应用概率统计　第 35 卷
第 2 期　2019 年 4 月

Chinese Journal of Applied Probability and Statistics
Apr., 2019, Vol. 35, No. 2, pp. 178-192
doi: 10.3969/j.issn.1001-4268.2019.02.006

Gibbs Sampler Algorithm of Bayesian Weighted Composite Quantile Regression *

TIAN Yuzhu[1,2⋆]　WANG Liyong[1]　WU Xinqian[2]　TIAN Maozai[3]

([1] *School of Statistics and Mathematics, Central University of Finance and Economics, Beijing, 100081, China*)

([2] *School of Mathematics and Statistics, Henan University of Science and Technology, Luoyang, 471023, China*)

([3] *Center for Applied Statistics of Renmin University of China, Beijing, 100872, China*)

Abstract: Most regression modeling is based on traditional mean regression which results in non-robust estimation results for non-normal errors. Compared to conventional mean regression, composite quantile regression (CQR) may produce more robust parameters estimation. Based on a composite asymmetric Laplace distribution (CALD), we build a Bayesian hierarchical model for the weighted CQR (WCQR). The Gibbs sampler algorithm of Bayesian WCQR is developed to implement posterior inference. Finally, the proposed method are illustrated by some simulation studies and a real data analysis.

Keywords: CALD; Markov chain Monte Carlo (MCMC) algorithm; quantile regression; Gibbs sampler; hierarchical model; posterior inference

2010 Mathematics Subject Classification: 62F15; 62J05

Citation: TIAN Y Z, WANG L Y, WU X Q, et al. Gibbs sampler algorithm of Bayesian weighted composite quantile regression [J]. Chinese J Appl Probab Statist, 2019, 35(2): 178–192.

§1. Introduction

The linear regression models are common in statistical analysis for establishing the relationship between a response variable and covariates, and have been frequently applied to a variety of scientific fields such as medicine, finance, economics and environmental science. A general linear regression model can be specified as

$$y_i = x_i^{\mathsf{T}}\beta + \varepsilon_i, \qquad i = 1, 2, \ldots, n. \tag{1}$$

*The project was partly supported by the China Postdoctoral Science Foundation (Grant No. 2017M610156), the National Natural Science Foundation of China (Grant No. 11501167) and the Young Academic Leaders Project of Henan University of Science and Technology (Grant No. 13490008).

⋆Corresponding author, E-mail: pole1999@163.com.

Received November 21, 2017. Revised June 22, 2018.

where y_i is the i-th observation, x_i is a $p \times 1$ dimensional covariate, ε_i is the error term.

Model (1) is commonly estimated using the least square estimation (LSE) or maximum likelihood estimation (MLE). For linear regression models with non-normal errors or outliers, however, estimation results of the LSE or MLE are expected to be sensitive and estimation efficiency may be significantly decreased. As an alternative, median regression is expected to produce more robust estimation results. However, median regression may not be the best choice in some cases, especially for some non-normal error distributions. Zou and Yuan [1] pointed out that quantile regression may result in an arbitrarily small relative efficiency compared to the LSE. They proposed the CQR estimation which can pull information of multiple quantiles together to achieve estimation efficiency gain over a single quantile regression including median regression. Many authors have conducted a large number of studies work on the CQR. For instance, Kai et al. [2] studied local polynomial CQR for nonparametric regression models. They showed that the CQR performed better over both the LSE and median regression. Kai et al. [3] implemented the CQR estimation for semiparametric varying coefficients partially linear models. Jiang et al. [4] investigated CQR nonlinear models. Tang et al. [5] considered the CQR and variable selection procedures for random censored models. Recently, Zhao and Xiao [6] argued that the equal-weighted CQR is not a best way of using distributional information of quantile regressions. They developed a WCQR model by forcing different weight on different component of quantile regressions. Additionally, Jiang and Li [7] presented the penalized WCQR for the linear regression model under heavy-tailed autocorrelated errors. Tian et al. [8] studied the WCQR of linear regression models using the EM algorithm. Although a lot of work has been done to study the CQR, there are still limited literatures to conduct CQR from a Bayesian viewpoint. Bayesian CQR is an interesting issue in CQR literstures. For example, Huang and Chen [9] discussed Bayesian CQR based on a mixture asymmetric Laplace distribution, Alhamzawi [10] considered Bayesian CQR via a skewed Laplace distribution. In this paper, we develop the Bayesian WCQR based on the CALD which can address a weighted version of Bayesian CQR.

The remainder of this paper is organized as follows. In Section 2, we review Bayesian quantile regression. In Section 3, we discuss Bayesian WCQR and derive the full conditional posterior distributions of all parameters via Gibbs sampling algorithm. In Section 4, some simulations are conducted to illustrate the sample performance of the developed method. In Section 5, a data analysis is provided to illustrate the proposed procedure. In Section 6, some conclusions are drawn.

§2. Reviews

As to model (1), suppose that the τ-th quantile of ε_i is zero. Then, the τ-th conditional quantile of the response y_i can be specified as $Q_\tau(y_i \mid x_i) = x_i^{\mathsf{T}}\beta_\tau$, $0 < \tau < 1$. From Koenker and Bassett [11], we can obtain the τ-th quantile estimator $\widehat{\beta}_\tau$ of regression coefficient by minimizing the following objective loss function

$$\frac{1}{n}\sum_{i=1}^{n}\rho_\tau(y_i - x_i^{\mathsf{T}}\beta_\tau), \tag{2}$$

where $\rho_\tau(u) = u[\tau - I(u < 0)]$ is the check function. For quantile regression estimation, Yu and Moyeed [12] proposed an equivalent Bayesian estimator based on the asymmetric Laplace error distribution. The asymmetric Laplace distribution $\mathrm{ALD}(\mu,\sigma,\tau)$ has its probability density function as follows

$$f(y \mid \mu,\sigma,\tau) = \frac{\tau(1-\tau)}{\sigma}\exp\Big[-\rho_\tau\Big(\frac{y-\mu}{\sigma}\Big)\Big], \tag{3}$$

where μ is the location, σ is the scale, and $0 < \tau < 1$ is the skewness. Yu and Moyeed [12] argued that their empirical results are robust by forcing the ALD on errors even if it is a misspecification of the true errors. Recently, Sriram et al. [13] provided a justification for this usage by proving posterior consistency under the ALD misspecification. Bayesian quantile regression has attracted a lot of attention. For instance, Kottas and Gelfand [14] studied Bayesian semi-parametric median regression modeling by employing mixture model errors, Dunson and Taylor [15] considered approximate Bayesian inference of quantile regression, Kottas and Krnjajić [16] discussed Bayesian semiparametric quantile regression, Reich et al. [17] considered Bayesian quantile regression for independent and clustered data. Kozumi and Kobayashi [18] presented an efficient Gibbs sampling algorithm by utilizing a hierarchical representation of the ALD. This representation can be displayed as follows. Suppose $y \sim \mathrm{ALD}(\mu,\sigma,\tau)$, y can be decomposed as the mixture representation: $y = \mu + \theta_1 \upsilon + \sqrt{\theta_2\sigma\upsilon}\cdot e$, where $\theta_1 = (1-2\tau)/[\tau(1-\tau)]$, $\theta_2 = 2/[\tau(1-\tau)]$, $\upsilon \sim \mathrm{Exp}(1/\sigma)$, $e \sim \mathrm{N}(0,1)$, υ and e are independent with each other. The biggest advantage of this representation is that it can transform the ALD to the conditional normal distribution. Further, Reich et al. [19] considered Bayesian spatial quantile regression models using this mixture representation, Kobayashi and Kozumi [20] considered Bayesian quantile regression of censored dynamic panel data, Zhao and Lian [21] considered Bayesian Tobit quantile regression for single-index models, Tian et al. [22] studied Bayesian joint quantile regression for mixed effects models with censoring and errors in covariates.

§3. Bayesian WCQR

3.1 The CALD

The CQR was firstly proposed by Zou and Yuan [1] to estimate linear regression. CQR can combine multiple quantiles together and provide more robust estimation than traditional mean regression and median regression, especially for non-normal errors. For model (1), suppose Q be a set of quantiles under consideration: $Q=\{\tau_1,\tau_2,\ldots,\tau_K\}$, $0<\tau_1<\tau_2<\cdots<\tau_K$. The CQR estimator of regression coefficient vector β can be obtained by minimizing the following objective loss function

$$(\widehat{\alpha}_1,\widehat{\alpha}_2,\ldots,\widehat{\alpha}_K,\widehat{\beta}^{\mathrm{CQR}})=\arg\min_{\alpha_1,\alpha_2,\ldots,\alpha_K,\beta}\sum_{i=1}^{n}\sum_{k=1}^{K}\rho_{\tau_k}(y_i-x_i^{\mathsf{T}}\beta-\alpha_k), \tag{4}$$

where α_k is the τ_k-th quantile of the error ε_i which satisfies monotonicity: $\alpha_1<\alpha_2<\cdots<\alpha_K$. Generally, one can let $\tau_k=k/(K+1)$, $k=1,2,\ldots,K$, where K is the quantile level size. Loss function (4) is a combination of the objective functions of quantile regression models. When $K=1$, function (4) is the loss function of median regression. In order to ensure the identifiability of constant term in linear regression model (1), we can force a constraint condition $\sum\limits_{k=1}^{K}\alpha_k=0$ on parameters $\alpha_1,\alpha_2,\ldots,\alpha_K$. Additionally, loss function (4) is piecewise linear and is hard to solve by differentiating the objective function. We will provide a Bayesian treatment based on the CALD which was firstly proposed in [23]. The probability density function of the CALD is

$$h(y\,|\,\mu,\sigma)\propto\prod_{k=1}^{K}\frac{1}{\sigma_k}\exp\Big[-\rho_{\tau_k}\Big(\frac{y-\mu_k}{\sigma_k}\Big)\Big], \tag{5}$$

where $\mu=(\mu_1,\mu_2,\ldots,\mu_K)$ is the location vector, $\sigma=(\sigma_1,\sigma_2,\ldots,\sigma_K)$ is the scale vector, $\tau_1,\tau_2,\ldots,\tau_K$ are quantiles in equation (4).

Remark 1 The function $h(y\,|\,\mu,\sigma)$ is a density kernel via omitting the regularized constants. The similar thought was introduced by Sriram et al. [24] for implementing simultaneous Bayesian estimation of multiple quantiles. As showed by Sriram et al. [13], theoretically, omitting the regularizing constant does not produce large deviation for estimation.

Remark 2 For model (1), suppose y_i follows the PCALD conditionally on covariate x_i, where location parameters μ_k, $k=1,2,\ldots,K$ are taken as $\mu_k=x_i^{\mathsf{T}}\beta+\alpha_k$. It is easy to see that the MLE of regression parameters are equivalent to minimize the loss function $\sum\limits_{i}^{N}\sum\limits_{k=1}^{K}\sigma_k^{-1}\rho_{\tau_k}(y_i-\alpha_k-x_i^{\mathsf{T}}\beta)$. Compared this loss function (4), the CQR estimation is equivalent the MLE under the CALD error distributions, where σ_k^{-1}, $k=1,2,\ldots,K$ are K weights of quantile regressions. Specially, when $\sigma_1=\sigma_2=\cdots=\sigma_K$, the MLE is the common CQR estimation.

Remark 3 For the CALD, set $\mu_k = x_i^{\mathsf{T}}\beta_k + \alpha_k$, $k = 1, 2, \ldots, K$, the MLE is the simultaneous multiple quantiles estimation of model (1). In simultaneous multiple quantiles regression, a key is how to retain monotonicity of quantile estimators. In Bayesian WCQR, we only need to keep monotonicity of posterior estimators of parameters $\alpha_1, \alpha_2, \ldots, \alpha_K$. Other references of simultaneous multiple quantiles regression can refer to [24], [25] and [26].

3.2 Bayesian Inference

The likelihood function of the WCQR is not easy to maximize due to the complexity of the CALD. To solve this problem, the density function of CALD can be decomposed as follows

$$\begin{cases} g(y \,|\, \mu, \sigma, \upsilon) \propto \prod\limits_{k=1}^{K} \dfrac{1}{\sqrt{\theta_{2,k}\upsilon_k\sigma_k}} \exp\Big[-\dfrac{(y-\mu_k-\theta_{1,k}\upsilon_k)^2}{2\theta_{2,k}\sigma_k\upsilon_k}\Big], \\ \upsilon_k \sim \mathrm{Exp}\Big(\dfrac{1}{\sigma_k}\Big), \qquad k = 1, 2, \ldots, K, \end{cases} \tag{6}$$

where $\upsilon = (\upsilon_1, \upsilon_2, \ldots, \upsilon_K)$ is a latent vector, $\theta_{1,k} = (1-2\tau_k)/[\tau_k(1-\tau_k)]$, $\theta_{2,k} = 2/[\tau_k(1-\tau_k)]$, $k = 1, 2, \ldots, K$.

Suppose that the τ_k conditional quantile of the response y_i be

$$Q_{\tau_k}(y_i \,|\, x_i) = x_i^{\mathsf{T}}\beta + \alpha_k, \qquad k = 1, 2, \ldots, K,\ i = 1, 2, \ldots, n. \tag{7}$$

Using the mixture representation (6), the conditional density function of the response y_i can be represented as follows

$$\begin{cases} g(y_i \,|\, \upsilon_i, \mu_i, \sigma, \tau) \propto \prod\limits_{k=1}^{K} \dfrac{1}{\sqrt{\theta_{2,k}\sigma_k\upsilon_{ik}}} \exp\Big[-\dfrac{(y_i-\mu_{ik}-\theta_{1,k}\upsilon_{ik})^2}{2\theta_{2,k}\sigma_k\upsilon_{ik}}\Big], \\ \upsilon_{ik} \sim \mathrm{Exp}\Big(\dfrac{1}{\sigma_k}\Big), \qquad k = 1, 2, \ldots, K, \end{cases} \tag{8}$$

where $\upsilon = (\upsilon_i, \upsilon_{i+1}, \ldots, \upsilon_n)$, $\upsilon_i = \{\upsilon_{ik},\ k = 1, 2, \ldots, K\}$, $\mu_{ik} = x_i^{\mathsf{T}}\beta + \alpha_k$.

The joint hierarchical likelihood of the complete data $\{y, x, \upsilon\}$ is expressed as

$$L(y, x, \upsilon) = \prod_{i=1}^{n} \Big[g(y_i \,|\, x_i, \upsilon_i, \beta, \sigma, \alpha) \times \prod_{k=1}^{K} f(\upsilon_{ik} \,|\, \sigma_k)\Big], \tag{9}$$

where $g(y_i \,|\, x_i, \upsilon_i, \beta, \sigma)$ is the conditional density function of hierarchical model (8), $f(\upsilon_{ik} \,|\, \sigma_k)$ is the density function of exponential distribution $\mathrm{Exp}(1/\sigma_k)$.

To implement Bayesian analysis, one need to specify prior distributions for all parameters. We set priors of $\{\beta, \alpha, \sigma\}$ as follows

$$\pi(\beta, \alpha, \sigma) \propto \pi(\beta)\pi(\alpha)\pi(\sigma).$$

The prior of β is taken as the multivariate normal distribution: $\mathrm{N}(\beta_0, B_0)$, where β_0 and B_0 are known hyper parameters. For fixed K, the prior of α is taken as $\pi(\alpha) =$

$\prod_{k=1}^{K} \pi(\alpha_k)$, where $\alpha_k,\ k = 1, 2, \ldots, K$ are supposed to follow normal distributions $\mathrm{N}(\alpha_{k,0}, \varsigma_{k,0}^2)$, $\alpha_{k,0}$ and $\varsigma_{k,0}^2$ are known mean and variance. In addition, quantile parameters $\alpha_1, \alpha_2, \cdots, \alpha_K$ should satisfy the constraint: $\alpha_1 < \alpha_2 < \cdots < \alpha_K$. To remain the above monotonicity, the hyper parameters $\alpha_{k,0},\ k = 1, 2, \ldots, K$ can be set as $\alpha_{1,0} < \alpha_{2,0} < \cdots < \alpha_{K,0}$. The scale hyper parameters $\varsigma_{k,0}^2$ can be taken as the same value for simplicity. The prior of σ can be taken as $\pi(\sigma) = \prod_{k=1}^{K} \pi(\sigma_k)$, where $\sigma_k,\ k = 1, 2, \ldots, K$ are supposed to follow inverse Gamma distributions $\mathrm{IG}(c_{k,0}, d_{k,0})$ with shape parameter $c_{k,0}$ and scale parameter $d_{k,0}$. For simplicity, denote Π as the set composed of all posterior sampling variables and Π_- as complementary set of Π excluding the present sampling variables. Incorporating the above priors into the joint likelihood (9), the joint posterior density of set Π can be presented as

$$\pi(\Pi \,|\, y, x) \propto L(y, x, \upsilon) \cdot \pi(\beta)\pi(\alpha)\pi(\sigma). \tag{10}$$

Based on Gibbs sampling algorithm, we derive the full conditional posterior distributions of all parameters as follows

- ♦ $\pi(\beta \,|\, \Pi_-) \sim \mathrm{N}(\beta^*, B^*)$, where

$$\Delta_{ik} = y_i - \alpha_k - \theta_{1,k}\upsilon_{ik},$$

$$\beta^* = B^* \cdot \Big(\sum_{i=1}^{n} \sum_{k=1}^{K} \frac{x_i \Delta_{ik}}{\theta_{2,k}\sigma_k\upsilon_{ik}} + B_0^{-1}\beta_0 \Big), \qquad B^* = \Big(\sum_{i=1}^{n} \sum_{k=1}^{K} \frac{x_i x_i^{\mathsf{T}}}{\theta_{2,k}\sigma_k\upsilon_{ik}} + B_0^{-1} \Big)^{-1};$$

- ♦ $\pi(\sigma_k \,|\, \Pi_-) \sim \mathrm{IG}\Big(\frac{3n}{2} + c_{k,0},\ \sum_{i=1}^{n} \Big(\frac{e_{ik}^2}{2\theta_{2,k}\upsilon_{ik}} + \upsilon_{ik} \Big) + d_{k,0} \Big)$, where

$$e_{ik} = y_i - \alpha_k - x_i^{\mathsf{T}}\beta - \theta_{1,k}\upsilon_{ik}, \qquad k = 1, 2, \ldots, K,\ i = 1, 2, \ldots, n;$$

- ♦ $\pi(\upsilon_{ik} \,|\, \Pi_-) \sim \mathrm{GIG}\Big(\frac{1}{2},\ \frac{\eta_{ik}^2}{\theta_{2,k}\sigma_k},\ \frac{\theta_{1,k}^2 + 2\theta_{2,k}}{\theta_{2,k}\sigma_k}\Big)$, where

$$\eta_{ik} = y_i - \alpha_k - x_i^{\mathsf{T}}\beta, \qquad k = 1, 2, \ldots, K,\ i = 1, 2, \ldots, n;$$

- ♦ $\pi(\alpha_k \,|\, \Pi_-) \sim \mathrm{N}\big(\alpha_{k,0}^*,\ (\varsigma_{k,0}^2)^*\big)$, where

$$\epsilon_{ik} = y_i - x_i^{\mathsf{T}}\beta - \theta_{1,k}\upsilon_{ik}, \qquad k = 1, 2, \ldots, K,\ i = 1, 2, \ldots, n,$$

$$\alpha_{k,0}^* = (\varsigma_{k,0}^2)^* \Big(\sum_{i=1}^{n} \frac{\epsilon_{ik}}{\theta_{2,k}\sigma_k\upsilon_{ik}} + \frac{\alpha_{k,0}}{\varsigma_{k,0}^2} \Big), \qquad (\varsigma_{k,0}^2)^* = \Big(\sum_{i=1}^{n} \frac{1}{\theta_{2,k}\sigma_k\upsilon_{ik}} + \frac{1}{\varsigma_{k,0}^2} \Big)^{-1}.$$

Remark 4 $\mathrm{GIG}(\lambda, \chi, \psi)$ denote the generalized inverse Gaussian (GIG) distribution with density function

$$f(x) = \frac{\chi^{-\lambda}(\sqrt{\chi\psi})^{\lambda}}{2K_\lambda(\sqrt{\chi\psi})} x^{\lambda-1} \exp\Big[-\frac{1}{2}(\chi x^{-1} + \psi x) \Big], \qquad \chi > 0,\ \psi > 0,\ x > 0,$$

where K_λ denotes a modified Bessel function.

Remark 5 It is easy to see that $\mathrm{d}\alpha_{k,0}^{*}/\mathrm{d}\alpha_{k,0} > 0$, $k = 1, 2, \ldots, K$ which prove that $\alpha_{k,0}^{*}$ is the monotonically increasing function of $\alpha_{k,0}^{*}$. This conclusion implies that the monotonicity of hyper parameters $\alpha_{k,0}$, $k = 1, 2, \ldots, K$ can ensure the monotonicity of posterior estimates of parameters $\alpha_1, \alpha_2, \ldots, \alpha_K$. In fact, for parameters α_k, $k = 1, 2, \ldots, K$, we conduct posterior inference by making use of Gibbs posterior samples. We obtain the final estimation values by taking the average values of MCMC posterior samples which converge to the posterior mean values $\alpha_{k,0}^{*}$, $k = 1, 2, \ldots, K$. Hence, the monotonicity of hyper parameters $\alpha_{k,0}$, $k = 1, 2, \ldots, K$ can ensure the monotonicity of the final posterior estimates of parameters α_k, $k = 1, 2, \ldots, K$. Additionally, to guarantee the identifiability condition $\sum\limits_{k=1}^{K} \alpha_k = 0$, we can subtract a average value from the posterior estimates of $\alpha_1, \alpha_2, \ldots, \alpha_K$ to derive the final estimation values in MCMC algorithm.

By sampling repeatedly from the full posterior distributions of parameters $\{\beta, \alpha, \sigma\}$, we obtain a series of stationary samples $(\beta^{(1)}, \alpha^{(1)}, \sigma^{(1)}), (\beta^{(2)}, \alpha^{(2)}, \sigma^{(2)}), \ldots, (\beta^{(T)}, \alpha^{(T)}, \sigma^{(T)})$ to conduct posterior inference.

§4. Simulations

In this section, we conduct Monte Carlo simulations to illustrate the Bayesian WCQR procedure. For model (1), the true parameter is $\beta = (\beta_0, \beta_1, \beta_2, \beta_3)^{\mathsf{T}} = (0.5, 1, 2, 1)^{\mathsf{T}}$, two sample sizes $n = 50$, 100 are considered, the covariates are set as $x_i^{\mathsf{T}} = (1, x_{i1}, x_{i2}, x_{i3})$, where x_{ij}, $j = 1, 2, 3$ are generated independently from the standard normal distribution $\mathrm{N}(0, 1)$. Four error distributions are considered: standard normal distribution $(\mathrm{N}(0, 1))$, t distribution with three degrees of freedom (t_3), chi-square distribution with two degrees of freedom $(\chi^2(2))$ and standard Cauchy distribution $(\mathrm{C}(0, 1))$. Quantile level numbers $K = 1, 3, 5, 7, 9$ are considered, where the case of $K = 1$ represent median regression. The priors of all parameters are set as follows:

$$\beta_j \sim \mathrm{N}(0, 10^2), \quad j = 0, 1, 2, 3; \quad \sigma_k \sim \mathrm{IG}(0.5, 0.5), \quad \alpha_k \sim \mathrm{N}(\alpha_{k,0}, 10^2), \quad k = 1, 2, \ldots, K.$$

The hyper parameters $\alpha_{k,0}$, $k = 1, 2, \ldots, K$ are monotonously taken as K knots which divide interval $[-10, 10]$ into $K - 1$ equal subintervals. Initial value of β is set as the LSE. Initial values of scale parameters σ_k, $k = 1, 2, \ldots, K$ are taken as 1. We run 50 times repeated simulations for the Bayesian WCQR procedure and run 15 000 times Gibbs sampling algorithm for each simulation. From the trace plots of Gibbs samples, we find that MCMC chains of regression coefficients can converge to their stationary distributions rapidly. For each simulation, we remove the first 5 000 samples and preserve the tailed 10 000 samples to produce 1 000 samples with 10 thin steps for implementing posterior inference. We report posterior bias, posterior root mean square errors (RMSE) and 95% confidence interval (CI) of regression coefficients $\beta_1, \beta_2, \beta_3$ based on 50 repeated simulations. Estimation results are presented in Tables 1−2. Meanwhile, in order to compare with Bayesian WCQR results, we provide the LSE results in Table 3.

Table 1 Bayesian WCQR results for $n = 50$

		$\beta_1 = 1$			$\beta_2 = 2$			$\beta_3 = 1$		
Error	K	Bias	RMSE	95% CI	Bias	RMSE	95% CI	Bias	RMSE	95% CI
	1	0.029	0.183	(0.688,1.401)	-0.007	0.160	(1.737,2.351)	0.023	0.157	(0.770,1.295)
	3	-0.006	0.152	(0.668,1.248)	-0.005	0.173	(1.636,2.304)	0.013	0.151	(0.732,1.309)
$\mathrm{N}(0,1)$	5	-0.005	0.149	(0.722,1.289)	-0.010	0.144	(1.748,2.240)	-0.023	0.158	(0.680,1.272)
	7	0.027	0.142	(0.780,1.329)	0.013	0.143	(1.775, 2.296)	-0.002	0.257	(0.708,1.289)
	9	-0.019	0.169	(0.720,1.316)	0.009	0.153	(1.697, 2.304)	0.012	0.160	(0.727,1.289)
	1	0.047	0.219	(0.665,1.450)	0.006	0.193	(1.648,2.342)	-0.001	0.188	(0.638,1.313)
	3	-0.023	0.182	(0.587,1.267)	-0.046	0.222	(1.577,2.345)	0.046	0.190	(0.672,1.387)
t_3	5	-0.063	0.204	(0.617,1.351)	-0.003	0.203	(1.613, 2.350)	-0.007	0.217	(0.543,1.427)
	7	0.018	0.211	(0.680,1.456)	-0.071	0.187	(1.602,2.321)	-0.051	0.219	(0.623,1.300)
	9	-0.020	0.205	(0.653,1.307)	-0.045	0.169	(1.619,2.211)	-0.002	0.199	(0.653,1.391)
	1	-0.012	0.250	(0.602,1.454)	0.017	0.280	(1.600,2.688)	-0.085	0.273	(0.374,1.376)
	3	0.028	0.172	(0.717,1.309)	0.002	0.168	(1.682, 2.298)	0.039	0.247	(0.574,1.514)
$\chi^2(2)$	5	0.063	0.218	(0.702,1.553)	0.026	0.197	(1.619,2.387)	-0.003	0.184	(0.626,1.256)
	7	-0.006	0.193	(0.609,1.339)	-0.033	0.181	(1.682,2.368)	0.015	0.224	(0.642,1.430)
	9	-0.006	0.153	(0.764,1.350)	-0.013	0.188	(1.628,2.345)	0.005	0.202	(0.489,1.316)
	1	0.025	0.358	(0.216,1.709)	0.002	0.241	(1.596,2.442)	0.046	0.312	(0.551,1.594)
	3	0.019	0.350	(0.319,1.840)	-0.010	0.250	(1.575,2.429)	-0.009	0.311	(0.501,1.540)
$\mathrm{C}(0,1)$	5	0.013	0.357	(0.489,1.912)	0.027	0.320	(1.403,2.646)	0.005	0.307	(0.338,1.496)
	7	-0.078	0.355	(0.290,1.698)	0.064	0.365	(1.489,2.879)	0.082	0.288	(0.612,1.621)
	9	0.070	0.360	(0.291,1.823)	-0.044	0.355	(1.124,2.385)	-0.026	0.357	(0.332,1.616)

From Tables 1−3, we see Bayesian WCQR performs better consistently for different error distributions and different quantile level number K. For most of cases, Bayesian WCQR for $K = 3, 5, 7, 9$ illustrate better estimates with smaller RMSEs and shorter confidence intervals than the results for $K = 1$ (median regression) and the traditional LSE, especially for small sample size ($n = 50$) and non-normal errors (such as $\chi^2(2)$ and $\mathrm{C}(0,1)$). Compared to the traditional estimation methods, Bayesian WCQR behaves more robust for relatively bigger K and non-normal errors. For $\mathrm{C}(0,1)$ error, the Bayesian WC-QR presents more advantages than the LSE. For same K, Bayesian WCQR results under $\mathrm{N}(0,1)$ error outperform those results under t_3, $\chi^2(2)$ and $\mathrm{C}(0,1)$ errors. Additionally, the

Table 2 Bayesian WCQR results for $n = 100$

		$\beta_1 = 1$			$\beta_2 = 2$			$\beta_3 = 1$		
Error	K	Bias	RMSE	95% CI	Bias	RMSE	95% CI	Bias	RMSE	95% CI
	1	-0.007	0.117	(0.735,1.208)	-0.026	0.110	(1.715,2.163)	0.011	0.111	(0.800,1.217)
	3	-0.019	0.108	(0.745,1.163)	-0.012	0.099	(1.829,2.182)	-0.008	0.098	(0.802,1.161)
N(0, 1)	5	-0.006	0.085	(0.837,1.121)	-0.012	0.093	(1.814,2.146)	0.007	0.121	(0.759,1.227)
	7	-0.013	0.106	(0.848,1.207)	0.028	0.115	(1.857,2.224)	0.000	0.096	(0.797,1.141)
	9	0.023	0.115	(0.821,1.258)	-0.001	0.103	(1.779,2.149)	-0.003	0.077	(0.866,1.149)
	1	0.007	0.131	(0.745,1.256)	-0.004	0.139	(1.776,2.249)	-0.005	0.136	(0.732,1.245)
	3	0.005	0.148	(0.755,1.348)	-0.006	0.129	(1.694,2.183)	-0.031	0.145	(0.683,1.197)
t_3	5	-0.007	0.122	(0.761,1.269)	0.004	0.137	(1.758,2.291)	-0.026	0.137	(0.750,1.231)
	7	-0.013	0.129	(0.688,1.203)	0.025	0.139	(1.802,2.261)	0.077	0.129	(0.870,1.262)
	9	-0.005	0.146	(0.709,1.306)	0.019	0.133	(1.770,2.251)	-0.008	0.116	(0.781,1.192)
	1	0.002	0.165	(0.743,1.315)	0.002	0.165	(1.713,2.277)	0.014	0.153	(0.764,1.267)
	3	-0.018	0.135	(0.724,1.230)	-0.010	0.130	(1.710,2.239)	0.023	0.137	(0.792,1.303)
$\chi^2(2)$	5	-0.012	0.099	(0.805,1.140)	0.011	0.114	(1.784,2.229)	0.022	0.123	(0.789,1.276)
	7	-0.002	0.106	(0.770,1.179)	0.019	0.134	(1.775,2.257)	-0.015	0.112	(0.786,1.187)
	9	-0.018	0.120	(0.760,1.182)	0.006	0.097	(1.831,2.173)	0.012	0.115	(0.801,1.232)
	1	-0.050	0.183	(0.636,1.325)	0.009	0.231	(1.792,2.345)	-0.032	0.212	(0.650,1.225)
	3	0.005	0.200	(0.679,1.439)	-0.012	0.163	(1.721,2.404)	-0.005	0.213	(0.550,1.370)
C(0, 1)	5	-0.012	0.215	(0.577,1.306)	0.031	0.222	(1.698,2.411)	-0.033	0.227	(0.551,1.498)
	7	0.004	0.239	(0.603,1.451)	-0.059	0.234	(1.507,2.336)	-0.024	0.183	(0.569,1.343)
	9	-0.020	0.200	(0.595,1.359)	-0.007	0.196	(1.637,2.360)	-0.010	0.223	(0.527,1.303)

Table 3 LSE results

		$\beta_1 = 1$			$\beta_2 = 2$			$\beta_3 = 1$		
Error	n	Bias	RMSE	95% CI	Bias	RMSE	95% CI	Bias	RMSE	95% CI
N(0, 1)	50	0.027	0.136	(0.787,1.314)	-0.018	0.157	(1.705,2.254)	0.001	0.138	(0.749,1.303)
	100	0.007	0.114	(0.832,1.246)	0.001	0.079	(1.848,2.152)	0.003	0.101	(0.807,1.190)
t_3	50	0.008	0.218	(0.611,1.390)	0.017	0.203	(1.685,2.362)	0.033	0.225	(0.647,1.388)
	100	-0.003	0.166	(0.662,1.264)	-0.009	0.164	(1.666,2.220)	0.012	0.156	(0.743,1.245)
$\chi^2(2)$	50	0.111	0.295	(0.666,1.575)	0.022	0.281	(1.534,2.533)	0.001	0.229	(0.630,1.436)
	100	-0.027	0.232	(0.575,1.414)	-0.025	0.239	(1.522,2.379)	0.046	0.194	(0.717,1.368)
C(0, 1)	50	–	–	–	–	–	–	–	–	–
	100	–	–	–	–	–	–	–	–	–

estimation efficiency of Bayesian WCQR does not always increase as K becomes bigger. From simulations, taking the size of K from 3 to 9 seems enough to derive ideal results.

§5. Real Data Analysis

In this section, we use Bayesian WCQR to evaluate Engel food expenditure data which can be found in [27]. This data set consists of 235 observations of household income and expenditure on food. Consider the linear regression model

$$y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \varepsilon_i, \qquad i = 1, 2, \ldots, n, \tag{11}$$

where y_i is the i-th observation for the response variable $\log_{10}$(Expenditure), x_i is the i-th observation for covariate $\log_{10}$(Income).

We apply Bayesian WCQR method to Engel data set based on model (11). In estimation procedure, the initial values of regression parameters are set as the LSE. Priors and initial values of other parameters are taken as the same to simulations in Section 4. Ten quantile levels $K = 1, 2, 3, 4, 5, 6, 7, 8, 9, 10$ are considered. For each quantile level K, we run 15 000 times Gibbs samplers and only the latest 10 000 posterior samples are remained to produce 1 000 samples with 10 thin steps. Bayesian WCQR estimates (Est.) and estimated standard errors (St.E.), 95% CI based on 1000 posterior samples are presented in Table 4, where the case for $K = 1$ mean the result of median regression. We also provide estimation results of the LSE in Table 4.

Table 4 Bayesian CQR estimates and LME(s)

		β_0			β_1			β_2		
Method	K	Est.	St.E.	95% CI	Est.	St.E.	95% CI	Est.	St.E.	95% CI
	1	-1.854	0.688	(-3.070,-0.396)	2.268	0.464	(1.281,3.102)	-0.236	0.078	(-0.371,-0.070)
	2	-1.742	0.479	(-2.635,-0.697)	2.182	0.322	(1.469,2.798)	-0.220	0.054	(-0.324,-0.104)
	3	-1.560	0.404	(-2.317,-0.723)	2.058	0.272	(1.499,2.576)	-0.200	0.045	(-0.285,-0.105)
	4	-1.448	0.356	(-2.148,-0.719)	1.985	0.239	(1.492,2.452)	-0.188	0.040	(-0.266,-0.105)
	5	-1.378	0.313	(-2.010,-0.750)	1.936	0.211	(1.518,2.361)	-0.179	0.035	(-0.252,-0.109)
BCQR	6	-1.374	0.299	(-1.937,-0.803)	1.933	0.201	(1.542,2.315)	-0.179	0.033	(-0.243,-0.113)
	7	-1.361	0.260	(-1.873,-0.855)	1.924	0.175	(1.589,2.273)	-0.177	0.029	(-0.236,-0.121)
	8	-1.349	0.244	(-1.846,-0.881)	1.915	0.165	(1.599,2.250)	-0.175	0.027	(-0.232,-0.122)
	9	-1.328	0.237	(-1.786,-0.862)	1.901	0.159	(1.592,2.212)	-0.173	0.026	(-0.226,-0.122)
	10	-1.322	0.215	(-1.754,-0.916)	1.896	0.145	(1.626,2.196)	-0.172	0.025	(-0.222,-0.126)
LSE		-1.352	0.646	(-2.626,-0.078)	1.922	0.432	(1.070,2.774)	-0.178	0.072	(-0.320,-0.035)

From Table 4, we see that Bayesian WCQR results perform better for K quantile numbers. For most of cases, the estimated standard errors decrease and 95% CI become narrower as K becomes bigger. Bayesian WCQR result of $K > 1$ perform better consistently than the results of $K = 1$ (median regression) and the LSE. In fact, for this data set, quantile level number $K = 5$ seems enough to provide a group of more robust estimates than other K. As a illustration, Figure 1 display the trace plots of 15 000 Gibbs samplers for $K = 5$, autocorrelation plots and hist plot of 1 000 produced posterior samples of two regression coefficients. The corresponding predictive plot between income and expenditure for this case is displayed in Figure 2.

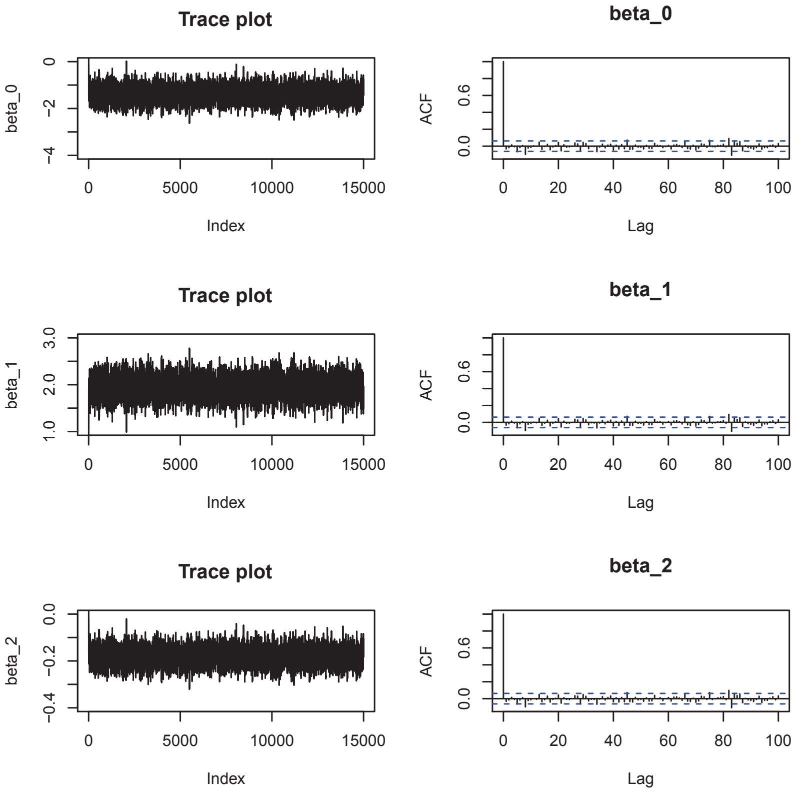

Figure 1 The left column represents the trace plots of 15 000 Gibbs samples and the right column represents autocorrelation plots of 1 000 produced samples of regression coefficients for $K = 5$

§6. Conclusion

We discuss Bayesian WCQR of linear regression based on the CALD. By using the hierarchical mixture of the CALD, we build a hierarchical Bayesian model and derive the full conditional posterior distributions of all unknown parameters using the Gibbs sampling algorithm. Some simulations are implemented and a data set is analyzed to illustrate the proposed procedure. Compared to the LSE and Bayesian median regression, Bayesian WCQR can provide more robust and accurate estimation results for a variate of error distributions, especially for non-normal error distributions. Bayesian WCQR is an

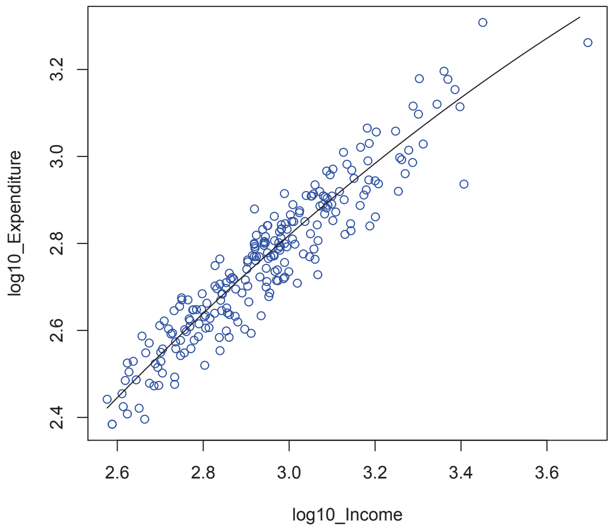

Figure 2 Fitted curve of Engel data for $K = 5$

interesting issue in CQR literatures and has not been fully discussed. There are still a lot of work to do in future research. For example, how to define a evaluation criterion to select the optimal K and how to apply the Bayesian WCQR to complex and high dimensional regression models are both interesting research subjects that can be discussed. We are continuing to pay attention to these issues in future work.

Acknowledgements The authors thank reviewers for their constructive comments and valuable suggestions which have greatly improved the paper.

Appexdix

The joint posterior distribution in equation (10) is expressed as

$$\begin{aligned}\pi(\Pi\,|\,y,x) &\propto L(y,x,\upsilon)\cdot\pi(\beta,\alpha,\sigma)\\ &\propto \prod_{i=1}^{n}\Big[g(y_i\,|\,x_i,\upsilon_i,\beta,\sigma,\alpha)\times\prod_{k=1}^{K}\frac{1}{\sigma_k}\exp\Big(-\frac{\upsilon_{ik}}{\sigma_k}\Big)\Big]\cdot\pi(\beta)\pi(\alpha)\pi(\sigma)\\ &\propto \prod_{i=1}^{n}\prod_{k=1}^{K}\Big\{\frac{1}{\sqrt{\theta_{2,k}\sigma_k\upsilon_{ik}}}\exp\Big[-\frac{(y_i-x_i^{\mathsf T}\beta-\alpha_k-\theta_{1,k}\upsilon_{ik})^2}{2\theta_{2,k}\sigma_k\upsilon_{ik}}\Big]\times\frac{1}{\sigma_k}\exp\Big(-\frac{\upsilon_{ik}}{\sigma_k}\Big)\Big\}\\ &\quad\cdot \mathrm{N}(\beta_0,B_0)\cdot\prod_{k=1}^{K}\mathrm{N}(\alpha_{k,0},\varsigma_{k,0}^2)\cdot\prod_{k=1}^{K}\mathrm{IG}(c_{k,0},d_{k,0}),\end{aligned}$$

where Π denotes the set composed of all posterior sampling variables, Π_- denotes the complementary set of Π excluding the present sampling variables, $\mathrm{N}(\beta_0,B_0)$ denotes the probability density function (pdf) of multivariate normal distribution with mean β_0 and variance-covariate matrix B_0, $\mathrm{IG}(c_{k,0}, d_{k,0})$ denotes the pdf of inverse Gamma distribution with shape parameer $c_{k,0}$ and scale parameter $d_{k,0}$.

Based on Gibbs sampling algorithm on the joint posterior distribution $\pi(\Pi\,|\,y,x)$, the full conditional posterior distributions of all parameters can be expressed as follows

♦
$$\begin{aligned}\pi(\beta\,|\,\Pi_-) &\propto \exp\Big[-\sum_{i=1}^{n}\sum_{k=1}^{K}\frac{(y_i - x_i^{\mathsf T}\beta - \alpha_k - \theta_{1,k}\upsilon_{ik})^2}{2\theta_{2,k}\sigma_k\upsilon_{ik}}\Big]\\ &\quad\cdot\exp\Big[-\frac{1}{2}(\beta-\beta_0)^{\mathsf T}B_0^{-1}(\beta-\beta_0)\Big]\\ &\propto \exp\Big[-\sum_{i=1}^{n}\sum_{k=1}^{K}\frac{(\beta^{\mathsf T}x_i - \Delta_{ik})^2}{2\theta_{2,k}\sigma_k\upsilon_{ik}}\Big]\cdot\exp\Big[-\frac{1}{2}(\beta-\beta_0)^{\mathsf T}B_0^{-1}(\beta-\beta_0)\Big]\\ &\sim \mathrm{N}(\beta^*, B^*),\end{aligned}$$

where $\Delta_{ik} = y_i - \alpha_k - \theta_{1,k}\upsilon_{ik}$, $k = 1,2,\ldots,K$, $i = 1,2,\ldots,n$,

$$\beta^* = B^*\cdot\Big(\sum_{i=1}^{n}\sum_{k=1}^{K}\frac{x_i\Delta_{ik}}{\theta_{2,k}\sigma_k\upsilon_{ik}} + B_0^{-1}\beta_0\Big),\qquad B^* = \Big(\sum_{i=1}^{n}\sum_{k=1}^{K}\frac{x_i x_i^{\mathsf T}}{\theta_{2,k}\sigma_k\upsilon_{ik}} + B_0^{-1}\Big)^{-1};$$

♦
$$\begin{aligned}\pi(\sigma_k\,|\,\Pi_-) &\propto \prod_{i=1}^{n}\Big\{\frac{1}{\sqrt{\sigma_k}}\exp\Big[-\frac{(y_i - x_i^{\mathsf T}\beta - \alpha_k - \theta_{1,k}\upsilon_{ik})^2}{2\theta_{2,k}\sigma_k\upsilon_{ik}}\Big]\times\frac{1}{\sigma_k}\exp\Big(-\frac{\upsilon_{ik}}{\sigma_k}\Big)\Big\}\\ &\quad\cdot\Big(\frac{1}{\sigma_k}\Big)^{c_{k,0}+1}\exp\Big(-\frac{d_{k,0}}{\sigma_k}\Big)\\ &\propto \Big(\frac{1}{\sigma_k}\Big)^{3n/2+c_{k,0}+1}\exp\Big\{-\frac{1}{\sigma_k}\Big[\sum_{i=1}^{n}\Big(\frac{e_{ik}^2}{2\theta_{2,k}\upsilon_{ik}} + \upsilon_{ik}\Big) + d_{k,0}\Big]\Big\}\\ &\sim \mathrm{IG}\Big(\frac{3n}{2} + c_{k,0},\ \sum_{i=1}^{n}\Big(\frac{e_{ik}^2}{2\theta_{2,k}\upsilon_{ik}} + \upsilon_{ik}\Big) + d_{k,0}\Big),\end{aligned}$$

where $e_{ik} = y_i - \alpha_k - x_i^{\mathsf T}\beta - \theta_{1,k}\upsilon_{ik}$, $k = 1,2,\ldots,K$, $i = 1,2,\ldots,n$;

♦
$$\begin{aligned}\pi(\upsilon_{ik}\,|\,\Pi_-) &\propto \frac{1}{\sqrt{\upsilon_{ik}}}\exp\Big[-\frac{(\eta_{ik} - \theta_{1,k}\upsilon_{ik})^2}{2\theta_{2,k}\sigma_k\upsilon_{ik}} - \frac{\upsilon_{ik}}{\sigma_k}\Big]\\ &\propto \frac{1}{\sqrt{\upsilon_{ik}}}\exp\Big[-\frac{1}{2}\Big(\frac{\eta_{ik}^2}{\theta_{2,k}\sigma_k}\upsilon_{ik}^{-1} + \frac{\theta_{1,k}^2 + 2\theta_{2,k}}{\theta_{2,k}\sigma_k}\upsilon_{ik}\Big)\Big]\\ &\sim \mathrm{GIG}\Big(\frac{1}{2}, \frac{\eta_{ik}^2}{\theta_{2,k}\sigma_k}, \frac{\theta_{1,k}^2 + 2\theta_{2,k}}{\theta_{2,k}\sigma_k}\Big),\end{aligned}$$

where $\eta_{ik} = y_i - \alpha_k - x_i^{\mathsf T}\beta$, $k = 1,2,\ldots,K$, $i = 1,2,\ldots,n$;

♦
$$\begin{aligned}\pi(\alpha_k\,|\,\Pi_-) &\propto \exp\Big[-\sum_{i=1}^{n}\frac{(y_i - x_i^{\mathsf T}\beta - \alpha_k - \theta_{1,k}\upsilon_{ik})^2}{2\theta_{2,k}\sigma_k\upsilon_{ik}}\Big]\cdot\exp\Big[-\frac{(\alpha_k - \alpha_{k,0})^2}{2\varsigma_{k,0}^2}\Big]\\ &\propto \exp\Big[-\sum_{i=1}^{n}\frac{(\alpha_k - \epsilon_{ik})^2}{2\theta_{2,k}\sigma_k\upsilon_{ik}} - \frac{(\alpha_k - \alpha_{k,0})^2}{2\varsigma_{k,0}^2}\Big]\end{aligned}$$

$$\sim \mathrm{N}\big(\alpha_{k,0}^{*},\, (\varsigma_{k,0}^{2})^{*}\big),$$

where $\epsilon_{ik} = y_i - x_i^{\mathsf{T}}\beta - \theta_{1,k}\upsilon_{ik}$, $k = 1, 2, \ldots, K$, $i = 1, 2, \ldots, n$,

$$\alpha_{k,0}^{*} = (\varsigma_{k,0}^{2})^{*}\Big(\sum_{i=1}^{n}\frac{\epsilon_{ik}}{\theta_{2,k}\sigma_k\upsilon_{ik}} + \frac{\alpha_{k,0}}{\varsigma_{k,0}^{2}}\Big), \qquad (\varsigma_{k,0}^{2})^{*} = \Big(\sum_{i=1}^{n}\frac{1}{\theta_{2,k}\sigma_k\upsilon_{ik}} + \frac{1}{\varsigma_{k,0}^{2}}\Big)^{-1}.$$

References

[1] ZOU H, YUAN M. Composite quantile regression and the oracle model selection theory [J]. *Ann Statist*, 2008, **36(3)**: 1108–1126.

[2] KAI B, LI R Z, ZOU H. Local composite quantile regression smoothing: an efficient and safe alternative to local polynomial regression [J]. *J R Stat Soc Ser B Stat Methodol*, 2010, **72(1)**: 49–69.

[3] KAI B, LI R Z, ZOU H. New efficient estimation and variable selection methods for semiparametric varying-coefficient partially linear models [J]. *Ann Statist*, 2011, **39(1)**: 305–332.

[4] JIANG X J, JIANG J C, SONG X Y. Oracle model selection for nonlinear models based on weighted composite quantile regression [J]. *Statist Sinica*, 2012, **22(4)**: 1479–1506.

[5] TANG L J, ZHOU Z G, WU C C. Weighted composite quantile estimation and variable selection method for censored regression model [J]. *Statist Probab Lett*, 2012, **82(3)**: 653–663.

[6] ZHAO Z B, XIAO Z J. Efficient regressions via optimally combining quantile information [J]. *Econometric Theory*, 2014, **30(6)**: 1272–1314.

[7] JIANG Y L, LI H. Penalized weighted composite quantile regression in the linear regression model with heavy-tailed autocorrelated errors [J]. *J Korean Statist Soc*, 2014, **43(4)**: 531–543.

[8] TIAN Y Z, ZHU Q Q, TIAN M Z. Estimation of linear composite quantile regression using EM algorithm [J]. *Statist Probab Lett*, 2016, **117**: 183–191.

[9] HUANG H W, CHEN Z X. Bayesian composite quantile regression [J]. *J Stat Comput Simul*, 2015, **85(18)**: 3744–3754.

[10] ALHAMZAWI R. Bayesian analysis of composite quantile regression [J]. *Stat Biosci*, 2016, **8(2)**: 358–373.

[11] KOENKER R, BASSETT JR G. Regression quantiles [J]. *Econometrica*, 1978, **46(1)**: 33–50.

[12] YU K M, MOYEED R A. Bayesian quantile regression [J]. *Statist Probab Lett*, 2001, **54(4)**: 437–447.

[13] SRIRAM K, RAMAMOORTHI R V, GHOSH P. Posterior consistency of bayesian quantile regression based on the misspecified asymmetric Laplace density [J]. *Bayesian Anal*, 2013, **8(2)**: 479–504.

[14] KOTTAS A, GELFAND A E. Bayesian semiparametric median regression modeling [J]. *J Amer Statist Assoc*, 2001, **96(456)**: 1458–1468.

[15] DUNSON D B, TAYLOR J A. Approximate Bayesian inference for quantiles [J]. *J Nonparametr Stat*, 2005, **17(3)**: 385–400.

[16] KOTTAS A, KRNJAJIĆ, M. Bayesian semiparametric modelling in quantile regression [J]. *Scand J Stat*, 2009, **36(2)**: 297–319.

[17] REICH B J, BONDELL H D, WANG H X J. Flexible Bayesian quantile regression for independent and clustered data [J]. *Biostatistics*, 2010, **11(2)**: 337–352.

[18] KOZUMI H, KOBAYASHI G. Gibbs sampling methods for Bayesian quantile regression [J]. *J Stat Comput Simul*, 2011, **81(11)**: 1565–1578.

[19] REICH B J, FUENTES M, DUNSON D B. Bayesian spatial quantile regression [J]. *J Amer Statist Assoc*, 2011, **106(493)**: 6–20.

[20] KOBAYASHI G, KOZUMI H. Bayesian analysis of quantile regression for censored dynamic panel data [J]. *Comput Statist*, 2012, **27(2)**: 359–380.

[21] ZHAO K F, LIAN H. Bayesian Tobit quantile regression with single-index models [J]. *J Stat Comput Simul*, 2015, **85(6)**: 1247–1263.

[22] TIAN Y Z, LI E Q, TIAN M Z. Bayesian joint quantile regression for mixed effects models with censoring and errors in covariates [J]. *Comput Statist*, 2016, **31(3)**: 1031–1057.

[23] TIAN Y Z. Censored quantile regression for longitudinal mixed measurement error models and some related topics [D]. Beijing: Renmin University of China, 2014.

[24] SRIRAM K, RAMAMOORTHI R V, GHOSH P. Simultaneous Bayesian estimation of multiple quantiles with an extension to hierarchical models [OL]. 2012 [2012-04-27]. IIM Bangalore Research Paper No. 359. http://dx.doi.org/10.2139/ssrn.2117722.

[25] WU Y C, LIU Y F. Stepwise multiple quantile regression estimation using non-crossing constraints [J]. *Stat Interface*, 2009, **2(3)**: 299–310.

[26] BONDELL H D, REICH B J, WANG H X. Noncrossing quantile regression curve estimation [J]. *Biometrika*, 2010, **97(4)**: 825–838.

[27] KOENKER R. *Quantile Regression* [M]. Cambridge: Cambridge University Press, 2005.

贝叶斯复合分位回归的 Gibbs 抽样算法

田玉柱[1,2] 王立勇[1] 武新乾[2] 田茂再[3]

([1]中央财经大学统计与数学学院, 北京, 100081; [2]河南科技大学数学与统计学院, 洛阳, 471023)

([3]中国人民大学应用统计科学研究中心, 北京, 100872)

摘 要: 大多数基于传统均值回归的建模方法都对非正态误差表现出不稳健的估计结果. 和传统均值回归相比, 复合分位回归 (CQR) 可以产生稳健的估计. 基于一个复合反对称 Laplace 分布 (CALD), 我们建立了加权复合分位回归 (WCQR) 的贝叶斯分层模型. Gibbs 抽样算法被发展用于 WCQR 的后验推断. 最后, 我们提供了一些模拟研究和一个实际数据分析来验证所提方法.

关键词: 复合反对称 Laplace 分布 (CALD); 马尔可夫链蒙特卡洛 (MCMC) 算法; 分位回归; Gibbs 抽样; 分层模型; 后验推断

中图分类号: O212.1; O212.8