Journal of the Korean Statistical Society 43 (2014) 133–147

Contents lists available at ScienceDirect

Journal of the Korean Statistical Society

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

On Bayesian estimation of regression models subject to uncertainty about functional constraints

Hea-Jung Kim [a], Taeryon Choi [b,*]

[a] Department of Statistics, Dongguk University, Seoul, Republic of Korea

[b] Department of Statistics, Korea University, Republic of Korea

ARTICLE INFO

Article history:
Received 7 May 2012
Accepted 20 March 2013
Available online 18 April 2013

AMS 2000 subject classifications:
62F15
62J02
62H12

Keywords:
Functional constraints
Hierarchical priors
Posterior distribution
Predictive distribution
Rectangle screened multivariate normal distribution

ABSTRACT

In this paper, we provide a Bayesian estimation procedure for the regression models when the constraint of the regression function needs to be incorporated in modeling but such a restriction is uncertain. For this purpose, we consider a family of rectangle screened multivariate Gaussian prior distributions in order to reflect uncertainty about the functional constraint, and propose the Bayesian estimation procedure of the regression models based on two stages of a prior hierarchy of the functional constraint, referred to as *hierarchical screened Gaussian regression models* (*HSGRM*). Specifically, we explore theoretical properties of the proposed estimation procedure by deriving the posterior distribution and predictive distribution of the unknown parameters under *HSGRM* in analytic forms, and discuss specific applications to regression models with uncertain functional constraints that can be explained in the context of *HSGRM*.

1. Introduction

Functional constraints in regression models naturally arise in a wide variety of practical problems. For examples, the linear inequality constraints for the coefficients of linear regression models are common in applied econometrics and other quantitative sciences, and the shape constraints on the regression function such as monotonicity and convexity are often assumed to be known a priori or plausible in theory in the nonparametric regression model. Thus, functional constraints may be incorporated in the estimation procedure of the regression models, and these a priori functional constraints should be reflected in prior distributions from the Bayesian point of view. In this regard, several Bayesian estimation procedures have been considered in the literature for regression models subject to functional constraints. O'Hagan (1973) considered a quadratic regression with constraint, and Chen and Deely (1996) and Geweke (1996) considered linear models subject to linear inequality constraints. For Bayesian nonparametric estimation of regression function subject to shape constraints, Lavine and Mockus (1995) introduced monotone regression using Dirichlet process priors, and Bornkamp and Ickstadt (2009), Brezger and Steiner (2008), Meyer, Hackstadt, and Hoeting (2011) and Neelon and Dunson (2004) considered shape restricted regression models using splines. In addition, Chang, Chien, Hsiung, Wen, and Wu (2007) proposed shape restricted regressions based on Bernstein polynomials. Shively, Sager, and Walker (2009) proposed two Bayesian methods for estimating monotone function using a Winer process prior and regression splines and extended their approaches to shape restricted regression models that include monotonicity, convexity and other constraints in Shively, Walker, and Damien (2011). In

* Corresponding author. Tel.: +82 2 3290 2245; fax: +82 2 924 9895.
E-mail addresses: trchoi@gmail.com, trchoi@korea.ac.kr (T. Choi).

1226-3192/$ – see front matter
http://dx.doi.org/10.1016/j.jkss.2013.03.005

addition to functional constraints, Bayesian inference on order-constrained parameters have also been studied in the literature (e.g. Dunson & Neelon, 2003, Hoijtink, Klugkist, & Boelen, 2008, Klugkist & Hoijtink, 2007, Mulder, Hoijtink, & Klugkist, 2010, Oh & Shin, 2011, and references therein).

However, it is often the case that the actual observations may violate functional constraints on account of measurement errors or due to some other reasons, and even the data may provide strong evidence that the constraints are inappropriate and may appear to contradict the theory. In this respect, it is expected to take uncertainty about the constraint into account in the estimation procedure, and O'Hagan and Leonard (1976) indeed proposed a family of skew prior distributions which reflect uncertainty about parameter constraints. They constructed a family of prior distributions with skew densities for estimating the unknown normal mean parameter when the prior considerations suggests the positive constraint on the normal mean parameter but it is not completely convinced. Specifically, they specified a positively skew distribution via the two stages of a prior hierarchy for the normal mean problem instead of using the positively truncated prior distribution. The concept of a prior hierarchy suggested by O'Hagan and Leonard (1976) can be generalized into the complex models. For examples, based on the two stages of a prior hierarchy, Madi, Leonard, and Tsui (2000) considered estimating treatment effects with uncertain order constraints and Kim (2011) considered estimation of simple normal models under uncertain inequalities constraints, and Liseo and Loperfido (2003) adopted the approach of O'Hagan and Leonard (1976) to obtain a hierarchical interpretation of two skew-normal densities in the multivariate case. The family of multivariate skew-normal distributions obtained by Liseo and Loperfido (2003), known as *hierarchical skew-normal* (*HSN*) family can be viewed as a special case of multivariate skewed distributions arising from selections (Arellano-Valle & Azzalini, 2006; Arellano-Valle, Branco, & Genton, 2006; Arellano-Valle, Genton, & Loschi, 2009) or screening (Boys & Dunsmore, 1986; Kim, 2008; Kim & Kim, 2012). In particular, the hierarchical Bayesian multivariate normal structure in Kim (2011) can be adapted to the general Bayesian regression problems with the uncertain functional constraints, and we notice that a class of rectangle-screened multivariate normal distributions in Kim and Kim (2012) would be useful for Bayesian estimation of the problems.

To the best of our knowledge, however, a Bayesian approach has not previously been investigated for regression models with functional constraints subject to uncertainty. Accordingly, influenced by the seminal work of O'Hagan and Leonard (1976), we provide a novel Bayesian estimation procedure for the regression models when the constraint of the regression function needs to be incorporated in modeling but such a restriction is uncertain. For this purpose, we consider a family of screened multivariate Gaussian prior distributions in order to reflect uncertainty about the functional constraint, and propose the Bayesian estimation procedure of regression models based on the two stages of screened multivariate Gaussian prior distributions, referred to as *hierarchical screened Gaussian regression models* (*HSGRM*). The proposed method is also a unified Bayesian estimation procedure since it is able to deal with the parametric regression model as well as nonparametric regression model in the context of *HSGRM* as discussed in the paper.

The remainder of this paper is organized as follows. In Section 2, we give a basic description of the general Gaussian regression model and provide fundamental properties of screened-Gaussian distribution that will base the estimation procedure of *HSGRM*. In Section 3, we explore theoretical properties of the proposed estimation procedure by analytically deriving the posterior distribution and predictive distribution of the unknown parameters under *HSGRM*. Specific Bayesian regression models with uncertain functional constraints are illustrated in the context of *HSGRM* in Section 4. Finally, concluding remarks with a discussion are made in Section 5.

2. Problem description

2.1. General regression model

Consider the following regression model for normally distributed data,

$$\boldsymbol{y}_n = \boldsymbol{\eta}_n(\mathbf{x}) + \boldsymbol{\epsilon}_n, \quad \boldsymbol{\epsilon}_n \sim N_n(\mathbf{0}, \boldsymbol{C}), \tag{1}$$

where $\boldsymbol{y}_n = (y_1, \ldots, y_n)^\top$ is $n \times 1$ vector of responses, $y_i = \eta(x_i) + \epsilon_i$, $\boldsymbol{\eta}_n(\mathbf{x}) = (\eta(\mathbf{x}_1), \ldots, \eta(\mathbf{x}_n))^\top$ is $n \times 1$ vector of regression function values evaluated at each of $p(\geq 1)$-dimensional predictor $\mathbf{x}_i$, $i = 1, \ldots, n$, $\mathbf{x} = (\mathbf{x}_1, \ldots, \mathbf{x}_n)^\top$ is the $n \times p$ design matrix, and $\boldsymbol{\epsilon}_n = (\epsilon_1, \ldots, \epsilon_n)^\top$ is a $n \times 1$ vector of random noises that has a n-variable normal distribution with zero mean vector and covariance matrix $\boldsymbol{C}$. For simplicity, we assume that the random noises $\boldsymbol{\epsilon}_n$ are independent and identically distributed, and thus $\boldsymbol{C}$ is a diagonal matrix, $\boldsymbol{C} = \sigma^2\boldsymbol{I}$. In the basic model structure of (1), $\boldsymbol{\eta}_n(\mathbf{x})$ may have various functional forms of the predictors $\mathbf{x}$. For example, the simple linear regression model is concerned with $\boldsymbol{\eta}_n(\mathbf{x}) = (\beta_0 + \beta_1 x_1, \ldots, \beta_0 + \beta_1 x_n)^\top$, where β_0 and β_1 are unknown parameters to be estimated. In the case of a nonparametric regression model, the parametric form of the regression function is not assumed and $\boldsymbol{\eta}_n(\mathbf{x}) = (\eta(\mathbf{x}_1), \ldots, \eta(\mathbf{x}_n))^\top$, but $\boldsymbol{\eta}_n(\mathbf{x})$ is assumed to have specific types of functional structure. For instance, $\boldsymbol{\eta}_n(\mathbf{x})$ can be represented with a Fourier series (Lenk, 1999), splines (Fahrmeir & Kneib, 2011), kernels (Chakraborty, Ghosh, & Mallick, 2012), Gaussian processes (Rasmussen & Williams, 2006) and others.

Bayesian analysis of the regression model in (1) begins with the specification of prior distributions for unknown parameters $\boldsymbol{\eta}_n(\mathbf{x})$ and the noise variance σ^2. Specifically, we assign a multivariate normal prior distribution for $\boldsymbol{\eta}_n(\mathbf{x})$ and an inverse gamma prior for σ^2, which are commonly used in normal regression models as conjugate priors,

$$\boldsymbol{\eta}_n \sim N_n(\boldsymbol{\eta}_0, \Sigma_0) \quad \text{and} \quad \sigma^2 \sim \mathrm{IG}(c, d), \tag{2}$$