International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Bayesian Quantile Regression using Normal-Compound Gamma Priors

Ahmed Alhamzawi[1], Gorgees Shaheed Mohammad[2]

Submitted: 19/01/2023 **Accepted:** 30/03/2023

Abstract

There are several robust regression methods that can be used to select parsimonious models in regression. In this paper, a study of the quantile regression with a normal-compound gamma scale mixture prior is presented. The Monte Carlo Markov Chain (MCMC) is derived for posterior inference. Finally, the robustness of this model is demonstrated using both real and simulated data. The results show that the proposed method performs very well compared to some of the existing methods.

Keywords: Quantile Regression, Monte Carlo Markov Chain, EM algorithm, Normal-Compound Gamma.

1. Introduction

One popular alternative method to standard mean regression which may provide greater information is quantile regression, which has in recent years proved to be a robust approach to the analysis of regression models (Koenker and Bassett, 1978; Li et al., 2010; Yu and Moyeed, 2001). The advantage of using this method becomes more readily apparent when trying to accomplish a more complete investigation of the relationship between a response variables and covariates. The flexibility of quantile regression models have made them into such an attractive tool that has numerous applications in the real world applications such as economics, agricultural, biology and several other fields.

Let $y = (y_1, \cdots, y_n)^T$ be a vector of observations and consider the linear model

$$y = X\beta + \epsilon, \qquad (1)$$

with a $n \times p$ design matrix of covariates $X = (x_1, \cdots, x_p)$, a $p \times 1$ of vector of unknown regression coefficient $\beta = (\beta_1, \cdots, \beta_p)^T$ and $\epsilon = (\epsilon_1, \cdots, \epsilon_n)^T$ where $\epsilon_i \sim N(0, \sigma^2)$. We can calculate the θth quantile regression model using the inverse cumulative distribution function $Q_{y_i}(\theta|x_i)$ of y_i given x_i defined as

$$Q_{y_i}(\theta|x_i) = \beta_0 + x_i'\beta. \qquad (2)$$

[1]University of AL-Qadisiyah, College of Science, Department of Mathematics, Iraq
ahmed.alhamzawi@qu.edu.iq
[2]University of AL-Qadisiyah, College of Education, Department of Mathematics, Iraq
gorgees.alsalamy@qu.edu.iq

We can estimate the value of the coefficient β by minimizing the following

$$\sum_{i=1}^{n} \rho_\theta\,(y_i - x_i^T\beta), \qquad (3)$$

with quantile check loss function ρ_θ defined by

$$\rho_\theta(x) = \{\theta x, if\ x \geq 0, (\theta - 1)x, if\ x < 0. \qquad (4)$$

Equivalently, we can write the check function as

$$\rho_\theta(x) = x\theta - xI(x \leq 0) \qquad (5)$$

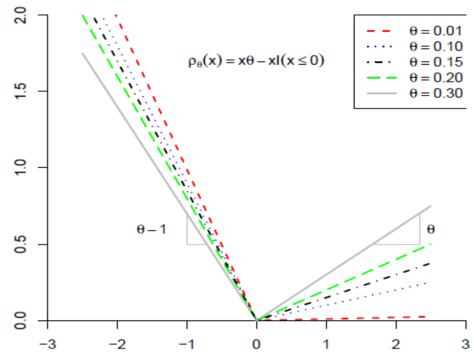

Figure 1. A plot showing check loss function ρ_θ for different values of θ: The dashed red line for $\theta = 0.01$, the dotted blue line for $\theta = 0.10$, the dashed doted black line for $\theta = 0.15$, the long dashed green linefor $\theta = 0.20$ and the solid gray line for $\theta = 0.30$.

Since the linear check function (4) is not differentiable at 0, there is no closed form solution available. However, the minimization in (3) can obtained using linear programming methods (Koenker and d'Orey, 1987). From a Bayesian point of view, the minimization check function in (3) is equivalent to maximize the likelihood function when the error distribution has asymmetric Laplace distribution of the form

$$f(\epsilon_n|\sigma) = \theta(1-\theta)\sigma\{-\sigma\rho_\theta(\epsilon_n)\}, \quad (6)$$

where the likelihood of the response variable of interest y is given by

$$f(y|X,\beta,\theta) = \prod_{i=1}^{n} \theta(1-\theta)\sigma\{-\sigma\rho_\theta(y_i - x_i^T\beta)\} \quad (7)$$

It is straightforward to see that maximizing (7) is equivalent to minimizing (3). In subsequent work of Kozumi and Kobayashi (2011), it was showed that the asymmetric Laplace distribution (6) can be written as a mixture of an exponential distribution with a scaled normal distribution. Thus, we have the Bayesian hierarchical model

$$y_i = x_i^T\beta + \xi_1 v_i + \xi_2\tau^{-1/2}\sqrt{v_i}w_i \quad (8)$$

$$v|\tau \sim \prod_{i=1}^{n} \tau(-\tau v_i) \quad (9)$$

$$w \sim \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}}\left(-\frac{1}{2}w_i^2\right) \quad (10)$$

where $v = (v_1, \dots, v_n)$, $w = (w_1, \dots, w_n)$

$$\xi_1 = \frac{1-2\theta}{\theta(1-\theta)} \quad and \quad \xi_2 = \sqrt{\frac{2}{\theta(1-\theta)}}. \quad (11)$$

In the following sections, we will introduce the NCG prior for quantile regression coefficients and the Gibbs sampler for this model will be constructed. Then, we will compare the results with other existing models using both simulated and real data examples.

2. Scale Mixture of Compound Gamma Distribution

We will study the quantile regression for a scale mixture model with compound gamma density

$$\pi(x) = \int_0^\infty \dots \int_0^\infty \left[\prod_{i=1}^{N} \frac{z_{i+1}^{c_i}}{\Gamma(c_i)} z_i^{c_i-1}\{-z_i z_{i+1}\}\right] dz_2 \dots dz_N \quad (12)$$

where $z_1 = x$ and $z_{N+1} = \phi$ is a constant. The behavior of our prior is clearly demonstrated in Figure 1 for specific values of N and c_1. One can clearly see that for higher values c_1, we have flatter head with a heavier tail for increasing values of N. However, if we consider values of c_1 closer to zero, then we have pole at the origin with a sharper pole as the value N increases. The values for these hyperparameters can be changed depending on the sparsity of our model. The above density is generalization of different types of priors considered in the past such as: the Three-Parameter Beta Distribution, the Scaled Beta2 (SBeta2) family (Armagan et al., 2011; P´erez et al., 2017) for $N = 2$, the Beta prime distribution when $N = 2$ and $\phi = 1$ (Bai and Ghosh, 2018) and horseshoe prior for $N = 4, c_3 = 1/2$ and $c_4 = 1/2$ (Carvalho et al., 2010) given by the hierarchical model

$$\beta_i|rest \sim N(0, \sigma^2 z_1),$$

$$z_1^{1/2} \sim C^+(0, z_2),$$

$$z_2^{1/2} \sim C^+(0,1).$$

The complexity of the prior in (12) may be presented more straightforwardly by using the following equivalence

Proposition 1. If $z_1 \sim CG(c_1, \dots, c_N, \phi)$, then

(1) $z_1 \sim G(c_1, z_2), z_2 \sim G(c_2, z_3), \dots, z_N \sim G(c_N, \phi)$

(2) $z_1 \sim G(c_1, 1), z_2 \sim IG(c_2, 1), \dots, z_N \sim \{G(c_N, \phi)\ odd\ \ N\ IG(c_N, \phi)\ even\ \ N$

where $G(\alpha, \delta)$ is the gamma distribution with shape parameter α and inverse scale (rate) parameter δ and $IG(\alpha, \delta)$ is the inverse gamma distribution with shape parameter α and scale parameter δ.

Proof. The proof of this equivalence is giving in (Alhamzawi and Mohammad, 2022).

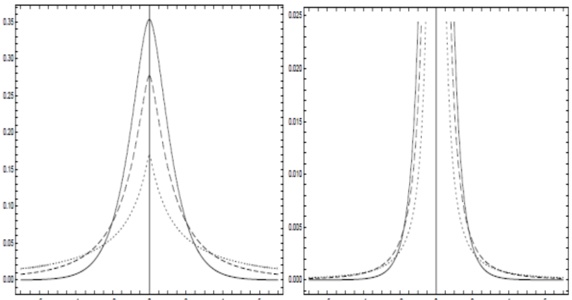

Figure 2. A demonstration of the density (12) for various values of c_1 and N. The solid, dashed and dotted lines represent $N = 2, N = 4$ and $N = 8$, respectively.

To proceed a Bayesian analysis, we assign a normal-compound gamma prior on β and have the following hierarchal model

$$y_i = x_i^T\beta + \xi_1 v_i + \xi_2\tau^{-1/2}\sqrt{v_i}w_i \quad (15)$$

$$v|\tau \sim \prod_{i=1}^{n} \tau(-\tau v_i) \quad (16)$$

$$w \sim \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}}\left(-\frac{1}{2}w_i^2\right) \quad (17)$$

$$\beta_i|z_1, z_2, \dots, z_N, \tau \sim N(0, \tau^{-1}\prod_{i=1}^{N} z_i) \quad (18)$$

$$\tau \sim G(c_0, d_0) \quad (19)$$

The full conditional posterior distributions for the above model can be calculated easily to construct the Gibbs sampler. Additionally, we will propose a method for updating the hyperparameters that will be incorporated into the Gibbs sampler.

3. Posterior Inference

The parameters of interest can be drawn from their conditional posteriors as follows:

- **Update** β

$$P(\beta|X,y,\dots) \propto P(y|\beta,\dots)\pi(\beta),$$

$$\propto \left\{-\frac{(y - X\beta - \xi_1 v)^T V^{-1}(y - X\beta - \xi_1 v)}{2\tau^{-1}\xi_2^2}\right\} \times \left\{-\tau\frac{\beta^T Z^{-1}\beta}{2}\right\} \quad (20)$$

$$\propto \left\{-\frac{\tau}{2}[-2\xi_2^{-2}(y - \xi_1 v)^T V^{-1}X\beta + \xi_2^{-2}\beta^T X^T V^{-1}X\beta + \beta^T Z^{-1}\beta]\right\}$$

where $\mu_\beta = \xi_2^{-2}\Sigma^{-1}X^T V^{-1}(y - \xi_1 v)$,$V = diag(v_1,\dots,v_n)$, $Z = daig(\prod_{i=1}^N z_{k1},\dots,\prod_{i=1}^N z_{kp})$ and $\Sigma = \xi_2^{-2}X^T V^{-1}X + Z^{-1}$. Therefore, we have the normal distribution $N(\mu_\beta, \Sigma^{-1}\sigma^2)$.

- **Update** z_k (if k is odd number)

$$P(z_k|X,y,\dots) \propto \pi(\beta_i|z_1,z_2,\dots,z_N,\tau)\pi(z_k)$$

$$\propto \frac{1}{\sqrt{z_k}}\left\{-\frac{\tau\beta^T Z^{-1}\beta}{2}\right\} \times (z_k)^{c_k-1}\{-z_k\}, \quad (21)$$

$$\propto (z_k)^{\left(c_k-\frac{1}{2}\right)-1}\left\{-\frac{1}{2}[\tau\beta^T Z_{-k}^{-1}\beta z_k^{-1} + 2z_k]\right\},$$

where $Z_{-k} = daig(\prod_{i=1,i\neq k}^N z_{k1},\dots,\prod_{i=1,i\neq k}^N z_{kp})$. Thus, we have the generalized gaussian distribution $GIG\left(\tau\beta^T Z_{-k}^{-1}\beta, 2, c_k - \frac{1}{2}\right)$.

- **Update** z_k (if k is even number)

$$P(z_k|X,y,\dots) \propto \pi(\beta_i|z_1,z_2,\dots,z_N,\tau)\pi(z_k)$$

$$\propto \frac{1}{\sqrt{z_k}}\left\{-\frac{\tau\beta^T Z^{-1}\beta}{2}\right\} \times (z_k)^{-c_k-1}\{-Z_k^{-1}\}, \quad (22)$$

$$\propto (z_k)^{-\left(c_k+\frac{1}{2}\right)-1}\left\{-\left[\frac{\tau\beta^T Z^{-1}\beta}{2} + 1\right](z_k)^{-1}\right\},$$

which is the inverse-gamma distribution $IG\left(c_k + \frac{1}{2}, \frac{\tau\beta^T Z^{-1}\beta}{2} + 1\right)$.

- **Update** v

$$P(v|X,y,\dots) \propto P(y|\beta,\dots)\pi(v_i|\tau)$$

$$\propto \frac{1}{\sqrt{v}}\left\{-\frac{(y - X\beta - \xi_1 v)^T V^{-1}(y - X\beta - \xi_1 v)}{2\tau^{-1}\xi_2^2}\right\} \times \{-\tau v\} \quad (23)$$

$$\propto v^{\frac{1}{2}-1}\left\{-\frac{1}{2}\left[\tau\xi_2^{-2}(y - X\beta)^T V^{-1}(y - X\beta) + \tau\left(\frac{\xi_1^2}{\xi_2^2} + 2\right)v\right]\right\}$$

Hence, we have the generalized gaussian distribution $GIG\left(\frac{\tau(y_i - x_i^T\beta)^2}{\xi_2^2}, \frac{\tau\xi_1^2}{\xi_2^2} + 2\tau, \frac{1}{2}\right)$.

- **Update** τ

$$P(\tau|X,y,\dots) \propto P(y|\beta,\dots)\pi(\beta_i|z_1,z_2,\dots,z_N,\tau)\pi(\tau)$$

$$\propto \left(\tau^{n/2}\left\{-\frac{(y - X\beta - \xi_1 v)^T V^{-1}(y - X\beta - \xi_1 v)}{2\tau^{-1}\xi_2^2}\right\}\right) \quad (24)$$

$$\times \left(\tau^{p/2}\left\{-\frac{\beta^T Z^{-1}\beta}{2\tau^{-1}}\right\}\right)(\tau^n(-\tau v))\tau^{c_0-1}\{-d_0\tau\}$$

$$\propto \tau^{c_0+\left(\frac{3n+p}{2}\right)-1}\left\{-\tau\left[\sum_{i=1}^n \frac{(y_i - x_i^T\beta - \xi_1 v_i)^2}{2\xi_2^2 v_i} + \sum_{i=1}^p \frac{\beta_i^2}{2\prod_{i=k}^N z_k} + d_0\right]\right\}$$

this is again the gamma distribution $G\left(c_0 + \frac{3n+p}{2}, \sum_{i=1}^n \frac{(y_i - x_i^T\beta - \xi_1 v_i)^2}{2\xi_2^2 v_i} + \frac{\beta^T\beta}{2\prod_{i=k}^N z_k} + d_0\right)$.

- **Update** c_k

We use the same method provided in (Alhamzawi, 2022) by calculating the expectation of the log complete-data likelihood

$$Q(\xi,\xi^{old}) = \sum_{k=1}^N \sum_{i=1}^p (-1)^{k+1} c_k E_{c_k^{old}}[(z_{ki})|y] + c_N(\phi) - \sum_{k=1}^N (\Gamma(c_k)) + C \quad (25)$$

where $\xi^{old} = (c_1^{old},\dots,c_N^{old})$ and C all the terms not containing $c_1, c_2, \dots, c_N$. Then, we have

$$\Gamma'(c_k) = \sum_{i=1}^p (-1)^{k+1} E_{c_k^{old}}[(z_{ki})|y] + c_N(\phi)I(k = N) \quad (26)$$

4. Simulation Studies

In this section, we will illustrate the proposed model using simulation studies. Here, we compare the proposed method (NCG10) with the Beta prime prior for scale parameters (NCG2, Bai and Ghosh, 2018), Bayesian quantile regression (Alhamzawi et al., 2011), Bayesian quantile regression with lasso penalty (Li et al., 2010), and Bayesian quantile regression with the elastic net penalty (Li et al., 2010). The data are simulated from the following model

$$y = X\beta + \epsilon, \quad \epsilon_i \sim N(0, \sigma^2 I_n)$$

Methods are evaluated based on the mean squared error (MSE), the false positive rate (FPR) and the false negative rate (FNR).

Example 1 (Very sparse model)

In this example we consider a very sparse model by setting $\beta = (4,0,0,0,0,0,0,0,0,0,0,0)$ and $\sigma^2 = \{1,9,25\}$. We consider three different scenarios by setting $\theta = \{0.05, 0.50, 0.95\}$. The covariates are simulated independently from $N(0,\Sigma)$ where the $(i,j)^{th}$ elements of Σ is $0.5^{|i-j|}$. We run 50 simulations. The results are summarized in Table 1. All results are averaged based on 100 replications. The results show that the proposed method perform very well compare to other existing methods. It produces smallest MSE for all quantiles considered in this example. The proposed method exhibits promising performance in terms of variable selection described by FPRs and FNRs.

Table 1. Results for Example 1. All results are averaged over 100 replications and their associated standard deviations (sd) are listed in the parentheses.

	θ	**MSE (sd)**	**FPR (sd)**	**FNR (sd)**
NCG2	0.05	0.4064 (0.2996)	0.0500 (0.2611)	0.0000 (0.0000)
NCG10	0.05	**0.3002** (0.2669)	0.0200 (0.1407)	0.0000 (0.0000)
Bqr	0.05	1.1100 (0.4859)	0.7400 (0.9600)	0.0000 (0.0000)
qr.lasso	0.05	0.8253 (0.3898)	0.4100 (0.6977)	0.0000 (0.0000)
qr.enet	0.5	0.8446 (0.4838)	0.4700 (0.7311)	0.0000 (0.0000)
NCG2	0.5	0.3541 (0.3548)	0.0800 (0.3674)	0.0000 (0.0000)
NCG10	0.5	**0.2543** (0.3213)	0.0200 (0.1407)	0.0000 (0.0000)
Bqr	0.5	1.0565 (0.5507)	0.5900 (0.8420)	0.0000 (0.0000)
qr.lasso	0.5	0.7540 (0.4490)	0.3800 (0.6784)	0.0000 (0.0000)
qr.enet	0.5	0.8446 (0.4838)	0.4700 (0.7311)	0.0000 (0.0000)
NCG2	0.95	0.3811 (0.3058)	0.0700 (0.2564)	0.0000 (0.0000)
NCG10	0.95	**0.2837** (0.2710)	0.0400 (0.1969)	0.0000 (0.0000)
Bqr	0.95	1.0976 (0.4529)	0.6200 (0.7756)	0.0000 (0.0000)
qr.lasso	0.95	0.7868 (0.3749)	0.3800 (0.5993)	0.0000 (0.0000)
qr.enet	0.95	0.8803 (0.3964)	0.4400 (0.6715)	0.0000 (0.0000)

Example 2 (Sparse model)

Here we consider a simple sparse model. The covariates are set similar to the above example except that we set $\beta = (3,1.5,0,0,1,2,0,0,5,0)$. The results are summarized in Table 2. Again, the proposed method NCG10 performs very well in terms of both covariate selection and prediction accuracy.

Table 2. Results for Example 2.

	θ	**MSE (sd)**	**FPR (sd)**	**FNR (sd)**
NCG2	0.05	1.0262 (0.4964)	0.2900 (0.5738)	0.5000 (0.5946)
NCG10	0.05	**0.9531** (0.4755)	0.1300 (0.3667)	0.6500 (0.6723)
Bqr	0.05	1.1325 (0.5284)	0.3800 (0.5993)	0.4000 (0.5505)
qr.lasso	0.05	1.0312 (0.4876)	0.3200 (0.5840)	0.4200 (0.5352)
qr.enet	0.05	1.0825 (0.5095)	0.3700 (0.6139)	0.3800 (0.5464)
NCG2	0.5	0.9588 (0.4252)	0.3000 (0.6113)	0.3500 (0.5389)
NCG10	0.5	**0.8632** (0.4073)	0.1400 (0.3766)	0.5800 (0.6225)
Bqr	0.5	1.0816 (0.4397)	0.4600 (0.6730)	0.2700 (0.4683)
qr.lasso	0.5	0.9713 (0.4017)	0.3100 (0.5808)	0.3100 (0.5064)
qr.enet	0.5	1.0199 (0.4175)	0.4200 (0.5891)	0.2500 (0.4578)
NCG2	0.95	1.0337 (0.4636)	0.3700 (0.6460)	0.3600 (0.4824)
NCG10	0.95	**0.9253** (0.4598)	0.1800 (0.4353)	0.5200 (0.5942)
Bqr	0.95	1.1534 (0.4869)	0.4500 (0.7017)	0.3100 (0.4648)
qr.lasso	0.95	1.0459 (0.4539)	0.3500 (0.5925)	0.3000 (0.4606)
qr.enet	0.95	1.1018 (0.4693)	0.4200 (0.6694)	0.3100 (0.4648)
NCG2	0.95	1.0337 (0.4636)	0.3700 (0.6460)	0.3600 (0.4824)

5. A real data example

Here, we compare the performance of the five methods in Section 4, NCG2, NCG10, Bqr, qr.lasso, qr.enet, on the Prostate cancer data (Stamey et al., 1989), where there are 97 observations and 8 covariates. This data set is available in the R package Brq (Alhamzawi and Ali, 2020). The response of interest is the logarithm of prostate-specific antigen. Table 3 lists briefly a description of the covariates (clinical measures).

Table 3. Description of the clinical measures

Covariates (clinical measures)	Description of the clinical measures
lcavol	The logarithm of cancer volume
lweight	The logarithm of prostate weight
age	Age
lbph	The logarithm of the amount of benign prostatic hyperplasia
svi	Seminal vesicle invasion
lcp	The logarithm of capsular penetration
gleason	The Gleason score
pgg45	The percentage Gleason score 4 or 5

Following Mallick and Yi (2014) and Alhamzawi and Mallick (2020), we analyze the prostate cancer data set by dividing it into a subset of 67 observations (training observations) and a subset of 30 observations (testing observations). Model fitting is carried out on the first subset (67 observations named training observations) and performance is evaluated with the prediction error (MSE) on the second subset (30 observations named testing observations). The results of mean squared errors and standard deviations are summarized in Table 4. We can see that the proposed method performs very well in terms of prediction accuracy. In addition, the trace plots (see for example Figure 3 when $\theta = 0.95$) and histograms (see for example Figure 4 when $\theta = 0.95$) show that the proposed Gibbs sampler converge very fast to the stationary distribution. Hence, the simulations and real data analyses show the importance of the proposed method.

Table 4. Results of the real data

	MSE (sd)	MSE (sd)	MSE (sd)
	$\theta = 0.05$	$\theta = 0.50$	$\theta = 0.95$
NCG**2**	0.5235 (0.7032)	0.5212 (0.7046)	0.5206 (0.7032)
NCG**10**	**0.4980** (0.6813)	0.5027 (0.6789)	0.4981 (0.7046)
Bqr	0.5502 (0.7456)	0.5499 (0.7463)	0.5507 (0.7037)
qr.lasso	0.5208 (0.7113)	0.5233 (0.7079)	0.5221 (0.6813)
qr.enet	0.5208 (0.6972)	0.5248 (0.7149)	0.5220 (0.6789)

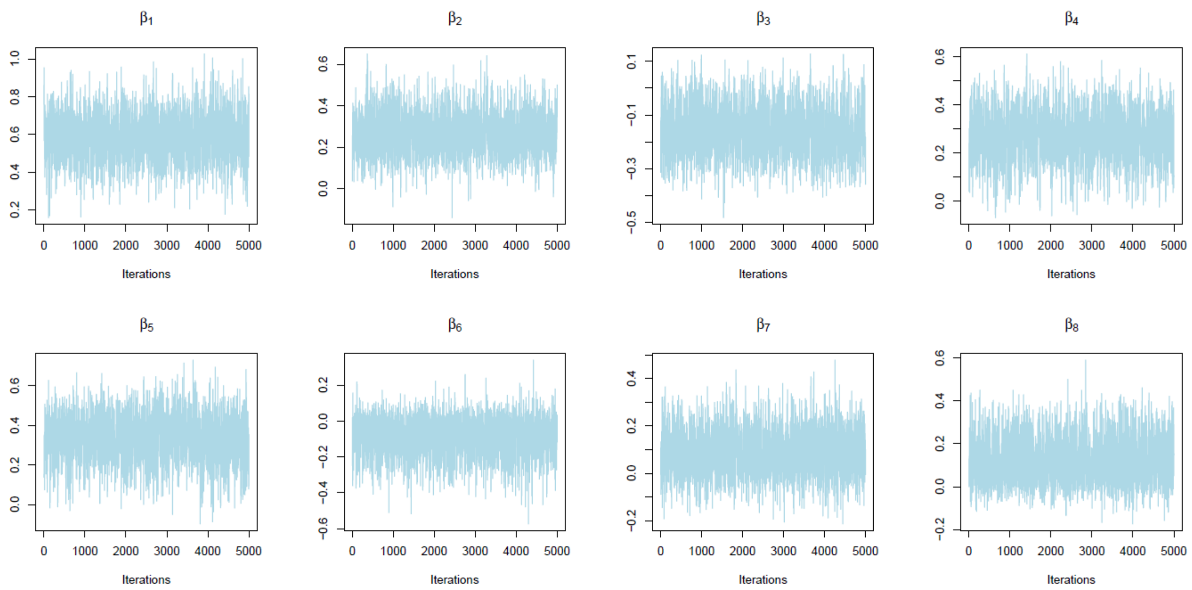

Figure 3. Trace plots of prostate cancer data covariates when $\theta = 0.95$.

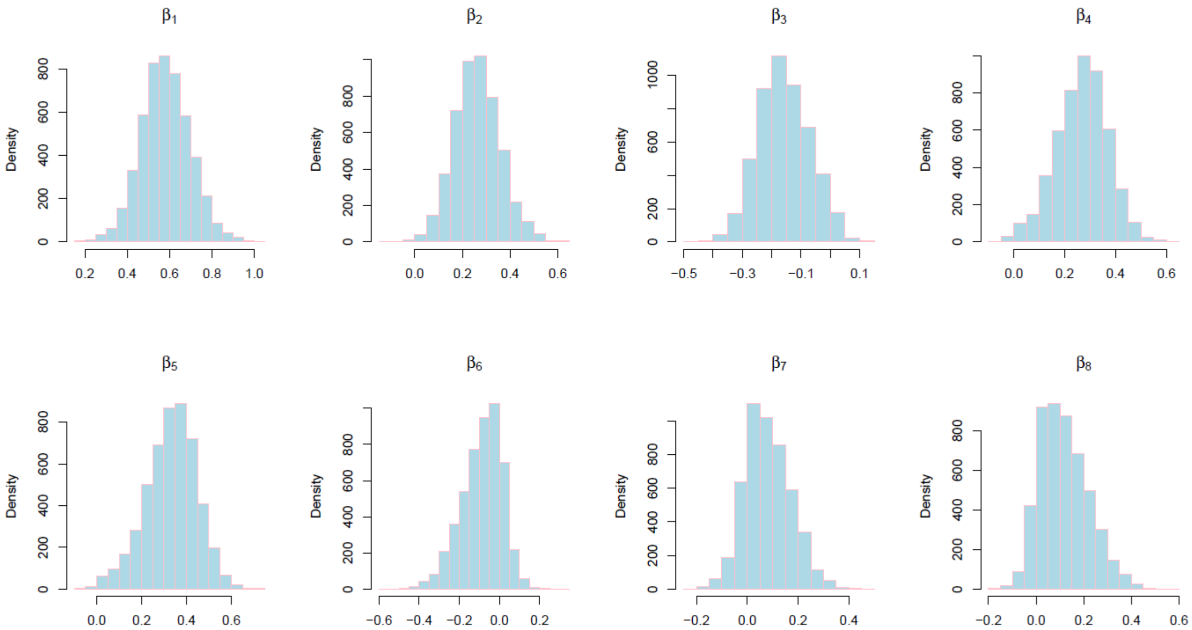

Figure 4. Histograms of prostate cancer data covariates when $\theta = 0.95$.

6. Concluding Remarks

In this paper, we have proposed a new method for Bayesian quantile regression using a normal-compound gamma prior for the regression coefficients. We have compared the proposed method (NCG10) with the Beta prime prior for scale parameters (NCG2, Bai and Ghosh, 2018), Bayesian quantile regression (Alhamzawi et al., 2011), Bayesian quantile regression with lasso penalty (Li et al., 2010), and Bayesian quantile regression with the elastic net penalty (Li et al., 2010). The results of the simulation studies and real data analyses show that NCG10 performs very well compared with the above existing methods. The proposed method can be extended easily to other existing methods such as: tobit quantile regression and binary quantile regression.

References

[1] Alhamzawi, A. and G. S. Mohammad (2022). Statistical inference with normal-compound gamma priors in regression models. arXiv preprint arXiv:2212.13144 .

[2] Alhamzawi, R. and H. T. M. Ali (2020). Brq: An r package for Bayesian quantile regression. METRON 78 (3), 313–328.

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

[4] Alhamzawi, R., K. Yu, V. Vinciotti, and A. Tucker (2011). Prior elicitation for mixed quantile regression with an allometric model. Environmetrics 22 (7), 911–920.

[5] Armagan, A., M. Clyde, and D. Dunson (2011).

Generalized beta mixtures of gaussians. Advances in neural information processing systems 24.

[6] Bai, R. and M. Ghosh (2018). On the beta prime prior for scale parameters in high-dimensional bayesian regression models. arXiv preprint arXiv:1807.06539 .

[7] Carvalho, C. M., N. G. Polson, and J. G. Scott (2010). The horseshoe estimator for sparse signals. Biometrika 97 (2), 465–480.

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

[9] Koenker, R. W. and V. d'Orey (1987). Algorithm as 229: Computing regression quantiles. Journal of the Royal Statistical Society. Series C (Applied Statistics) 36 (3), 383–393.

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

[11] Li, Q., R. Xi, N. Lin, et al. (2010). Bayesian regularized quantile regression. Bayesian Analysis 5 (3), 533–556.

[12] Mallick, H. and N. Yi (2014). A new Bayesian lasso. Statistics and its interface 7 (4), 571.

[13] P´erez, M.-E., L. R. Pericchi, and I. C. Ram´ırez (2017). The scaled beta2 distribution as a robust prior for scales. Bayesian Analysis 12 (3), 615–637.

[14] Stamey, T. A., J. N. Kabalin, J. E. McNeal, I. M. Johnstone, F. Freiha, E. A. Redwine, and N. Yang (1989). Prostate specific antigen in the diagnosis and treatment of adenocarcinoma of the prostate. ii. radical prostatectomy treated patients. The Journal of urology 141 (5), 1076–1083.

[15] Yu, K. and R. A. Moyeed (2001). Bayesian quantile regression. Statistics & Probability Letters 54 (4), 437–447.