

Received : 22/9/2019

Accepted :25/11/2019

New Bayesian Single Index Quantile Regression Based on Uniform Scale Mixture

Samer Alalaq^a, Taha Alshaybawee,^b

^a General Directorate of Education, Misan, Iraq

^b Department of Statistics , College of Administration and Economics, University of Al-Qadisiyah, Al Diwaniyah, Iraq

E-mail addresses: taha.alshaybawee@qu.edu.iq.

sam371979371979@gmail.com

Abstract:

To scale back the dimensionality while holding a lot of flexibility of a nonparametric model Wu, et al. (2010) proposed a single index conditional quantile regression model. In this paper, a new Bayesian lasso for single index quantile regression model is proposed based on a scale mixture uniform. In addition, we construct an efficient and sampling Gibbs algorithm for posterior inference based on a uniform scale mixture representation for Laplace distribution. Simulation study have considered to evaluate our proposed method compare to the existing methods. The results of simulations indicate that the new Bayesian algorithm performs well.

Keywords: Bayesian inference, single index, quantile regression, prior distribution, uniform scale mixture.

1. Introduction

In real applications, the exact response variable cannot be predicted from the prediction variables. Classical regression focuses on predicting a modal response variable Y for a single or set of predictor variables X , prediction of the response variable is given by prediction variables $E(Y|X)$, called regression function (Gujarati 2003; Weisberg 2005). Nevertheless, many

researchers claim that the fitting of the conditional mean is unable to give a complete picture of underlying interrelationships. For example, Mosteller and Tukey (1977) indicate “. The regression curve gives a large summary of the averages corresponding distributions to the set x ’s. We could go further and compute several different regression curves corresponding to the various percentage points of the distributions and thus get a

more complete picture of the set. Ordinarily, this is not done, and so regression often gives a rather incomplete picture. Just as the mean gives an incomplete picture of a single distribution, so the regression curve gives a correspondingly incomplete picture for a set of distributions". Quantile regression (QR) suggested by Koenker and Basset (1978), allowing the person to thoroughly review the conditional distribution of Y over X in completely different locations, thus giving a complete view of the interrelationships between Y and X. Subsequently, quantile regression has become a unite statistical methodology for estimating models of conditional quantile functions Wu & Liu (2009). QR usually provides a more complete picture description of the distribution of the response over the mean, it offers a practically necessary alternative to classical mean regression. Quantile regression is additionally a valuable alternative to the usually utilized Cox proportional risk model and the accelerated failure time model (AFT) utilized in survival analysis Koenker, R., & Geling, O. (2001). For instance, in a cancer study, it's known that treatment can cause different impacts among patients in lower or higher quantiles of the survival distribution Wu, et al. (2010). Nonparametric quantile regression was studied by Koenker (2005). Kernel primarily based methods and spline smoothing ways, have been used to accommodate nonlinear

relationships; see for more details, Yu and Jones (1998), He and Shi (1994) and Koenker et al. (1994). There are problems in the non-parameter estimation of the existence of high dimensions that can be solved using the single index model and avert the "curse of dimensionality" (Bellman et al., 1966) in nonparametric problems by that the response is just associated with a single linear set of the covariates. In addition, the main problem of non-parameter quantile regression that the estimated function is often hard to see and interpret with multivariable variables. To scale back the dimensionality while holding a lot of flexibility of a nonparametric model Wu, et al. (2010) proposed a single index conditional quantile regression model. They are the most common and necessary models in statistics as well as applied quantitative sciences such as econometrics and psychology due to their ability to reduce dimensions (Ichimura, 1993). Whereas this model is conditional on the quantile regression so it becomes more immune to outliers and includes a heavier tail errors distribution than the mean regression. The mathematical type for the τ th quantile level SIQR model is shows by:

$$y = m_{\tau}(x_i^{\tau} \beta_{\tau}) + \varepsilon_{\tau} \quad (1)$$

where $\tau \in (0,1)$, y is real response variable, x is that the p-dimensional vector of explanatory variables β_{τ} , is that the index parameter vector for the τ^{th}

quantile, and $g_{\tau}(\cdot)$ is the unknown nonparametric link function for the τ^{th} quantile, that is estimated for a single variable $(x_i^{\tau} \beta_{\tau})$.

The algorithm for estimating single index vector, based on linear quantile regression and with the unknown nonparametric function estimated by the local linear method was constructed by Wu et al. (2010). Computational algorithm for estimating and variable selection was advanced by Alkenani, Yu (2013), Lv et al. (2014) and Kuruwita (2016). This model has been given a great deal of interest from many researchers on both the applied and theoretical side. Some of the drawbacks experienced the frequentist approach are mentioned by Benoit and Van den Poel (2012); the difficult optimization of the estimated parameters and the structure of the confidence interval. Hu et al. (2013) were developed new Bayesian hierarchical model to estimate nonparametric link function in the single index model (SIM) conditional to QR. The Bayesian analysis method has become vastly applied, outcome of its ability to interest from all ready information in the analysis. Many studies newly presented to develop and expand the single index quantile regression model, for instance; Completely Bayesian method was sophisticated by Hu et al. (2013). A similar Bayesian approach was utilized by Zhao and Lian (2013) for the censored data in the single index quantile regression model. Alshaybawee et al. (2017) proposed a Bayesian approach with elastic net penalty to estimate and selected variables at the single index quantile regression model. In

this paper, a new Bayesian lasso for single index quantile regression model is proposed based on a scale mixture uniform to the prior parameters. In addition, we construct an efficient and simple Gibbs sampling algorithm for posterior inference based on a uniform scale mixture representation for Laplace distribution.

The rest of this paper is orderly as follows: In Section 2, we briefly review the single index model and the prior distribution. The Bayesian hierarchical model and the posterior inference for updating the parameters is described in section 3. Simulation study include in section 4 and some conclusions are given in section 5.

2. Single index quantile regression model and prior assumptions

The model of single index quantile regression was given as:-

$$y_i = m_{\tau}(x_i^{\tau} \beta_{\tau}) + \epsilon_{i\tau} \quad i = 1, 2, \dots, n \quad (2)$$

At τ^{th} quantile, the quantile error term $\epsilon_{i\tau}$ distributed as an asymmetric Laplace distribution with the density:

$$\pi(\epsilon | \sigma, \tau) = \frac{\tau(1-\tau)}{\sigma} \exp\left\{-\frac{1}{\sigma} \rho_{\tau}(\epsilon)\right\} \quad (3) \text{ where}$$

σ is that the scale parameter, $\rho_{\tau}(\cdot)$ is the check loss function defined by $\rho_{\tau}(\epsilon) = \epsilon(\tau - I\{\epsilon < 0\})$ for quantile level $\tau \in (0, 1)$.

The likelihood function for

$\epsilon_{i\tau} = y_i - m_\tau(x_i^T \beta_\tau), \quad i = 1, 2, \dots, n.$
where n is a sample size, shown as:

$$\begin{aligned} \pi(y|m, \beta, \sigma) &= \frac{\tau^n (1-\tau)^n}{\sigma} \exp \left\{ -\frac{1}{\sigma} \sum_{i=1}^n \rho_\tau(y_i - m(x_i^T \beta)) \right\} \end{aligned} \quad (4)$$

Quantile regression is often based on minimizing of the check function. However, direct use of the likelihood function is not convenient for Bayesian inference. A location-scale mixture (Kozumi and Kobayashi 2011) have used to the ALD which is useful here. We can rewrite Eq. (2) based on location – scale representation as follows:

$y_i = m(x_i^T \beta) + \delta v_i + \sqrt{\theta \sigma v_i} u_i$
where v_i is a standard exponential random variable, u_i is a standard normal random, $\delta = \frac{1-2\tau}{\tau(1-\tau)}$ and $\theta = \frac{2}{\tau(1-\tau)}$. Therefore, the likelihood function are rewritten as Kozumi, H., & Kobayashi, G. (2011).

$$\begin{aligned} \pi(y_i|\beta, m_\tau, v_i) &= \prod_{i=1}^n (2\pi\sigma\theta v_i)^{-\frac{1}{2}} \cdot \exp \left\{ -\frac{1}{2\sigma\theta v_i} (y_i - m_\tau(x_i^T \beta) - \delta v_i)^2 \right\} \end{aligned} \quad (5)$$

$\alpha \{ (\det[D^{-1}])^{-\frac{1}{2}} \exp \left\{ -\frac{(y-m_n-\delta v_i)^T D^{-1} (y-m_n-\delta v_i)}{2} \right\}$

Here

$$y_i = (y_1, y_2, \dots, y_n)^T, v_n = (v_1, v_2, \dots, v_n)^T$$

$$\begin{aligned} , D^{-1} &= \delta \sigma \text{diag}(v_1, v_2, \dots, v_n), \text{ and} \\ m_n &= (m_1, m_2, \dots, m_n)^T \\ &= (m(x_1^T \beta), m(x_2^T \beta), \dots, m(x_n^T \beta))^T \end{aligned}$$

As in Choi et al. (2011) and Gramacy and Lian (2012), the Gaussian method distribution is put as a prior distribution for the unknown nonparametric link function $m_\tau(\cdot)$. More specially, the prior distribution of $m_\tau(\cdot)$ is a Gaussian process, with the mean zero and the square covariance exponential function as follows:

$$\begin{aligned} m_\tau(\cdot) &\sim GP(0, E(\cdot, \cdot)), E(x_i, x_j) \\ &= \partial \exp \left(-\frac{(x_i - x_j)^2}{w} \right) \end{aligned}$$

Where ∂ and w are hyperparameters.

Therefore, The prior distribution for unknown function can be shown as follows:

$$\begin{aligned} \pi(m_\tau|\beta_\tau, \partial) &= \det[E_n]^{-\frac{1}{2}} \exp \left\{ -\frac{m_n^T E_n^{-1} m_n}{2} \right\} \end{aligned}$$

where E_n is the covariance matrix with dimension $(n \times n)$ and it can be written as

$$E(x_i^T \beta, x_j^T \beta) = \partial \exp \left\{ -(x_i - x_j)^T \beta \beta^T (x_i - x_j) \right\},$$

We have seen that the variance function $E(x_i, x_j) =$

$$\partial \exp \left(-\frac{(x_i^T \beta - x_j^T \beta)^2}{w} \right), \text{ follow Gramacy and}$$

Lian (2012), when the Gaussian process is employed as the prior for the nonparametric function, $\frac{\beta}{\sqrt{w}}$ is identifiable without the necessity for the constraint $\|\beta\| = 1$. Therefore, the parameter β will be instead of $\frac{\beta}{\sqrt{w}}$.

$$E(x_i, x_j) = \partial \exp \{-(x_i^T \beta - x_j^T \beta)^2\}$$

The inverse gamma distribution is considered as a hyper prior for ∂ , which implies that $\partial \sim IG(a_\gamma, b_\gamma)$ where a_γ and b_γ are the hyperparameters.

Scale mixture of uniform distribution (SMUD) for regression models has been utilized in few of occasions in literature. Walker et al. (1997) utilized SMU distribution in natural regression models in non-Bayesian framework. Damien and Qin (2000), & Walker, S. G. (1997) provided Gibbs sampler by utilized SMU in variance regression models and additionally to derive Gibbs sampler for autocorrelated heteroscedastic regression models. Nevertheless, its use has been specified in penalized regression setting. Choy et al. (2009) consider it in random volatility model by employing a two-stage scale mixture representation of the t-student distribution. It is important to note that the potential density function of Laplace distribution can be expressed in different forms of mixture representation. This

representation is much importance in research and studies, especially using the pyramidal method. This representation is important in facilitating the estimation process and increasing its efficiency. Some examples of this representation are important in this research. See for instance, Mallick and Yi (2014) they were developed the former methods by utilizing method called a new Bayesian Lasso, such suggested procedure for variable selection (VS), estimation of linear regression coefficient. Husseini (2017) proposed the Bayesian Lasso quantile regression method for (VS) to allocate an independent scale mixture of uniform distributions for regression coefficients. Alhamzawi & Ali (2018) offered a complete Bayesian method for tobit QR to reach the L_1 based on the representation of a mixture of the deviant Laplace distribution. A simple and effective Bayesian MCMC algorithm was introduced. The results of Mallick in his research were ideal and proved their competence from the previous Bayesian processes utilized.

Malik's results in his research were exemplary and proved efficient from previous Bayesian processes utilized.

Hierarchical model and MCMC sampler

A Laplace distribution function can be written as a scale mixture of uniform

distribution and a specific gamma distribution Mallick, H., Yi, N., 2014, i.e.

$$\pi(\beta_{j\tau}|\lambda_j, \sigma) = \prod_{j=1}^p \left(\frac{\lambda_j}{2\sigma}\right) \exp\left\{-\left(\frac{\lambda_j|\beta_{j\tau}|}{\sigma}\right)\right\}, \lambda > 0 \quad (6)$$

$$\left(\frac{\gamma}{2}\right) \exp\{-\left(\gamma|\beta_{j\tau}| \right)\}$$

$$= \int_{s>|\beta|} \frac{1}{2s} \frac{\gamma^2}{\Gamma_2} s^{2-1} e^{-\gamma s} ds \quad (7)$$

Based on the formulas above (6) and (7) we can rewrite the hierarchical representation model as follows:

$$y|\beta, \sigma, m_\tau, v \sim N(m_\tau(x_i^T \beta) + \delta v, \theta \sigma u)$$

$$, m_\tau|x_i, \beta, \partial \sim GP(0, E)$$

$$\beta_\tau | s \sim \prod_{j=1}^p \text{uniform}(-\sqrt{s_j}, \sqrt{s_j}), \quad (8)$$

$$u_i \sim N(0, 1) \quad , \quad v_i \sim \frac{1}{\sigma} \exp\left(-\frac{v_i}{\sigma}\right)$$

$$, s_j|\gamma \sim \prod_{j=1}^p \text{Gamma}(2, \gamma)$$

$$\gamma \sim \text{Gamma}(a_\gamma, b_\gamma),$$

$$\sigma \sim \text{Inv. Gamma}(a_\sigma, b_\sigma),$$

$$\partial \sim \text{Inv. Gamma}(a_\partial, b_\partial)$$

For all parameters the posterior can be given as:

$$p(m_n, \beta, v, s, \gamma, \partial, \sigma/y)$$

$$\propto \{(\det[D^{-1}])\}^{-\frac{1}{2}} \exp\left\{-\frac{(y-m_n-\delta v)^T D^{-1}(y-m_n-\delta v)}{2}\right\} \times$$

$$\det[E_n]^{-\frac{1}{2}} \exp\left\{-\frac{m_n^T E_n^{-1} m_n}{2}\right\} \times$$

$$\prod_{j=1}^p \frac{1}{2s} \frac{\gamma^2}{\Gamma_2} s^{2-1} e^{-\gamma s} \times \prod_{i=1}^n \frac{1}{\sigma}$$

$$\exp\left(-\frac{v_i}{\sigma}\right) \times \prod_{j=1}^p \frac{b^a}{\Gamma_a} s^{a-1} e^{-bs} \times$$

$$\prod_{j=1}^p (\gamma_j)^{a_\gamma-1} \exp(-b_\gamma \gamma_j) \times$$

$$(\partial)^{-a_\partial-1} \exp\left(-\frac{b_\partial}{\partial}\right) \times (\sigma)^{-a_\sigma-1}$$

$$\exp\left(-\frac{b_\sigma}{\sigma}\right) \quad (9)$$

Where a, b, c and d are hyperparameters,

here $y_i = (y_1, y_2, \dots, y_n)^T$ $v_n = (v_1, v_2, \dots, v_n)^T$, $D^{-1} = \delta \sigma \text{diag}(v_1, v_2, \dots, v_n)$, and $m_n = (m_1, m_2, \dots, m_n)^T =$

$(m(x_1^T \beta), m(x_2^T \beta), \dots, m(x_n^T \beta))^T$, E is the covariance matrix with dimension $(n \times n)$

and elements $E(\dots)$ given in Equation

$E(x_1^T \beta, x_1^T \beta) = \partial \exp \exp \{-\left(x_i - x_j\right)^T \beta \beta^T (x_i - x_j)\}$. For the parameters and

potential variables, the conditional posterior distributions will simply be derived via a

Gibbs sampler algorithm for the Bayesian single index quantile regression. In the

simulation study the kernel matrix E is almost be singular, in order that m_τ is

integrated to avert the singularity problem

and therefore the inverse is calculated for $(E + D)$ the matrix (Hu et al. (2013) and

Zhao and Lian (2015)).The conditional posterior densities of all the parameters and

variables, excepting for β and ∂ , are

common distributions. The conditional distributions utilized in the sampling are given below. m_τ is marginalized when look considering the posterior conditional distribution of β and ∂ . For the τ th quantile an easy and efficient Gibbs sampler can be shown, works as follows:

Sampling β

$$1 - \pi(\beta|m_n, v_i, \gamma, \sigma, s, \partial, y) \propto \pi(y|v_n, \sigma, m_n) \times \pi(m_n|\beta, \partial) \times \pi(\beta|s)$$

Full conditional posterior distributions

$$\propto \exp\left\{-\frac{(y-\delta v_i)^T(D+E_n)^{-1}(y-\delta v_i)}{2}\right\} \times (\det[E_n + D])^{-\frac{1}{2}} \times \prod_{j=1}^p \frac{1}{2s}$$

Updating β The full conditional distribution of β_j is truncated normal With mean $\bar{\beta}_j$ and variance S_β^2 where

$$\bar{\beta}_j = S_\beta^2 \sum_{i=1}^n \frac{\sigma}{2\theta v_i} (y_i - m_\tau(x_i^T \beta_\tau) - \delta v_i)$$

$$I\{|\beta_j| < -\sqrt{s}\}, \quad \text{and } S_\beta^2 = \left(\sum_{i=1}^n \frac{\sigma}{2\theta v_i}\right)^{-1}$$

Sampling m_n

$$2 - \pi(m_n|\beta, v_i, \sigma, s, \gamma, \partial, y) \propto \pi(y|v_i, m_n, \sigma, \beta) \times \pi(m_n|\beta, \partial)$$

Full conditional posterior distributions

$$\pi(m_n|\beta, v_i, \sigma, \partial, y) \sim N(\mu_n, \Sigma_n)$$

Sampling σ

$$3 - \pi(\sigma|\beta, m_n, v_i, \gamma, \partial, y) \propto \pi(y|v_i, m_n, \sigma) \times \pi(v_i/\sigma) \times \pi(\sigma)$$

Full conditional posterior distributions

$$\alpha_\sigma = \frac{-3n}{2} - \alpha_\sigma$$

$$\vartheta_\sigma = \left\{ \sum_{i=1}^n \frac{(y-m_n-\delta v_i)^T(y-m_n-\delta v_i)}{2\theta v_i} + v_i + b_\sigma \right\} \pi(\sigma|m_n, \beta_n, v_i, \gamma, y) \sim IG(\alpha_\sigma, \vartheta_\sigma),$$

Sampling γ

$$4 - \pi(\gamma|\sigma, s, m_n, v_i, \partial, y) \propto \pi(\beta/s) \times \pi(\gamma)$$

Full conditional posterior distributions

$$\pi(\gamma|s, \beta) \sim IG\left(a_\gamma + 2p, b_\gamma + \sum_{j=1}^p s_j\right)$$

Sampling s

$$5 - \pi(s|m_\tau, v_i, \gamma, \sigma, \beta, \partial, y) \propto \pi(\beta/s) \times \pi(s)$$

Full conditional posterior distributions

$$\pi(s|m_\tau, \gamma, \beta, y) \sim \text{Gamma}(p - pa, \sum_{j=1}^p (\gamma + b))$$

Sampling ∂

$$6 - \pi(\partial|\beta, v_n, \sigma, y) \propto \pi(y|v_n, \sigma, m_n) \times \pi(m_n|\beta, \partial) \times \pi(\partial)$$

$$\propto \{(\det[D^{-1}])^{-\frac{1}{2}} \exp\left\{-\frac{(y-m_n-\delta v_i)^T D^{-1}(y-m_n-\delta v_i)}{2}\right\}\}$$

$$\times \det[E_n]^{-\frac{1}{2}} \exp\left\{-\frac{m_n^T E_n^{-1} m_n}{2}\right\} \times$$

$$(\partial)^{-a\partial-1} \exp\left(-\frac{b\partial}{\partial}\right)$$

Full conditional posterior distributions

We use the Metropolis algorithm for sampling ∂ .

Sampling v_i

The full conditional distribution for v_i is a generalized inverse Gaussian distribution (GIG),

$$7 - \pi(v_i | \sigma, \beta, m_n, \gamma, \partial, y) \alpha$$

$$\pi(y | \sigma, m_n, v_i) \times \pi(v_i)$$

Full conditional posterior distributions

$$\pi(v_i | \sigma, \beta, m_n, \gamma, \partial, y) =$$

$$\pi(v_i | \sigma, m_n, y) \sim GIG\left(\frac{1}{2}, \sqrt{\frac{(y-m_n)^2}{\theta\sigma}}, \sqrt{\frac{\delta^2}{\theta\sigma}} + \sqrt{\frac{2}{\sigma}}\right)$$

where the likelihood density function of $GIG(\rho, w, z)$ is

$$f(\rho, w, z) =$$

$$\frac{\left(\frac{z}{w}\right)^\rho}{2K_\rho(wz)} x^{\rho-1} \exp\left\{-\frac{1}{2}(w^2 x^{-1} + z^2 x)\right\}$$

$x > 0, -\infty < \rho < \infty, w \geq 0, z \geq 0$, and K_ρ is the changed Bessel function of the third kind (Barndorff-Nielsen and Shephard 2001).

3. Simulation

The simulation study considered in this section to evaluate our suggested method Bayesian single index quantile regression with a uniform scale mixture (BUSIQR) and compared it's with other existing methods Bayesian quantile regression (BQR) and Bayesian single index quantile regression (BSIQR) and Bayesian lasso quantile regression (BLQR). Here, simulation example is considered in this study. This example was utilized by Wu et al. (2010) and have been very popular often it employed by most researchers whose study single index model conditional to quantile regression [See, (Hu et al. (2013), Alkenani and Yu (2013), Lv et al. (2014), Zhao and Lian (2015), Kuruwita (2015) and Alshaybawee et al. (2016))], but with some differences. Methods performance is evaluated based on the median mean absolute deviations referred to as (MMAD), bias and standard deviation. These measurements are calculated based on 100 replications. We have run the proposed Bayesian algorithm for 10000 iteration and the first 2000 iterations was burn in. Computations were done based on R package.

Simulation Example

In this example we have considered three samples with different size 50, 100 and 200 these samples are generated based on the following model:

$$y = m(x_i^T \beta) + \sqrt{(\sin(x_i^T \beta) + 1)} \epsilon$$

where $m(h) = 10 \sin(0.75 h)$,

where $x_i, i = 1, \dots, 6$ are i.i.d normal distribution with $N \sim [0, (0,25)^2]$, $\beta = (\beta_1, \dots, \beta_6)^T = \frac{1}{\sqrt{2}} (0,1,0,0,1,0)^T$, and the error term (ϵ) is standard normal distribution (SND),

the models were supply with three various quantiles, $T = \{ 0.25, 0.50, 0.75 \}$.

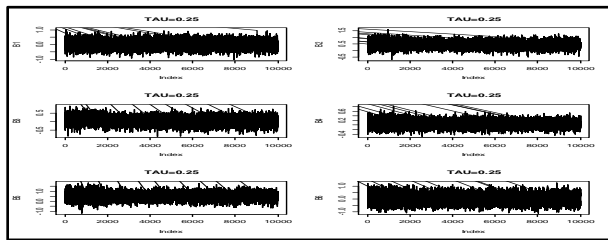


Figure (1) shows the trace plots for the estimated parameters at quantile 0.25 for simulation Example with 100 sample size.

We have display the trace plots in Figures (1) and (2), to check the rapprochement of the parameter estimates for the chain of the parameters β at two quantiles (0.25, 0.75). Based on the trace plots, we have seen clear that the chains of our suggested algorithm given a good performance.

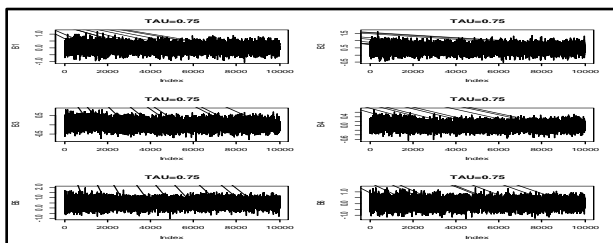


Figure (2) shows the trace plots for the estimated parameters at quantile 0.75 for simulation Example with 100 sample size.

The standard deviation parameter estimates of this simulation example for the samples 50,100 and 200 and based on 100 replications are summarized in Figures 1, 2 and 3 of the parameter estimates of each quantile.

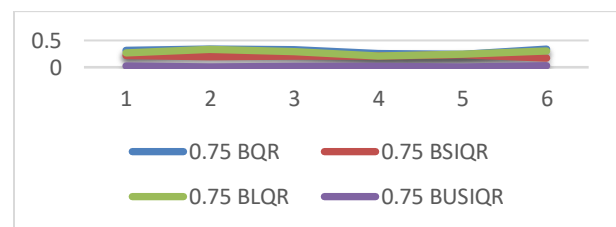
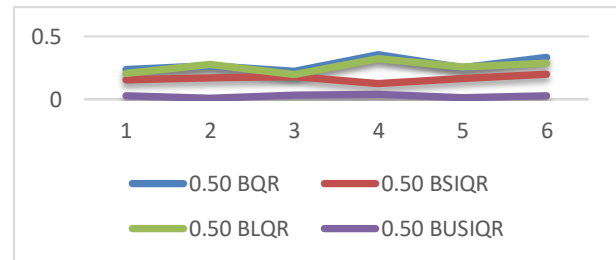
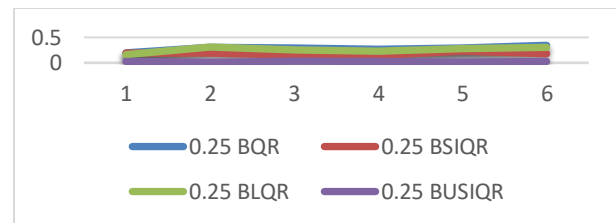


Figure (3) Shows the SD of the estimated parameters for the quantiles (0.25, 0.50, 0.75) in 100 replications when the sample size is 50.

From Figure (3), We can view that the proposed parameter estimates the BUSIQR method better than other methods, whereas this method have the smallest SD compare to the others. The small values of SD mean that the parameter

estimates are homogeneous. On the contrary, the BQR and BLQR methods are got the largest values of SD. Also we can see that the BSIQR method are got SD smaller than the BLQR and BQR method.

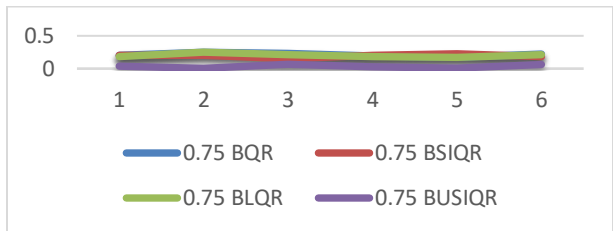
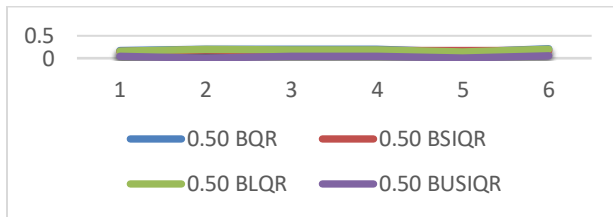
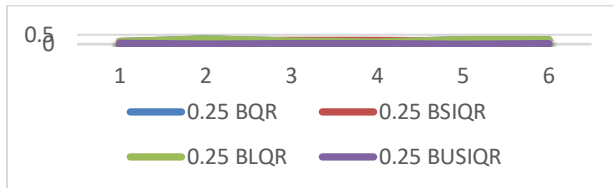


Figure (4) Shows the SD of the estimated parameters for the quantiles (0.25, 0.50, 0.75) in 100 replications when the sample size is 100.

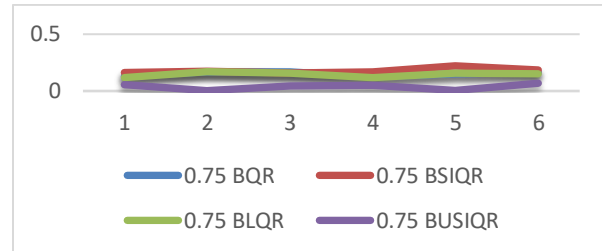
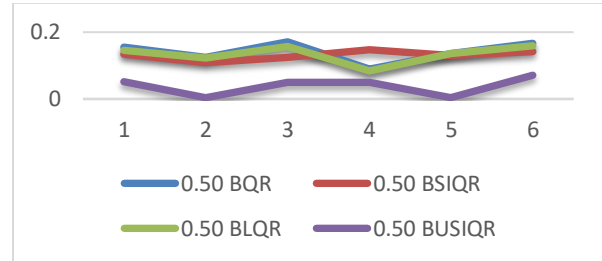
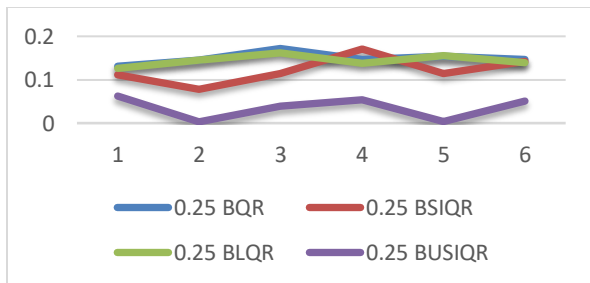


Figure (5) Shows the SD of the estimated parameters for the quantiles (0.25, 0.50, 0.75) in 100 replications when the sample size is 200.

As same as Figure (3), Figures (4) and (5), shows that the proposed method BUSIQR get the smallest SD for most parameter estimates at the different quantiles. In addition, the BQR method get the largest SD for all parameter estimates at all quantiles and in the different sample size 50,100 and 200. Also we can see that when the sample size are increase the SD for the BLQR and BSIQR are converge.

Table (1) shows the bias for the different quantiles when the sample size is 50

τ	Methods	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$	$\hat{\beta}_6$
0.25	BQR	0.898	0.939	0.73	0.557	0.923	0.97
	BSIQR	0.096	0.033	0.255	0.184	0.377	0.213
	BLQR	0.161	0.877	0.623	0.459	0.888	0.743
	BUSIQR	0.021	0.01	0.054	0.072	0.105	0.09
0.50	BQR	1.018	1.875	0.687	0.821	2.829	0.725
	BSIQR	0.829	1.047	0.277	0.552	0.097	0.261

	BLQR	0.934	1.306	0.277	0.737	2.793	0.611
	BUSIQR	0.217	0.228	0.069	0.112	0.086	0.094
0.75	BQR	0.145	1.815	0.428	0.571	1.901	0.866
	BSIQR	0.256	0.986	0.206	0.216	1.062	0.271
	BLQR	0.144	1.763	0.327	0.369	1.867	0.651
	BUSIQR	0.126	0.386	0.019	0.087	0.414	0.173

Table (2) shows the bias for the different quantiles when the sample size is 100

τ	Methods	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$	$\hat{\beta}_6$
0.25	BQR	0.252	1.081	0.727	1.327	0.881	1.065
	BSIQR	0.118	0.849	0.533	0.720	0.702	0.926
	BLQR	0.149	0.949	0.726	1.031	0.854	1.058
	BUSIQR	0.023	0.338	0.279	0.368	0.200	0.412
0.50	BQR	0.350	0.914	1.044	1.077	0.890	1.136
	BSIQR	0.178	0.690	0.898	0.461	0.463	0.801
	BLQR	0.211	0.885	1.045	0.994	0.871	0.931
	BUSIQR	0.109	0.170	0.125	0.244	0.132	0.319
0.75	BQR	0.523	1.334	0.954	1.329	0.855	1.006
	BSIQR	0.187	0.881	0.221	0.954	0.421	0.874
	BLQR	0.433	0.903	0.851	1.029	0.838	0.962
	BUSIQR	0.132	0.278	0.141	0.312	0.103	0.262

Table (3) shows the bias for the different quantiles when the sample size is 200

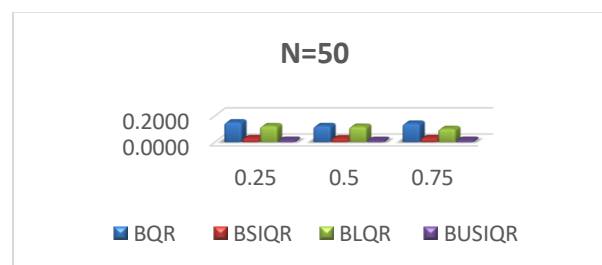
τ	Methods	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$	$\hat{\beta}_6$
0.25	BQR	1.404	0.827	1.033	0.915	0.865	1.026
	BSIQR	0.904	0.625	0.546	0.781	0.521	0.918
	BLQR	1.128	0.814	0.828	0.813	0.752	1.004
	BUSIQR	0.302	0.272	0.245	0.202	0.188	0.226
0.50	BQR	1.255	0.856	1.214	1.037	0.886	1.019
	BSIQR	0.943	0.717	0.837	0.901	0.624	0.767
	BLQR	1.104	0.841	0.911	0.935	0.875	1.018
	BUSIQR	0.283	0.367	0.321	0.319	0.214	0.315
0.75	BQR	1.012	0.917	1.014	0.855	0.967	0.927
	BSIQR	0.873	0.409	0.826	0.605	0.732	0.791
	BLQR	0.910	0.899	1.013	0.655	0.854	0.726
	BUSIQR	0.240	0.272	0.104	0.129	0.314	0.237

Table (1), (2) and (3) shows the bias for the proposed BUSIQR method and the existing

BQR, BLQR and BSIQR methods in the different quantiles and different samples of size. We can see clear that the performance of the proposed method was better than the other existing methods. Whereas the bias for the BUSIQR method were less than those for the others in all the different quantiles and the different samples of size. In addition, we can see that the bias for the BSIQR method is smaller than BQR and BLQR methods. So that, we can say that the performance of the BSIQR method is better than the BQR and BLQR.

Table(4) Comparison of MMAD for SQR, BSIQR, BLQR and BUSIQR methods based on 100 replications

n	τ	BQR	BSIQR	BLQR	BUSIQR
50	0.25	0.2262	0.0298	0.2233	0.0137
	0.50	0.2243	0.0285	0.2192	0.0129
	0.75	0.2273	0.0308	0.2218	0.0135
100	0.25	0.2072	0.0299	0.2058	0.0095
	0.50	0.2080	0.0270	0.2063	0.0132
	0.75	0.2116	0.0278	0.2100	0.0128
200	0.25	0.2086	0.0267	0.2080	0.0127
	0.50	0.2110	0.0272	0.2101	0.0091
	0.75	0.1851	0.0254	0.1839	0.0134





Figure(6) shows the MMAD for SQR, BSIQR, BLQR and BNSIQR methods based on 100 replications

Table (4) and Figure (6) show that the median of mean absolute deviations (MMAD) of the new proposed method BUSIQR method is got the smallest values of MMAD compare to the other three existing methods at all the quantiles and the different sample size. BQR method is got the largest MMAD values, which indicate not good performance. In addition, the MMAD of the BSIQR method is smaller than BLQR and its better.

5. Conclusion

In this paper, the Bayesian estimation and variable selection for single index conditional quantile regression model based on scale mixture uniform is proposed. We have construct a Bayesian hierarchical

model of the Lasso method and the MCMC algorithm is considered for posterior inference.

Simulation example was considered to compare our proposed method, BUSIQR, with three other existing methods, BQR, BSIQR and BLQR. Based on the results in the tables and figures of the simulation study, we have seen that the BUSIQR provides substantial improvement compare to the other methods. Therefore, we concluded that the perform of our proposed method is better than the onther existing methods.

References

- Alhamzawi, R., & Ali, H. T. M. (2018). The Bayesian adaptive lasso regression. *Mathematical biosciences*, 303, 75-82.
- Alhousseini, F. H. H. (2017). Bayesian Quantile Regression with Scalemixture of Uniform Prior Distributions. *International Journal of Pure and Applied Mathematics*, 115(1), 77-91.
- Alkenani, A., & Yu, K. (2013). Penalized single-index quantile regression. *International Journal of Statistics and Probability*, 2(3), 12.
- Alshaybawee, T., Midi, H., & Alhamzawi, R. (2017). Bayesian elastic net single index

quantile regression. *Journal of Applied Statistics*, 44(5), 853-871.

Barndorff-Nielsen, O. E., & Shephard, N. (2001). Non-Gaussian Ornstein–Uhlenbeck-based models and some of their uses in financial economics. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(2), 167-241.

Bellman, R., and Bellman, R. Adaptive Control Processes: A Guided Tour. Princeton University Press (1966).

Benoit, D. F., & Van den Poel, D. (2012). Binary quantile regression: a Bayesian approach based on the asymmetric Laplace distribution. *Journal of Applied Econometrics*, 27(7), 1174-1188.

Choi, T., Shi, J. Q., & Wang, B. (2011). A Gaussian process regression approach to a single-index model. *Journal of Nonparametric Statistics*, 23(1), 21-36.

Choy, B., Wan, W. Y., & Chan, C. M. (2009). Bayesian student-t stochastic volatility models via scale mixtures. *Wai Yin and Chan, Chun Man, Bayesian Student-T Stochastic Volatility Models Via Scale Mixtures (August 18, 2009)*.

Damien, P., Qin, Z. S., & Walker, S. G. (2000). Uniform scale mixture models with applications to variance regression.

Gramacy, R. B., & Lian, H. (2012). Gaussian process single-index models as emulators for computer experiments. *Technometrics*, 54(1), 30-41.

Gujarati, D. N., & Porter, D. C. (2003). Basic econometrics-Economic series McGraw-Hill international editions: Economic series.....

He, X., & Shi, P. (1994). Convergence rate of B-spline estimators of nonparametric conditional quantile functions. *Journal of Nonparametric Statistics*, 3(3-4), 299-308.

Hu, Y., Gramacy, R. B., & Lian, H. (2013). Bayesian quantile regression for single-index models. *Statistics and Computing*, 23(4), 437-454.

Hu, Y., Gramacy, R. B., & Lian, H. (2013). Bayesian quantile regression for single-index models. *Statistics and Computing*, 23(4), 437-454.

Hu, Y., Gramacy, R. B., & Lian, H. (2013). Bayesian quantile regression for single-index models. *Statistics and Computing*, 23(4), 437-454.

Ichimura, H. (1993). Semiparametric least squares (SLS) and weighted SLS estimation of single-index models. *Journal of Econometrics*, 58(1-2), 71-120.

Koenker, R., & Basset, G. (1978). Asymptotic theory of least absolute error

regression. *Journal of the American Statistical Association*, 73(363), 618-22.

Koenker, R., & d'Orey, V. (1994). Remark AS R92: A remark on algorithm AS 229: Computing dual regression quantiles and regression rank scores. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 43(2), 410-414.

Koenker, R., & Geling, O. (2001). Reappraising medfly longevity: a quantile regression survival analysis. *Journal of the American Statistical Association*, 96(454), 458-468.

Koenker. (2005). *Quantile Regression (Econometric Society monographs; no. 38)*. Cambridge university press.

Kozumi, H., & Kobayashi, G. (2011). Gibbs sampling methods for Bayesian quantile regression. *Journal of statistical computation and simulation*, 81(11), 1565-1578.

Kuruwita, C. N. (2016). Non-iterative Estimation and Variable Selection in the Single-index Quantile Regression Model. *Communications in Statistics-Simulation and Computation*, 45(10), 3615-3628.

Lv, Y., Zhang, R., Zhao, W., & Liu, J. (2014). Quantile regression and variable selection for the single-index model. *Journal of Applied Statistics*, 41(7), 1565-1577.

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

Mosteller, F., & Tukey, J. W. (1977). Data analysis and regression: a second course in statistics. *Addison-Wesley Series in Behavioral Science: Quantitative Methods*.

Walker, S. (1997). On scale mixtures of uniform distributions and the latent weighted least squares method.

Weisberg, S. (2005). *Applied linear regression* (Vol. 528). John Wiley & Sons.

Wu, T. Z., Yu, K., & Yu, Y. (2010). Single-index quantile regression. *Journal of Multivariate Analysis*, 101(7), 1607-1621.

Wu, T. Z., Yu, K., & Yu, Y. (2010). Single-index quantile regression. *Journal of Multivariate Analysis*, 101(7), 1607-1621.

Wu, Y., & Liu, Y. (2009). Variable selection in quantile regression. *Statistica Sinica*, 19(2), 801.

Yu, K., & Jones, M. C. (1998). Local linear quantile regression. *Journal of the American statistical Association*, 93(441), 228-237.

Yu, Y., & Ruppert, D. (2002). Penalized spline estimation for partially linear single-index models. *Journal of the American Statistical Association*, 97(460), 1042-1054.



Available online at <http://qu.edu.iq/journalsc/index.php/JOPS>

ISSN 2411-3514 ONLINE

Al-Qadisiyah Journal Of Pure Science (QJPS)

ISSN 1997-2490 PRINTED

Vol. 24, No. 4, pp 9–23, Year 2019

Zhao, K., & Lian, H. (2015). Bayesian Tobit quantile regression with single-index models. *Journal of Statistical Computation and Simulation*, 85(6), 1247-1263.