American Scientific Research Journal for Engineering, Technology, and Sciences (ASRJETS)
ISSN (Print) 2313-4410, ISSN (Online) 2313-4402

http://asrjetsjournal.org/

New Bayesian Lasso Composite Quantile Regression

Fadel Hamid Hadi Alhusseini*

Department of Statistics and Economic Informatics, University of Craiova, Romania University of Al-Qadiseya , Iraq

Email: fadhelfadhel222@yahoo.com

Abstract

In this paper, we propose a new Bayesian lasso inference scheme for variable selection in composite quantile regression model (C Quantile Reg). The suggested approach is to construct a hierarchical structure within the Gibbs sampling under the assumption that the residual term comes from skew Laplace distribution (asymmetric Laplace distribution) and assign scale mixture uniform (SMU) as prior distributions on the coefficients of composite quantile regression model. Our proposed method was compared to some other existing methods by testing the performance of these methods through simulation studies and real data examples.

Keywords: New Bayesian Lasso; Posterior distributions; composite Quantile regression; Scale mixture of uniform.

1. Introduction

Since the seminal pioneering work of [11], quantile regression (Q Reg) has become more and more popular in numerous fields of science, for instance, the microarray study [25], agricultural economics [10], ecological studies [8], growth chart [26], etc. In addition, the quantile regression has good properties and it is a suitable regression model to non-normal errors, as quantile regression is more robust compared to ordinary least-square (OLS) regression [12]. QReg is more flexible, thus, it offers an extensive coverage of the response variable and its covariates correlations QReg operates without an assumption about the random error, providing greater statistical efficiency than traditional regression models when the error is non-normal. So far, QReg has proven to be an enhanced model of the conventional regression model.

----

* Corresponding author.

The linear quantile regression model assumes that the response variable $\boldsymbol{y_i}$ written as

$$y_i = \alpha_\theta + x_i^T\beta_\theta + u_i \qquad i = 1,2,\dots,n \qquad (1)$$

where y_i $[i = 1,2,\dots..n]$ is response variable

θ is level quantile, and θ belongs to the open interval $(0,1)$. $\theta \in (0,1)$

α_θ is the intercept.

β_θ is vector of unknown parameters

x_i^T is vector of covariates

u_i is the random error term with θ-th quantile equal to zero

The parameters α_θ and β_θ to the quantile regression model are estimated by solving the following equation:

$$\min_{\alpha_\theta,\beta_\theta} \sum_{i=1}^{n} \rho_\theta\,(y_i - \alpha_\theta - x_i^T\beta_\theta) \qquad (2)$$

where $\rho_\theta(u)$ is the check (loss) function

$$\rho_\theta(u) = \begin{cases} \theta u & if\ u \geq 0 \\ -(1-\theta)u & if\ u < 0 \end{cases} \qquad (3)$$

Since equation (2) is not differentiable at the origin, there is no exact form of the solution for equation (2) [15]. The minimization of equation (2) can be achieved by a linear programming algorithm [13]. A Bayesian approach enables the exact estimation of quantile regression parameters. One important thing in building a regression model is the selection of the active covariates. The selection process aims to increase the prediction accuracy and to get high interpretation [4]. Recently, there has been considerable attention on regression models that include all covariates and use informative priors to shrink inactive regression coefficients toward zero exactly, for instance, the Lasso method [23], the adaptive Lasso, proposed by [26], dantzig selector [7], Lasso QReg [17] and adaptive Lasso QReg [28]. A comprehensive account of these and other recent methods can be found in the work of [24]. Similarly, from a Bayesian framework, [20] proposed Bayesian Lasso for traditional linear regression models by specifying scale mixture of normal (SMN) prior distributions. Reference [21] proposed Bayesian adaptive Lasso by using different shrinkage parameters. In later approaches, Reference [18] suggested Bayesian Lasso QReg and [4] proposed Bayesian adaptive Lasso Qreg. Some further extensions of the Lasso Qreg models have also been suggested by [5,1,27,29,2], among others. All aforementioned methods were focused on a single quantile level and scale mixture of normal (SMN) priors. [30,6].; proposed methods that focused on composite quantile regression model. [9] proposed a Bayesian approach of composite quantile regression by using scale mixture of normal (SMN) priors. In this paper our proposed method develops a

Bayesian approach for regularization in linear composite QReg by assigning the scale mixture of uniform (SMU) formulation of the Laplace density. This approach was provided by [19]. In other word, in this paper, we propose a new formulation for Bayesian lasso composite QReg by using the scale mixture of uniform (SMU) formulation. The following sections of this paper are arranged in the following order: section 2 presents New Bayesian lasso composite QReg. In section 3, we perform simulation studies. In section 4, we show a real data example, and the conclusions and recommendations are included in section 5.

2. New Bayesian Lasso Composite QReg

2.1. Bayesian Composite QReg

The quantile regression model is focused on modeling the relationship between the response variable and a set of covariates by single quantile level. From known, there are infinite quantile regression lines within $\theta \in (0,1)$ or $0 < \theta_1, \theta_2, \ldots\ldots \theta_H < 1$ where H represents different quantiles. But choosing the informative quantile line is challenging and to overcome this problem, the information is used over all quantile lines. [30] introduced the mathematical formula to composite quantile regression model, as follows:

$$\alpha_{\theta_1}, \alpha_{\theta_2}, \ldots\ldots, \alpha_{\theta_h}, \beta = \underset{\alpha_{\theta_1}, \alpha_{\theta_2}, \ldots\ldots, \alpha_{\theta_h}}{Min} \sum_{h=1}^{H} \sum_{i=1}^{n} \rho_{\theta h}(y_i - \alpha_h - x_i^T \beta) \qquad (4)$$

where

$h = 1,2, \ldots., H \; and \; i = 1,2, \ldots\ldots, n$

Hence, the check(loss) function of composite different quantiles takes the following form:

$$\rho_{\theta h}(u) = \begin{cases} \theta_h u & if \; u \geq 0 \\ -(1-\theta_h) u & if \; u < 0 \end{cases} \qquad (5)$$

The equation (4) is also not differentiable at 0, and the minimization of (4) can be achieved by a linear programming algorithm [14]. It is possible to estimate the parameters of composite quantile regression by using a Bayesian approach through a new Gibbs sampler and this is our proposed method to posterior distributions. [28] noted that the random error of quantile is close to asymmetric Laplace distribution (ALD).

$$f_{ALD}(u|\theta_h) = \frac{\theta_h(1-\theta_h)}{\sigma} \exp\left(\sigma^{-1} \rho_{\theta h}(u)\right) \qquad -\infty < x < \infty \qquad (6)$$

The function of asymmetric Laplace distribution with a scale parameter equal to 1 is

$$f_{ALD}(u|\sigma,\theta_h) = \theta_h(1-\theta_h)\ \exp\big(\rho_{\theta h}(u)\big) \qquad -\infty < x < \infty \qquad (7)$$

With mean $E(u) = \frac{(1-2\theta_h)}{\theta_h(1-\theta_h)}$ and the variance $var\ (u) = \frac{(1-2\theta_h - 2\theta_h^2)}{\theta_h(1-\theta_h)}$

The joint distribution of response variable $= (y_1, y_2, \dots\dots, y_n)^T$, given $X = (x_1, x_2, \dots., x_n)^T$, $\alpha = \left(\alpha_{\theta_1}, \alpha_{\theta_2}, \dots\dots, \alpha_{\theta_h}\right)^T$ and $\beta = (\beta_1, \beta_2, \dots\dots\beta_k)^T$ for composite quantile regression is:

$$f_{ALD}(y|X,\alpha,\beta) = \prod_{h=1}^{H} \theta_h^n(1-\theta_h^n)\ \exp\left(\sum_{i=1}^{n} \rho_{\theta h}(y-\alpha_h - x_i^T\beta)\right) \qquad (8)$$

Hence, maximizing the likelihood function of response variable y in equation (8) is equivalent to minimizing equation (4). It is very difficult to use the likelihood function of response variable y directly, because of the mixture of H components. Following [9], we employ a cluster assignment matrix C with its $(i,h)^{th}$ element, C_{ih} taking two values: one ($C_{ih} = 1$) if ith subject belongs to the $h\ th$ cluster, or zero ($C_{ih} = 0$) if the element C_{ih} is treated like missing value. Therefore, our likelihood takes the following form:

$$f_{ALD}(y|X,\alpha,\beta) = \prod_{h=1}^{H}\prod_{i=1}^{n}\left[\theta_h(1-\theta_h)\ \exp\Big(\rho_{\theta h}(y-\alpha_h - x_i^T\beta)\Big)\right]^{C_{ih}} \qquad (9)$$

Recently, [16] reformulated the distribution of the (ALD) as a mixture of normal distributions (SMN). More specifically, the random error is equal to $u = \zeta_{1h}z_i + \zeta_{2h}\sqrt{z_i}\,\epsilon$, where $z_i \sim Exp[\theta_h(1-\theta_h)]$ and ϵ is distributed standard normal distribution, $\epsilon \sim N(0,1)$. Therefore, the random error u is distributed normal distribution with mean $((1-2\theta_h)z_i$ and variance $(2z_i)$, $u \sim N(\big((1-2\theta_h)z_i, 2z_i\big)$. As a result, the response variable y is distributed normal distribution with mean $(\alpha_h + x_i^T\beta + (1-2\theta_h)z_i)$ and variance $(2z_i)$.

,$y_i|\alpha_h,\beta,z_i \sim N(\alpha_h + x_i^T\beta + (1-2\theta_h)z_i, 2z_i)$. According to the formulation of [16], our Bayesian composite quantile regression can be written as:

$$f(y|X,z_i,\beta,\alpha_h,C) = \prod_{i=1}^{n}\left[\frac{1}{\sqrt{4\pi z_i}}\right]^{C_{ih}} exp\left\{-\frac{1}{2}\sum_{i=1}^{n}\sum_{h=1}^{H}\frac{C_{ih}(y-\alpha_h - x_i^T\beta - (1-2\theta_h)z_i)}{2z_i}\right\} \qquad (10)$$

Under the of formulation equation (10), the composite quantile regression coefficients have good features for constructing a good and efficient Gibbs sampler algorithm.

2.2. New Bayesian Lasso Composite QReg (new B C QReg)

[23] proposed a new method of estimation and variables selection in linear regression model, known as lasso (Least Absolute Shrinkage and Selection Operator) . In the same context [23] mentioned Bayesian lasso in linear regression model is possible under assigned Laplace distribution as priors for the β_j. [12] proposed lasso

quantile regression that takes the following formula:

$$\min_{,\beta_\theta} \sum_{i=1}^{n} \rho_\theta \left(y_i - \alpha_\theta - x_i^T \beta\right) + \lambda \left||\beta|\right| \quad (11)$$

where $\lambda(\lambda \geq 0)$ is the tuning parameter.

[18] suggested Bayesian Lasso for linear quantile regression model by assigning Laplace prior for the β_j.The probability density function takes the specification of the form:

$$p(\beta_j|\lambda) = \frac{\lambda}{2}\, exp[-\lambda|\beta|] \quad (12)$$

It is difficult to use equation (12) directly in the estimation of the model parameters. Therefore, most researchers in the field of Bayesian lasso regression models use the transformation Laplace distribution to the scale mixture normal. The works of [20]; [17]; [4] and [9] are relevant. In this paper, we develop an alternative hierarchical Bayesian model by using a new lasso technique. According to the proposal of [19] the Laplace prior distribution on coefficients (*βj*) can be written as:

$$\frac{\lambda}{2}\, exp[-\lambda|\beta|] = \int_{s_j > |\beta_j|} \frac{1}{2s_j} \frac{\lambda^2}{\Gamma(2)}\, {s_j}^{2-1}\, exp\{-\lambda s_j\}\, ds_j \quad (13)$$

$$= \int_{s_j > |\beta_j|} \frac{\lambda^2}{2}\; exp\{-\lambda s_j\}\, ds_j$$

where $\Gamma(2) = (2-1)! = 1$

$$= \frac{\lambda^2}{2} \left[\frac{exp\{-\lambda s_j\}}{-\lambda}\right]_{|\beta_j|}^{\infty} = \frac{\lambda^2}{2}\left[\frac{exp\{-\lambda\infty\}}{-\lambda} + \frac{exp\{-\lambda|\beta|\}}{-\lambda}\right] = \frac{\lambda^2}{2}\left[0 + \frac{exp\{-\lambda|\beta|\}}{-\lambda}\right] = \frac{\lambda}{2}\, exp[-\lambda|\beta|]$$

where *sj* is a mixing variable. Our Bayesian hierarchical model can be formulated as below after assigning Gamma priors on λ.

$$f(y|X, z_i, \beta, \alpha_h, C) = \prod_{i=1}^{n} \left[\frac{1}{\sqrt{4\pi z_i}}\right]^{C_{ih}} exp\left\{-\frac{1}{2}\sum_{i=1}^{n}\sum_{h=1}^{H} \frac{C_{ih}(y - \alpha_h - x_i^T\beta - (1-2\theta_h)z_i)}{2z_i}\right\}$$

$$p(\alpha_h) \propto 1$$

$$z_i \sim exp\{\theta_h(1-\theta_h)\},$$

$$\beta_j | s_j \sim Uniform\left(-s_j, s_j\right),$$

$$s_j | \lambda \sim Gamma(2, \lambda)\,, \tag{14}$$

$$\lambda \sim \lambda^{a-1} exp(-b\lambda)\,.$$

where a, b are hyperparameter.

2.2.1. Posterior Computation Inference

The Bayesian theory is used for the conditional posterior distribution for the parameters of the model by combining two different sources of information.

The first source is the likelihood function of response variable y in equation (10) and the second source is a set of prior distributions to model parameters in equation (14).

The Bayesian hierarchical in (10) and (14) produces posterior distributions $\left[\alpha_h, \beta, z = (z_1, z_2, \dots, z_n)^T, s = (s_1, s_2, \dots, s_p)^T, \lambda \ and\ C\right]$. A simple and efficient Gibbs sampler algorithm is as follows:

- Updating z_i -the full conditional distribution of each z_i foe $i = 1,2, \dots \dots n,\ InvG(\delta_i^T, \varphi_i^T)$, where $\varphi^T = \frac{\sum_{h=1}^{H} c_{ih}}{2}$ and

$$\delta_i^T = \sqrt{\frac{\frac{\sum_{h=1}^{H} c_{ih}}{2}}{\frac{\sum_{h=1}^{H} c_{ih} \left(y_i - \alpha_h - x_i^T \beta\right)^2}{2}}} = \sqrt{\sum_{h=1}^{H} c_{ih} \Big/ \sum_{h=1}^{H} c_{ih} \left(y_i - \alpha_h - x_i^T \beta\right)^2}$$

- Updating α_h - the full conditional distribution of each α_h foe $h = 1,2, \dots \dots H,\ \ is\ N(\tilde{\alpha}_h, \hat{\sigma}_h^2)$,

$$\hat{\alpha}_h = \left(\hat{\sigma}_h^2 \sum_{i=1}^{n} c_{ih} \left[y_i - x_i^T \beta - (1 - 2\theta_h) z_i\right) \Big/ 2z_i \right),\ \hat{\sigma}_h^2 = \left[\frac{\sum_{i=1}^{n} c_{ih}}{2z_i}\right]^{-1}$$

- Updating β_j - the full conditional distribution of each β_j for $j = 1,2, \dots \dots k,\ \ is\ N(\tilde{\beta}_j, \hat{\sigma}_j^2)$, where:

$$\tilde{\beta}_j = \left(\tilde{\sigma}_j^2 \sum_{i=1}^{n} \sum_{h=1}^{H} c_{ih} x_{ij} \left(T_i - \alpha_h - (1 - 2\theta_h) z_i - \sum_{l \neq j} x_{ij}\, \beta_j\right) \Big/ 2z_i \right) I\{|\beta_j| \le s_j\}$$

$$\text{and } \tilde{\sigma}_j^2 = \left(\frac{\sum_{i=1}^{n} \sum_{h=1}^{H} c_{ih} x_{ij}^2}{2z_i} + s_j^{-1}\right)^{-1}.$$

Updating s_j -the full conditional distribution of s_j is a left-truncated exponential distribution given by

$$s_j \sim Exp(\lambda) I\{s_j > |\beta_j|\} \text{ . for } j = 1,2, \ldots \ldots k,.$$

Updating s_j can be done by using inversion method [19] as follows:

1-Update $s_j{}^*$ from $Exp(\lambda)$

2- Set $s_j = s_j{}^* + |\beta_j|$.

Updating λ_j - the full conditional distribution of each λ_j for $j = 1,2, \ldots \ldots k,$.is Gamma $\left(a + 2k, b + \sum_{j=1}^{k} |\beta_j|\right)$.

Updating C - the full conditional distribution of each $C_i = (C_{i1}, C_{i2}, \quad , C_{ih})^T$ is multinomial distribution with

$$\hat{p}_h = \frac{\exp\left(-\left(y_i - \alpha_h - x_i^T \beta - (1-2\theta_h) z_i\right)^2 \Big/ 4z_i \right)}{\sum_{m=1}^{M} \exp\left(-\left(y_i - \alpha_h - x_i^T \beta - (1-2\theta_h) z_i\right)^2 \Big/ 4z_i \right)}$$

After updating the model parameters as illustrated above, an efficient Gibbs sampler results for building (MCMC) algorithm for parameters estimation $[\alpha_h, \beta, z, s, \lambda \text{ and } C]$. Our algorithm is run for 11000 iterations. The first 1000 iterations were discarded as burn-in.

3. Simulation studies

In this section, the behavior of our proposed method is assessed by simulation studies. To evaluate our method, a comparison is performed, using a set of methods: Bayesian Lasso quantile regression proposed by [18] and classical frequentist approach using the R function rq() in the R package quantreg.

The methods are assessed based on two criteria - median of mean absolute deviations (MMAD), where MMAD $=$ median mean $\left(\left|x_i^T \beta^{Estimated} - x_i^T \beta^{true}\right|\right)$, and standard deviation (SD) of the MADs. In this approach, we chose a set $H = 3$ (three quantile levels, as follows: $\theta_1 = 0.25$ is the low quantile level, $\theta_2 = 0.55$ is the middle quantile level and $\theta_3 = 0.85$ is the high quantile level) for each simulation study.

The error terms generated from four distributions are standard normal distribution $N(0,1)$, $\chi^2_{(3)}$distribution with three degrees of freedom mixture normal distribution $u_i \sim \frac{1}{2} N\ (0,1) + \frac{1}{2} N(1,1)$ and mixture Laplace distribution $u_i \sim \frac{1}{2} Laplace\ (0,1) + \frac{1}{2} Laplace(1,1)$.

Our algorithm is run 11000 iterations and the first 1000 are discarded as the burn-in. To assess the performance of our proposed method, we used three approaches.

3.1 The First Approach

In this section, the simulated study is done with a very sparse case. This is clear from true parameters values, $\beta = (1,0,0,0,0,0,0,0)^T$. Therefore, the true model becomes as follows:

$$y_i = 0 + x_{1i} + u_i, \qquad i = 1,2, \dots \dots .100$$

We simulate eight covariates ($x_{1i}, x_{2i}, x_{3i}, x_{4i}, x_{5i}, x_{6i}, x_{7i}, x_{8i}$) from a multivariate normal distribution $X \sim N_8(\mu, \Sigma)$, where μ is mean vector $\mu \epsilon R^n$ and Σ is covariance matrix with $(\Sigma_x)_{ij} = (2^{-1})^{|i-j|}$.

3.2 The Second Approach

The simulated study is done with a very dense case. This is obvious from true parameters values $\beta = (0.85, 0.85, 0.85, 0.85, 0.85, 0.85, 0.85, 0.85)^T$.

Therefore, the true model becomes as follows:

$$y_i = 0.85x_{1i} + 0.85x_{1i} + 0.85x_{2i} + 0.85x_{3i} + 0.85x_{4i} + 0.85x_{5i} + 0.85x_{6i} + 0.85x_{7i} + 0.85x_{7i} + u_i,$$

$$i = 1,2, \dots \dots .100$$

We simulated eight covariates (x_{1i}, x_{2i}, x_{3i}, x_{4i}, x_{5i}, x_{6i}, x_{7i}, x_{8i}) from a multivariate normal distribution $X \sim N_8(\mu, \Sigma)$, where μ is mean vector $\mu \epsilon R^n$ and Σ is covariance matrix with $(\Sigma_x)_{ij} = (2^{-1})^{|i-j|}$.

3.3 The Third Approach

The simulated study is done with group structures. This is obvious from true parameters values $\beta = [(0,0,0), (2,2,2), (0,0,0), (2,2,2), (0,0,0)]^T$. Therefore, the true model takes the formula below:

$$y_i = 2x_{4i} + 2x_{5i} + 2x_{6i} + 2x_{10i} + 2x_{11i} + 2x_{12i} + u_i, \qquad i = 1,2, \dots \dots .100$$

where the covariates are simulated.

$x_{1i}, x_{2i}, x_{3i}, x_{4i}, x_{5i}, x_{6i}, x_{7i}, x_{8i,}x_{9i,}x_{10i,}x_{11i,}x_{12i,}x_{13i,}x_{14i,}x_{15i}$) according to a multivariate Gaussian distribution $X \sim N_{15}(\mu, \Sigma)$, where μ is mean vector $\mu \epsilon R^n$ and Σ is covariance matrix with $(\Sigma_x)_{ij} = (2^{-1})^{|i-j|}$.

Table 1 shows the results of (MMAD) and (SD) averaged over 100 independent simulations for the three methods under comparison, via three quantile levels. The method proposed (the New Lasso CQ Reg) displays better results than the rq and $Lasso\ N$ methods, as can be seen in Table 1.

In case of MMAD, for all considered distributions, the values are much smaller when New Lasso CQ Reg in employed, compared to rq and $Lasso\ N$ methods.

In case of SD, the values are also smaller than for the other two methods, for different distributions. Taking into consideration these aspects, the New Lasso CQ Reg has superior performance compared to the other two methods under test.

Table 1: Summary of MMAD and SD values. The results are averaged over one hundred simulated data set for the methods under comparison.

	Error distributions			
Methods	$N(0,1)$	$\chi^2_{(3)}$	Normal mixture	Laplace mixture
First approach				
$rq\theta_1 = 0.25$	**1.929** (0.870)	**2.776**(0.345)	**2.513** (0.663)	**2.067**(1.013)
$rq\theta_2 = 0.55$	**2.160**(0.588)	**3.154**(0.264)	**2.614** (0.566)	**2.218**(0.891)
$rq\theta_3 = 0.85$	**2.513**(0.661)	**3.744**(0.428)	**2.959** (0.510)	**2.678**(0.772)
$Lasso\ N\theta_1 = 0.25$	**1.411**(0.599)	**2.113**(0.343)	**1.847** (0.483)	**1.476**(0.821)
$Lasso\ N\theta_1 = 0.55$	**1.932**(0.606)	**2.933**(0.375)	**2.434** (0.474)	**2.031**(0.781)
$Lasso\ N\theta_1 = 0.85$	**2.415** (0.626)	**3.653**(0.414)	**2.889** (0.442)	**2.483** (0.758)
New Lasso CQ Reg	**0.902(0.200)**	**0.560(0.174)**	**0.715 (0.215)**	**0.741(0.310)**
Second approach				
$rq\theta_1 = 0.25$	**1.698** (0.171)	**2.520**(0.338)	**2.018** (0.384)	**2.007**(0.156)
$rq\theta_2 = 0.55$	**1.959** (0.258)	**2.987** (0.409)	**2.375** (0.295)	**2.329** (0.161)
$rq\theta_3 = 0.85$	**2.256** (0.268)	**3.681** (0.948)	**2.724** (0.432)	**2.741** (0.301)
$Lasso\ N\theta_1 = 0.25$	**1.264** (0.216)	**1.858** (0.932)	**1.546** (0.337)	**1.431** (0.186)
$Lasso\ N\theta_1 = 0.55$	**1.736** (0.258)	**2.754** (0.287)	**2.104** (0.318)	**2.057** (0.235)
$Lasso\ N\theta_1 = 0.85$	**2.142** (0.309)	**3.560** (0.437)	**2.592** (0.408)	**2.647** (0.331)
New Lasso CQ Reg	**0.765 (0.106)**	**0.703 (0.122)**	**0.621 (0.124)**	**0.596 (0.331)**
Third approach				
$rq\theta_1 = 0.25$	**0.930** (0.841)	**3.446** (0.410)	**2.752**(1.071)	**2.813**(0.689)
$rq\theta_2 = 0.55$	**0.876** (0.667)	**4.005** (0.425)	**3.200** (1.020)	**3.209**(0.508)
$rq\theta_3 = 0.85$	**0.837** (0.777)	**4.845** (0.504)	**3.927** (0.758)	**3.777**(0.631)
$Lasso\ N\theta_1 = 0.25$	**0.851**(0.769)	**2.623** (0.357)	**2.090** (1.020)	**2.221**(0.522)
$Lasso\ N\theta_1 = 0.55$	**0.885**(0.704)	**3.594** (0.361)	**2.923** (0.959)	**2.981**(0.551)
$Lasso\ N\theta_1 = 0.85$	**0.869**(0.686)	**4.707** (0.411)	**3.851** (0.814)	**3.719**(0.683)
New Lasso CQ Reg	**0.373**(0.420)	**1.089** (0.307)	**1.496** (0.268)	**1.157**(0.352)

The standard deviations of the MAD are mentioned in parentheses.

The Table 2 provides additional criteria that represent the parameters estimation in direct way. As can be observed in the table, the parameters estimation through our proposed method (New Lasso CQ Reg) was very close to true parameters comparison with rq and $Lasso\ N$ methods.

Therefore, the proposed method (New Lasso CQ Reg) has performed better than rq and $Lasso\ N$ methods.

Table 2: Posterior means for simulation of the first approach via three quantile levels and four error distribution; the results are averaged over 100 simulations

Error distribution	**Method**	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$	$\hat{\beta}_6$	$\hat{\beta}_7$	$\hat{\beta}_8$
$N(0,1)$	**True parameters**	**1**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
	$rq\theta_1 = 0.25$	-0.882	1.807	-0.384	-0.157	0.878	0.192	0.237	0.340
	$rq\theta_2 = 0.55$	-1.035	2.320	-0.717	-0.613	1.392	0.043	0.253	0.331
	$rq\theta_3 = 0.85$	-2.135	3.572	-1.852	-0.402	1.707	0.130	0.377	0.114
	$Lasso\ N\theta_1 = 0.25$	-0.686	1.790	-0.481	0.105	0.978	0.238	0.475	0.366
	$Lasso\ N\theta_1 = 0.55$	-1.301	2.535	-1.061	-0.184	1.414	0.116	0.444	0.387
	$Lasso\ N\theta_1 = 0.85$	-2.230	3.333	-1.952	-0.186	1.642	0.374	0.126	0.464
	New Lasso CQ Reg	**0.970**	**0.017**	**0.002**	**0.121**	**0.009**	**0.003**	**0.044**	**0.145**
$\chi^2_{(3)}$	**True parameters**	**1**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
	$rq\theta_1 = 0.25$	1.377	0.465	0.469	-0.784	1.954	1.366	-0.135	0.361
	$rq\theta_2 = 0.55$	1.358	0.096	1.747	-2.080	1.971	1.936	-0.207	1.341
	$rq\theta_3 = 0.85$	0.148	0.129	2.381	-1.942	3.008	2.265	-0.887	1.119
	$Lasso\ N\theta_1 = 0.25$	2.105	-0.021	0.992	-1.269	1.646	1.081	-0.307	0.662
	$Lasso\ N\theta_1 = 0.55$	1.383	0.009	1.663	-1.821	2.398	1.724	-0.637	0.987
	$Lasso\ N\theta_1 = 0.85$	0.198	0.322	2.366	-2.179	2.764	2.410	-1.041	1.075
	New Lasso CQ Reg	**1.242**	**0.065**	**0.285**	**0.342**	**0.183**	**0.283**	**0.218**	**0.327**
Mixture Normal	**True parameters**	**1**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
	$rq\theta_1 = 0.25$	1.275	0.288	0.290	-0.137	0.752	-0.234	1.014	-0.769
	$rq\theta_2 = 0.55$	2.012	0.682	-0.410	0.283	1.378	-0.732	1.921	-1.356
	$rq\theta_3 = 0.85$	1.539	1.335	-0.182	-0.472	1.582	-1.698	3.037	-1.573
	$Lasso\ N\theta_1 = 0.25$	2.489	0.631	-0.109	-0.083	0.856	-0.755	1.353	-0.901
	$Lasso\ N\theta_1 = 0.55$	2.116	0.885	-0.346	-0.036	1.271	-1.055	2.029	-1.186
	$Lasso\ N\theta_1 = 0.85$	1.715	1.286	-0.348	-0.234	1.579	-1.448	2.762	-1.413
	New Lasso CQ Reg	**1.244**	**0.121**	**0.026**	**0.238**	**0.131**	**0.160**	**0.234**	**0.070**
Mixture Laplace	**True parameters**	**1**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
	$rq\theta_1 = 0.25$	5.050	0.120	-1.650	1.393	0.807	0.301	-1.553	0.523
	$rq\theta_2 = 0.55$	5.046	-0.100	-1.558	2.040	0.929	-0.172	-1.443	0.450
	$rq\theta_3 = 0.85$	5.554	-0.689	-0.967	2.502	0.951	-0.694	-1.950	0.669
	$Lasso\ N\theta_1 = 0.25$	5.265	-0.373	-0.609	1.167	-0.146	-0.971	-0.971	0.393
	$Lasso\ N\theta_1 = 0.55$	5.285	-0.584	-0.924	1.858	0.740	-0.288	-1.447	0.489
	$Lasso\ N\theta_1 = 0.85$	5.346	-0.846	-0.999	2.426	0.943	-0.608	-1.941	0.618
	New B Lasso CQ Reg	**0.159**	**0.110**	**0.464**	**0.322**	**0.037**	**0.191**	**0.350**	**0.120**

Figure 1 provides a clear vision about table 2.

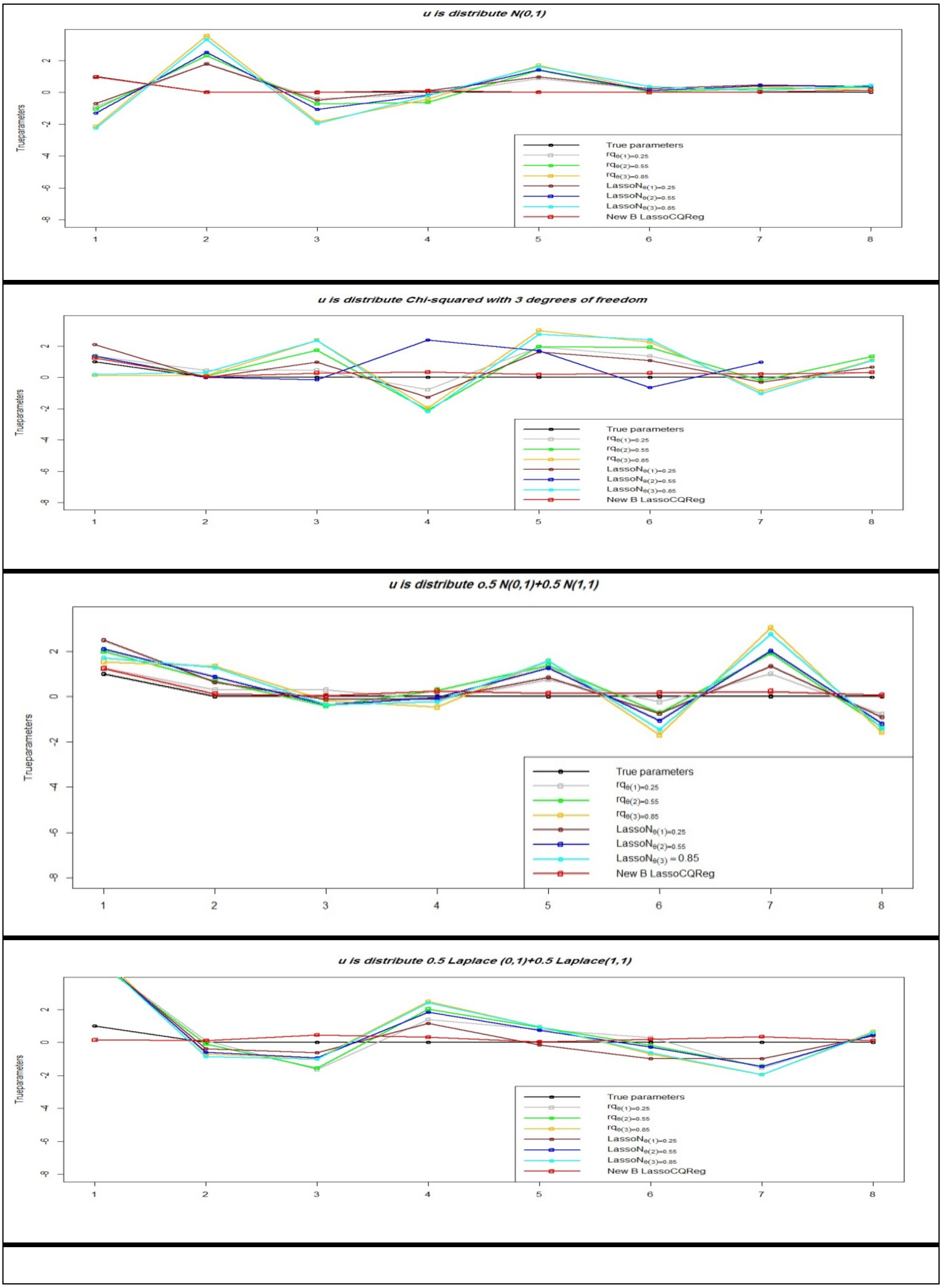

Figure 1: Bar chart summarizing the parameters of the studied methods close to true parameters

4. Real data example

In this section, we apply the three methods to prostate cancer data, first analysed by [22]. It was also analyzed using the model selection in quantile regression model by [3].

These data exist within "bayesQR" package in R. The sample size of prostate cancer data was of 97 observations, the response variable is the level of prostate antigen referred to as (lpsa) and there are eight covariates.

These covariates are x_1 logarithm of cancer amount, referred to as (lcavol), x_2 logarithm of the weight of prostate, referred to as (lweight), x_3 age , x_4 logarithm of the volume of benign enlargement of the prostate, referred to as (lbph), x_5 seminal vesicle invasion, referred to as (svi), x_6 logarithm of Capsular penetration in prostate cancer, referred to as (lcp), x_7 Gleason score, referred to as (gleason) and x_8 percentage of Gleason scores 4 or 5, referred to as (pgg45).

We applied a new Bayesian lasso composite quantile regression (**New B Lasso CQ Reg**) to the prostate cancer data and considered H=5 so that $\theta_h = [0.16 , 0.33, \ 0.50 \ , 0.66 \ ,0.83] \quad where\ h = 1,2,3,4,5) \ respectively$.

Modeling the relationship between response variable and the covariates via the three methods is summarized in Table 3.

Table 3: Parameter estimates via the three methods in the comparison

		rq						***Lasso N***			**New B Lasso CQ Reg**
Variables	$\theta_1 = 0.16$	$\theta_2 = 0.33$	$\theta_3 = 0.50$	$\theta_4 = 0.66$	$\theta_5 = 0.83$	$\theta_1 = 0.16$	$\theta_2 = 0.33$	$\theta_3 = 0.50$	$\theta_4 = 0.66$	$\theta_5 = 0.83$	M=5
Intercept	-0.521	-0.152	-0.057	0.170	0.519	-0.800	-0.305	-0.0003	0.291	0.763	0.459
lcavol	0.742	0.607	0.543	0.513	0.633	0.648	0.607	0.571	0.563	0.559	0.241
lweight	0.273	0.270	0.238	0.156	-0.007	0.227	0.245	0.219	0.185	0.162	0.155
age	-0.059	-0.169	-0.172	-0.133	-0.037	-0.093	-0.122	-0.140	-0.139	-0.119	0.152
lbph	0.042	0.182	0.201	0.182	0.183	0.144	0.164	0.168	0.153	0.121	0.208
svi	0.341	0.243	0.286	0.259	0.309	0.211	0.261	0.292	0.308	0.335	0.252
lcp	-0.349	-0.176	-0.158	-0.087	-0.112	-0.258	-0.223	-0.164	-0.101	-0.024	0.308
gleason	-0.041	0.046	0.127	0.032	-0.073	-0.023	0.005	0.032	0.040	0.041	0.114
pgg45	0.318	0.153	0.099	0.115	0.085	0.214	0.195	0.147	0.106	0.071	0.108

The parameter estimates in Table 3 are used to calculate the mean square error (MSE), as shown in table 4.

Table 4: Mean square error (MSE) for the three methods

Methods	$\theta_1 = 0.16$	$\theta_2 = 0.33$	$\theta_3 = 0.50$	$\theta_4 = 0.66$	$\theta_5 = 0.83$
rq	9.487	7.401	6.912	5.820	5.358
$\boldsymbol{Lasso\ N}$	11.247	8.234	6.618	5.252	5.412
New B Lasso CQ Reg	**5.077**				

Table 4 shows the value of (MSE) with our proposed method (New B Lasso CQ Reg). This value is **5.077.** Table 4 also indicates the results of (MSE) for the rq and $\boldsymbol{Lasso\ N}$ methods. The (MSE) calculated by our proposed method (New B Lasso CQ Reg) is much smaller than the (MSE) calculated by rq and $\boldsymbol{Lasso\ N}$ methods. This indicates that our proposed method performs better than the rq and $\boldsymbol{Lasso\ N}$ methods for all selected quantile levels.

5. Conclusions and recommendations

In this paper, we developed an efficient method (New B Lasso CQ Reg) to estimate model parameters and variables selection simultaneously, in composite quantile regression model via scale mixture uniform as prior distributions on the parameters, by using Bayesian approach. The simulation study illustrated that our proposed method is more efficient than other methods. The benefits of using the New B Lasso CQ Reg method are supported by the simulations performed as well as by the analysis of the real data. The results indicate higher accuracy in case of the New B Lasso CQ Reg method. So, we recommended, the researchers are working in field of quantile regression using composite quantile regression . Because of , this approach informative for all quantile level under study.

References

[1]- Alhamzawi. R., & Yu, K. “Bayesian Lasso-mixed quantile regression” Journal of Statistical Computation and Simulation, 84(4),pp. 868-880, 2014.

[2]- Alhamzawi, Rahim. “Bayesian elastic net Tobit quantile regression. ”Communications in Statistics-Simulation and Computation” 45(7),pp. 2409-2427, 2016

[3]- Alhamzawi, R. “Model selection in quantile regression models. ”Journal of Applied Statistics, 42(2), pp.445-458, (2015).

[4]- Alhamzawi, R. Yu, K., & Benoit, D. F. “Bayesian adaptive Lasso quantile regression. ” Statistical Modelling, 12(3), pp.279-297, (2012).

[5]- Benoit, D. F., Alhamzawi, R. and Yu, K. “Bayesian lasso binary quantile regression. ” Computational

Statistics, 28(6): pp.2861–2873. 2, (2013).

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

[7]- Candes, E., & TaoT. “The Dantzig selector: Statistical estimation when p is much larger than. ” The Annals of Statistics, pp2313-2351,(2007).

[8]- Cade, B. S., & Noon, B. R. “A gentle introduction to quantile regression for ecologists. ” Frontiers in Ecology and the Environment, 1(8), pp.412-420,(2003).

[9]- Huang, H., & Chen, Z. “Bayesian composite quantile regression. ”.Journal of Statistical Computation and Simulation, 85(18), pp.3744-3754,(2015).

[10]- Kostov, P., & Davidova, S. A. “quantile regression analysis of the effect of farmers’ attitudes and perceptions on market participation.” Journal of Agricultural Economics, 64(1), pp.112-132, (2013).

[11]- Koenker, R., & Bassett Jr, G. “Regression quantiles.” Econometrica : journal of the Econometric Society, pp.33-50, (1978).

[12]- Koenker, R., & Geling, O. “Reappraising medfly longevity.” a quantile regression survival analysis. Journal of the American Statistical Association,96(454),pp. 458-468 ,(2001).

[13]- Koenker, R. W., & d'Orey, V. “ Algorithm AS 229: Computing regression quantiles.” Journal of the Royal Statistical Society. Series C (Applied Statistics), 36(3), pp.383-393,(1987).

[14]- Koenker, R. W., & d'Orey, V. “Algorithm AS 229: Computing regression quantiles.” Journal of the Royal Statistical Society. Series C (Applied Statistics), 36(3), pp.383-393,(1987).

[15]- Koenker, R. "Quantile regression" . Cambridge university press: (2005). (No. 38).

[16]-Kozumi, Hideo, and Genya Kobayashi. "Gibbs sampling methods for Bayesian quantile regression." Journal of statistical computation and simulation 81, no. 11 ,pp.1565-1578. (2011).

[17]- Li, Y. and Zhu, J . “L1-norm quantile regression. ” Journal of Computational and Graphical Statistics, 17(1). pp. 163-185., (2008).

[18]- Li, Q., Xi, R., Lin, N., et al. “Bayesian regularized quantile regression. ” Bayesian Analysis, 5(3);pp. 533–556. (2010).

[19]- Mallick, H., & Yi, N. “A new Bayesian lasso. ” STATISTICS AND ITS INTERFACE, 7(4),pp. 571-582, (2014).

[20]- Park, T., & Casella, G. “The Bayesian lasso. ” Journal of the American Statistical Association, 103(482), pp.681-686,(2008).

[21]- Sun, W., Ibrahim, J. G., and Zou, F. “Genome wide multiple-loci mapping in experimental crosses by iterative adaptive penalized regression. ” Genetics, 185(1),pp.349–359. 2 , (2010).

[22]- Stamey, T. A., Kabalin, J. N., McNeal, J. E., Johnstone, I. M., Freiha, F., Redwine, E. A., & Yang, N. . “Prostate specific antigen in the diagnosis and treatment of adenocarcinoma of the prostate. ” II. Radical prostatectomy treated patients. ”The Journal of urology, 141(5),pp. 1076-1083,(1989).

[23]- Tibshirani, R. “Regression shrinkage and selection via the lasso. ” .Journal of the Royal Statistical Society. Series B (Methodological),pp. 267-288, (1996).

[24]- Tibshirani, R. “Regression shrinkage and selection via the lasso. ” a retrospective". Journal of the Royal Statistical Society: Series B (Statistical Methodology), 73(3), pp.273-282,(2011).

[25]- Wang, H. and He, X . “Detecting differential expressions in GeneChip microarray studies: a quantile approach. ” Journal of the American Statistical Association,102(477);pp. 104–112. 1, . (2007).

[26]- Wei, Y., Pere, A., Koenker, R., and He, X . “Quantile regression methods for reference growth charts.” .Statistics in medicine, 25(8); pp.1369–1382, . (2006).

[27]-Wu, Y. and Liu, Y. “Variable selection in quantile regression.”Statistica Sinica,19(2),pp. 801. 2, (2009).

[28]-Yu, K., & Moyeed, R. A. "Bayesian quantile regression. “ Statistics & Probability Letters, 54(4),pp. 437-447 ,(2001).

[29]- Zou, H. "The adaptive lasso and its oracle properties. “ Journal of the American statistical association, 101(476), pp.1418-1429 ,(2006).

[30]- Zou, H., & Yuan, M , “Composite quantile regression and the oracle model selection theory. ” The Annals of Statistics, pp. 1108-1126, . (2008).