backslash tj}\ , \pi\ (k^{-}\ \ y\ )^{1\backslash ti}\}$$

5. Update $S_{k(i+1,i)}$ according to $\left| r^{\mathrm{T}} \quad {}_{k(i+1,i\,)}\, X_j \right| : j = 1\ , \ldots\ldots, p$

end for

After deriving the posterior distribution for the parameters as showed above, an efficient and easy (MCMC) algorithm for parameters of posterior distribution shown above . Our proposed algorithm is run for 13000 iterations. The first three thousand iterations was exclude as burn-in . In order to our algorithm more stationary

4. Simulation study

In this section, the performance of our method is studied by simulation scenarios via two examples . The our proposed method will be used to variables selection and coefficients estimation in quantile regression model which it is named (Quantile Regression with reciprocal LASSO penalty), denoted by (B R LASSO Q R). three criterions are used in this study are relative mean square error, denoted by (RMSE), $RMSE = \frac{\|X^t\hat{\beta} - X^t\beta^R\|^2}{\sigma}$,where β^R is true parameters , $\hat{\beta}$ is estimation parameters , σ is stander deviation of random error. Median of mean absolute deviations denoted by (MMAD).$MMAD = median(mean(|X^t\hat{\beta} - X^t\beta^r|)$. where β^R is true parameters , $\hat{\beta}$ is estimation parameters. Mean absolute error denoted by (MAE) . This proposed method is also compared with two other method Bayesian Lasso quantile Regression denoted by (BQReg N) and Bayesian new lasso quantile regression denoted by (**BQReg U**) . via five quantile levels ($\tau = 0.10, \tau = 0.20, \tau = 0.50, \tau = 0.70\ and\ \tau = 0.95$).

First example (sparse case)

In this example, the true parameters are $\beta^r = (0,1,0,0,2,0,1.5,0,1)^t$. Therefore , We generate data set from quantile regression model, as follow formula

$$y = x_{2i} + 2x_{5i} + 1.5x_{7i} + x_{9i} + \varepsilon_{\tau i} \qquad {}_{[i=1,\ldots\ldots,100]\,,} \qquad 0 < \tau < 1$$

where $\varepsilon_i \sim N(0, \sigma^2)$

We generate Nine explanatory variables from a multivariate normal with mean 0, and cov-variance $(x_i\,, x_j) = 0.5^{|i+j|}$.

The RMSE, MMAD,MAE and standard deviations (SD) are inserted in table 1. It is Clearly observed that via all the quantile levels under study. The our proposed method (B R LASSO Q R) generate smaller RMSE, MMAD,MAE and (SD) comparison to other method (BQRegU, BQReg N,). This mean the our proposed method is very accurately. it is very good method to achieving variables selection and coefficient estimation. In general , our proposed method has a good performance via all quantile level . In finally, the method (B R LASSO Q R) is not different its performance via all quantile level with sparse models.

Table 1: Show results of relative mean square error, denoted by (RMSE) , Median of mean absolute deviations (MMAD) and Mean absolute error MAE via averaged over 50 replications

Comparison Methods		RMSE	MMAD	MAE
$\tau = 0.10$	**BQReg U**	0.377 (0.245)	0.492 (0.323)	0.419 (0.346)
	BQReg N	0.347 (0.229)	0.342 (0.261)	0.403 (0.379)
	B R LASSO Q R	**0.252 (0.119)**	**0.173 (0.036)**	**0.119 (0.023)**
$\tau = 0.20$	**BQRegU**	0.378 (0.278)	0.369 (0.284)	0.333 (0.205)
	BQReg N	0.383 (0.465)	0.451 (0.312)	0.469 (0.315)
	B R LASSO Q R	**0.229 (0.122)**	**0.250 (0.124)**	**0.287 (0.141)**
$\tau = 0.50$	**BQRegU**	0.363 (0.293)	0.483 (0.302)	0.411 (0.396)
	BQReg N	0.482 (0.389)	0.472 (0.385)	0.414 (0.352)
	B R LASSO Q R	**0.232(0.111)**	**0.242(0.144)**	**0.236(0.141)**
$\tau = 0.70$	**BQRegU**	0.472 (0.347)	0.482 (0.384)	0.492 (0.382)
	BQReg N	0.445 (0.346)	0.345 (0.145)	0.487 (0.491)
	B R LASSO Q R	**0.255 (0.143)**	**0.248 (0.147)**	**0.334 (0.173)**
$\tau = 0.95$	**BQRegU**	0.505 (0.353)	0.484 (0.445)	0.445 (0.359)
	BQReg N	0.584 (0.454)	0.479 (0.303)	0.445 (0.345)
	B R LASSO Q R	**0.363 (0.135)**	**0.383 (0.273)**	**0.234(0.106)**

Note: In the parentheses are SDs of the MAD

First example (very sparse case)

In this example, the true parameters are $\beta^r = (0,1,0,0,0,0,0,0,0)^t$. Therefore , We generate data set from quantile regression model, as follow formula

$$\boldsymbol{y = x_{2i} + \varepsilon_i} \qquad [\boldsymbol{i=1,\ldots\ldots,100}] \ , \qquad \boldsymbol{0 < \tau < 1}$$

where $\varepsilon_i \sim N(0, \sigma^2)$

We generate Nine explanatory variables from a multivariate normal with mean 0, and cov-variance $(x_i , x_j) = 0.5^{|i+j|}$.
The RMSE, MMAD,MAE and standard deviations (SD) are inserted in table 2. It is Clearly observed that via all the quantile levels under study. The our proposed method (B R LASSO Q R) generate smaller RMSE, MMAD,MAE and (SD) comparison to other method (BQRegU, BQReg N,). This mean the our proposed method is very accurately. it is very good method to achieving variables selection and coefficient estimation. In general , our proposed method has a good performance via all quantile level . In finally, the method (B R LASSO Q R) is not different its performance via all quantile level with sparse models.

Table 2: Show results of relative mean square error, denoted by (RMSE) , Median of mean absolute deviations (MMAD) and Mean absolute error MAE via averaged over 50 replications

Comparison Methods		RMSE	MMAD	MAE
$\tau = 0.10$	**BQRegU**	0.576 (0.433)	0.602 (0.453)	0.547 (0.492)
	BQReg N	0.413 (0.359)	0.579 (0.483)	0.447 (0.310)
	B R LASSO Q R	**0.283 (0.173)**	**0.315 (0.276)**	**0.362 (0.152)**
$\tau = 0.20$	**BQRegU**	0.454 (0.371)	0.462 (0.383)	0.364 (0.278)
	BQReg N	0.484 (0.384)	0.478 (0.367)	0.594 (0.437)
	B R LASSO Q R	**0.253 (0.183)**	**0.261 (0.162)**	**0.248 (0.172)**
$\tau = 0.50$	**BQRegU**	0.424 (0.354)	0.407 (0.395)	0.423 (0.290)
	BQReg N	0.458 (0.383)	0.374 (0.206)	0.338 (0.207)
	B R LASSO Q R	**0.240 (0.183)**	**0.296 (0.137)**	**0.232 (0.142)**
$\tau = 0.70$	**BQRegU**	0.562 (0.483)	0.567 (0.482)	0.528 (0.392)
	BQReg N	0.482 (0.353)	0.452 (0.352)	0.434 (0.355)
	B R LASSO Q R	**0.317 (0.263)**	**0.253 (0.192)**	**0.226 (0.138)**
$\tau = 0.95$	**BQRegU**	0.456 (0.356)	0.446 (0.395)	0.473 (0.272)
	BQReg N	0.484 (0.359)	0.473 (0.384)	0.583 (0.396)
	B R LASSO Q R	**0.219 (0.092)**	**0.216 (0.172)**	**0.261 (0.162)**

Note: In the parentheses are SDs of the MAD

From figure 1,2 . We can see the quantile regression parameters estimates via our proposed method is very closed from normal distribution through histogram graphs. Also, it convergence to stationary this clearly from trace plot (at $\tau = 0.50$). This mean the MCMC sampler is easy and effective.

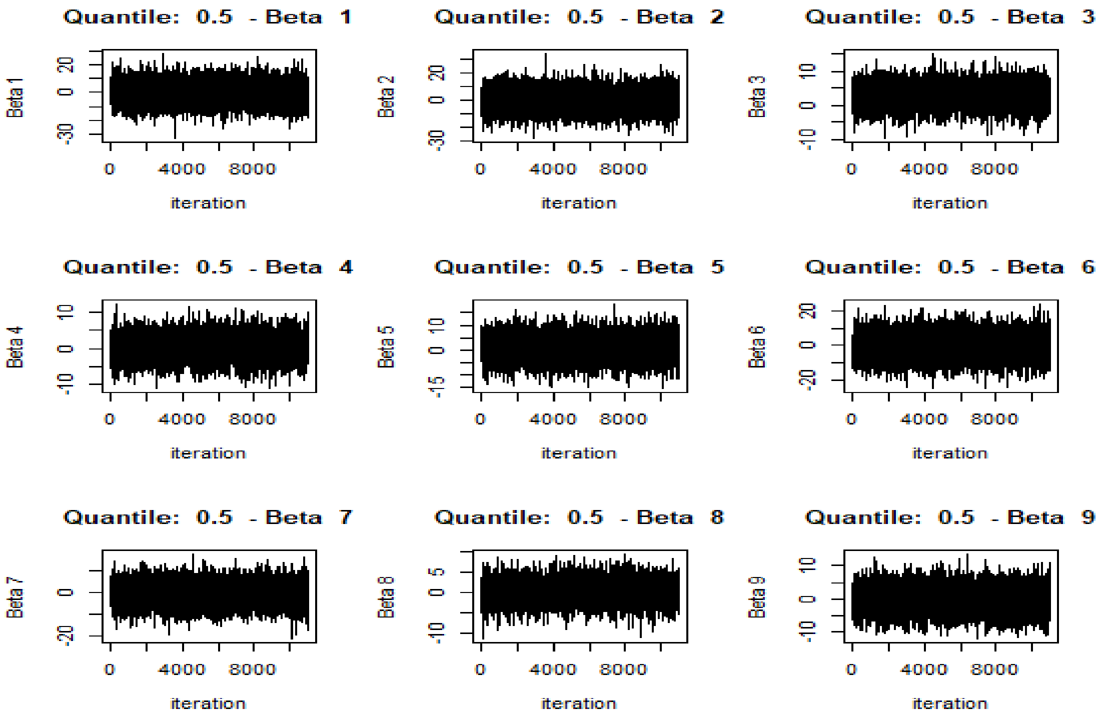

Figure 1. Trace plots of (B R LASSO Q R) with quantle ($\tau = 0.50$)

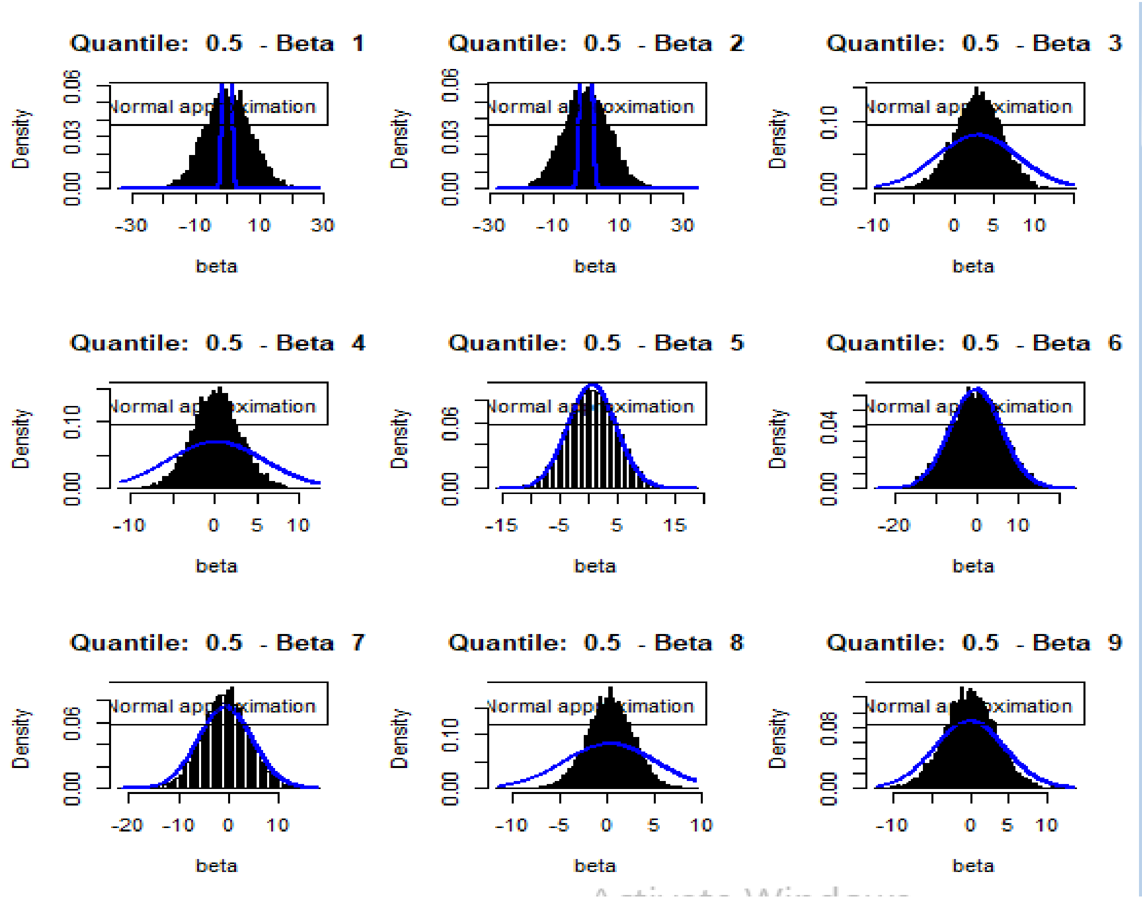

Figure 2 . Histogram of (B R LASSO Q R) parameter estimation with quantle ($\tau = 0.50$).

5 .Conclusion and Recommendation

5 . 1 Conclusion

In this paper, we will proposed good methods in quantile regression model by using Bayesian approach . (**B R LASSO Q R**) is focus by adding new contribution to achieving variables selection and coefficients estimation in quantile regression model with high efficiency . **B R LASSO Q R** has a good performance compared with other methods via all quantile levels , this clear through results of simulation and real data study.

5 . 2 Recommendation

We recommend the use of suggested extinction regularization hierarchical model with reciprocal LASSO penalty to Bayesian reciprocal LASSO Tobit regression model, Bayesian reciprocal LASSO Tobit quantile regression model Bayesian reciprocal LASSO binary regression model, and Bayesian reciprocal LASSO principal component regression model.

Bibliography

[1] Alhamzawi, R. and H. Mallick (2020). Bayesian reciprocal lasso quantile regression. Communications in Statistics-Simulation and Computation, 1{16.

[2] Alhamzawi, R. and K. Yu (2014). Bayesian lasso-mixed quantile regression. Journal of Statistical Computation and Simulation 84(4), 868{880.

[3] Alhamzawi, R., K. Yu, and D. F. Benoit (2012). Bayesian adaptive lasso quantile regression. Statistical Modelling 12(3), 279{297.

[4] Benoit, D. F., R. Alhamzawi, and K. Yu (2013). Bayesian lasso binary quantile regression. Computational Statistics 28(6), 2861{2873.

[5] Belloni, A., V. Chernozhukov, et al. (2011). ‘1-penalized quantile regression in high-dimensional sparse models. The Annals of Statistics 39(1), 82{130.

[6] Bradic, J., J. Fan, and W. Wang (2011). Penalized composite quasi-likelihood for ultrahigh dimensional variable selection. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 73(3), 325{349.

[7] Chen, C. W., D. B. Dunson, C. Reed, and K. Yu (2013). Bayesian variable selection in quantile regression. Statistics and its Interface 6(2), 261{274.

[8] Fan, J. and R. Li (2001). Variable selection via nonconcave penalized likelihood and its oracle properties. Journal of the American statistical Association 96 (456),1348{1360.

[9] Fan, J. and J. Lv (2008). Sure independence screening for ultrahigh dimensional feature space. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 70(5), 849{911.

[10] Hans, C. (2009). Bayesian lasso regression. Biometrika 96(4), 835{845.

[11] Hans, C., A. Dobra, and M. West (2007). Shotgun stochastic search for \large p" regression. Journal of the American Statistical Association 102 (478), 507{516.

[12] Johnson, V. and D. Rossell (2012). Bayesian model selection in high-dimensional settings. Journal of the American Statistical Association 107 (498), 649{660.

[13] Koenker, R. and G. Bassett (1978). Regression quantiles. Econometrica: Journal of the Econometric Society, 33{50.

[14] Kozumi, H. and G. Kobayashi (2011). Gibbs sampling methods for bayesian quantile regression. Journal of statistical computation and simulation 81 (11), 1565{1578.

[15] Li, Q., N. Lin, et al. (2010). The Bayesian elastic net. Bayesian analysis 5(1),151{170.

[16] Lamarche, C. (2010). Robust penalized quantile regression estimation for panel data. Journal of Econometrics 157(2), 396{408.

[17] Mallick, H., R. Alhamzawi, and V. Svetnik (2019). The reciprocal Bayesian lasso.

[18] Park, T., & Casella, G. (2008). The bayesian lasso. Journal of the American Statistical Association, 103(482), 681-686.

[19] Rossell, D., D. Telesca, and V. Johnson (2013). Statistical models for data analysis..

[20] Shin, M., A. Bhattacharya, and V. E. Johnson (2018). Scalable bayesian variable selection using nonlocal prior densities in ultrahigh-dimensional settings. Statistica Sinica 28(2), 1053.

[21] Song, Q. and F. Liang (2015). High-dimensional variable selection with reciprocal l1-regularization. Journal of the American Statistical Association 110 (512), 1607{ 1620.

[22] Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society: Series B (Methodological) 58 (1), 267{288.

[23] Zou, H. (2006). The adaptive lasso and its oracle properties. Journal of the American statistical association 101 (476), 1418{1429.

[24] Zou, H. and T. Hastie (2005). Regularization and variable selection via the elastic net. Journal of the royal statistical society: series B (statistical methodology) 67(2), 301{320.

[25] Zou, H. and H. H. Zhang (2009). On the adaptive elastic-net with a diverging number of parameters. Annals of statistics 37 (4), 1733

[26] Zheng, Q., C. Gallagher, and K. Kulasekera (2013). Adaptive penalized quantile regression for high dimensional data. Journal of Statistical Planning and Inference 143(6), 1029{1038.