Computational Statistics and Data Analysis 64 (2013) 209–219

Contents lists available at SciVerse ScienceDirect

Computational Statistics and Data Analysis

journal homepage: www.elsevier.com/locate/csda

Conjugate priors and variable selection for Bayesian quantile regression[☆]

Rahim Alhamzawi[a,b,*], Keming Yu[b]

[a] Statistics Department, College of Administration and Economics, Al-Qadisiyah University, Al Diwaniyah, Iraq
[b] Department of Mathematics, Brunel University, Uxbridge UB8 3PH, UK

ARTICLE INFO

Article history:
Received 24 June 2011
Received in revised form 6 January 2012
Accepted 7 January 2012
Available online 24 January 2012

Keywords:
Asymmetric Laplace distribution
g-prior
Gibbs sampler
Model selection
Quantile regression

ABSTRACT

Bayesian variable selection in quantile regression models is often a difficult task due to the computational challenges and non-availability of conjugate prior distributions. These challenges are rarely addressed via either penalized likelihood function or stochastic search variable selection. These methods typically use symmetric prior distributions such as a normal distribution or a Laplace distribution for regression coefficients, which may be suitable for median regression. However, an extreme quantile regression should have different regression coefficients from the median regression, and thus the priors for quantile regression should depend on the quantile. In this article an extension of the Zellners prior which allows for a conditional conjugate prior and quantile dependent prior on Bayesian quantile regression is proposed. Secondly, a novel prior based on percentage bend correlation for model selection is also used in Bayesian regression for the first time. Thirdly, a new variable selection method based on a Gibbs sampler is developed to facilitate the computation of the posterior probabilities. The proposed methods are justified mathematically and illustrated with both simulation and real data.

1. Introduction

Since the seminal work of Koenker and Bassett (1978), quantile regression has been the subject of great theoretical interest as well as numerous practical applications in a number of fields such as econometrics, finance, biomedical studies, social sciences, and survival analysis; Koenker (2005) and Yu et al. (2003) for a comprehensive review. One of the attractions of quantile regression over its classical mean regression counterpart lies in its flexibility in providing a more complete investigation of the entire distribution of the relationship between a response variable and its covariates. To this end, quantile regression has gradually emerged as a comprehensive extension to standard mean regression. Suppose that we have a sample $(\boldsymbol{x}'_1, y_1), \ldots, (\boldsymbol{x}'_n, y_n)$. Then, the pth quantile regression takes the form of

$$Q_{y_i}(p|\boldsymbol{x}_i) = \boldsymbol{x}'_i\boldsymbol{\beta}_p, \quad 0 < p < 1, \tag{1}$$

where each y_i is the response variable, $\boldsymbol{x}'_i = (x_{i1}, \ldots, x_{ik})$ is a $1 \times k$ vector denoting the ith row of the $n \times k$ matrix of covariates $\boldsymbol{X}$, the unknown quantity $\boldsymbol{\beta}_p$ is a vector of k regression parameters and $Q_{y_i}(\cdot) = F_{y_i}^{-1}(\cdot)$ is the inverse of the cumulative distribution function of the response variable y_i conditional on $\boldsymbol{x}'_i$. Koenker and Bassett (1978) demonstrate that the regression coefficient vector $\boldsymbol{\beta}_p$ can be estimated consistently as the solution to the minimization of

$$\sum_{i=1}^{n} \rho_p(y_i - \boldsymbol{x}'_i\boldsymbol{\beta}_p), \tag{2}$$

[☆] The article was originally scheduled to be part of the special issue Special issue on QUANTILE REGRESSION AND SEMIPARAMETRIC METHODS, however is now being published as a regular article because the article was missed to be published along with the special issue.

* Corresponding author at: Department of Mathematics, Brunel University, Uxbridge UB8 3PH, UK. Tel.: +44 (0) 1895 265618.
E-mail address: rahim.al-hamzawi@brunel.ac.uk (R. Alhamzawi).

0167-9473/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.csda.2012.01.014

where $\rho_p(\cdot)$ is the check function defined by

$$\rho_p(t) = \frac{|t| + (2p-1)t}{2}. \tag{3}$$

Since (2) is not differentiable at the origin, there is no closed form solution for $\boldsymbol{\beta}_p$ (Koenker, 2005). However, the minimization of (2) can be achieved through implementation of the algorithm proposed by Koenker and D'Orey (1987). Alternatively, Koenker and Machado (1999) noted that the check function (2) is closely related to the asymmetric Laplace distribution and consequently the unknown parameters $\boldsymbol{\beta}_p$ can be estimated through exploiting this link.

From a Bayesian point of view, Yu and Moyeed (2001) proposed a Bayesian modeling approach of quantile regression using the asymmetric Laplace error distribution and sampling $\boldsymbol{\beta}_p$ from its posterior distribution using Markov chain Monte Carlo (MCMC) methods. The authors used improper uniform prior distributions for all the regression coefficients.

A serious challenge in Bayesian quantile regression lies in specifying a quantile dependent prior. Additionally, despite having attracted a great deal of attention in the literature, see Tsionas (2003), Scaccia and Green (2003), Schennach (2005), Dunson and Taylor (2005), Geraci and Bottai (2007), Taddy and Kottas (2007), Yu and Stander (2007), Kottas and Krnjajic (2009), Kozumi and Kobayashi (2009), Reed and Yu (2009), Lancaster and Jun (2010), Li et al. (2010), Reich et al. (2010), Yuan and Yin (2010), Yue and Rue (2010), Benoit and Poel (2010), Taddy and Kottas (2007) and Gerlach et al. (2011), it is well known that for the Bayesian inference quantile regression proposed by Yu and Moyeed (2001) and Yu and Stander (2007) a standard conjugate prior distribution is not available. While the aforementioned Bayesian inference models cover parametric, semi-parametric as well as nonparametric approaches almost all of these models set priors independent of the values of quantiles. That is, the same prior used for modeling different order of quantiles. In so doing, this approach may result in inflexibility in quantile modeling. For example, a 95% quantile regression model should have different parameter values from the median quantile, and thus the priors used for modeling the quantiles should be different (Alhamzawi and Yu, 2011; Alhamzawi et al., 2011). It is therefore more reasonable to set different priors for different quantiles.

A second serious challenge in quantile regression lies in Bayesian variable selection, due to the challenge in specifying a quantile dependent prior over model space. At present time, all Bayesian variable selection approaches in quantile regression set priors independent of the value of quantiles over model space (see, Reed et al., 2009 and Ji et al., 2011, among others). Finally, another serious challenge encountered in modeling with Bayesian quantile regression lies in computational efficiency.

In this article we address these three issues. For the first, it is quite important to elicit a prior distribution for quantile regression coefficients that is as informative as possible, and more crucially, that depends on the quantile level. In order to address this challenge a quantile dependent conjugate prior distribution is proposed. For the second, the percentage bend correlation is used to obtain suitable priors over model space and to address the third difficulty a new Gibbs sampler is proposed to facilitate the computations.

The rest of this article is organized as follows. Section 2 introduces a modification of Zellners g-prior in quantile regression as well as presenting the Bayesian MCMC estimation procedure. An outline of prior assumptions and a simple Gibbs sampler for model selection are addressed in Section 3, and in Section 4 the appropriateness of the ALD-based posterior distribution is justified. In Section 5 simulation studies are conducted to examine the performance of the proposed approaches for model selection and estimation, Section 6 provide an illustration of the proposed methods using real data examples and finally, in Section 7, we conclude the article with a brief discussion.

2. Methods

2.1. Zellner's informative g-prior

It is well known that conjugate priors play the most crucial role in Bayesian analysis, as it is desirable to have conditional distributions as the prior in terms of the same functional form and similar properties (Chen and Ibrahim, 2003). In standard mean regression, various approaches for assessing the prior distribution for regression coefficients and the variance in the nature of conjugate form have been proposed over the years. However, it is difficult to assess the prior covariances matrix for regression coefficients (Zellner, 1983; Agliari and Parisetti, 1988). For this reason, Zellner (1983, 1986) proposed a procedure for assessing a conjugate prior distribution referred to as Zellner's informative g-prior, or simply, g-prior (Agliari and Parisetti, 1988).

Zellner's informative g-prior has been widely used in the context of Bayesian analysis for the mean regression models due to the fact that analytical results are more readily available, computational efficiency and its simple interpretation (Krishna et al., 2008). For a normal regression model

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{e}, \quad \boldsymbol{e} \sim N(0, \sigma^2 \boldsymbol{I}),$$

with a vector of regression coefficients $\boldsymbol{\beta}$, Zellner's informative g-prior based on n observations and k predictors can be written as

$$p(\boldsymbol{\beta}, \sigma | \boldsymbol{\beta}_a, \sigma_a, \boldsymbol{y}, \boldsymbol{X}) \propto \sigma^{-(n-k+1)} \exp\{-(n-k)\sigma_a^2/2\sigma^2\} \sigma^{-k} \exp\left\{-(\boldsymbol{\beta} - \boldsymbol{\beta}_a)'\boldsymbol{X}'\boldsymbol{X}(\boldsymbol{\beta} - \boldsymbol{\beta}_a)/2g\sigma^2\right\}, \tag{4}$$