Journal of Advances in Applied Mathematics, Vol. 2, No. 4, October 2017
https://dx.doi.org/10.22606/jaam.2017.24001

Study on Bayes Semiparametric Regression

Abdulhussein Saber AL-Mouel and Ameera Jaber Mohaisen

Mathematics Department College of Education for Pure Sciences, AL-Basrah University, Iraq
Email: ameera.jaber@yahoo.com

Abstract. In this paper, Bayesian approach based on Markov chain Monte Carlo (MCMC) to (fully) Semiparametric regression problems is described as a mixed model using a convenient connection between penalized splines and mixed models. We investigate the inferences on the model coefficients under some conditions on the prior, as well as studying some properties of the posterior distribution and identifying the analytic form of the Bayes factors.

Keywords: Semiparametric regression, penalized spline, mixed model, bayes approach, prior distribution, posterior distribution, bayes factor.

1 Introduction

Semiparametric regression models have been investigated by many researchers. Lenk (1999) presented the Bayesian inference of a semiparametric regression model using a Fourier representation[6]. Natio (2002) studied semiparametric regression with multiplicative adjustment[10]. Also Brezger et al. (2002) investigated and analyzed Bayesian semiparametric models[1]. Ruppert et al. (2003) introduced semiparametric regression models based on penalized regression splines and mixed models[12]. Tsiatis and MA (2004) studied locally efficient semiparametric estimators for functional measurements error models by defining after projecting the score vector with respect to the parameter β on to the nuisance tangent space for the nonparametric conditional distribution of X given Z, where Z is the predictor variable measured precisely[14]. In (2007), Yuan and DE Gooijer presented semiparametric regression with kernel error model[16]. Also in (2007), Jensen and Maheu studied bayesian semiparametric stochastic volatility modeling[3]. Choi, Lee and Roy (2008) investigated the large sample property of the Bayes factor for testing the parametric null model against the Semiparametric alternative model[2].

Wand (2009) presented semiparametric regression and graphical models[15]. Tarmaratram (2011) proposed a robust estimation method in semiparametric regression models for penalized regression splines that can be used in the presence of outliers in the response variable, and studied a robust version of the model selection criterion AIC, Akaiki's information criterion for regression models where S- and MM- estimators are used for estimation[13]. Pelenis (2012) studied Bayesian semiparametric regression that considered a Bayesian estimation of restricted conditional moment model with linear regression as a particular example[11]. Mohaisen and Abdulhussain investigated Bayesian semiparametric regression based on penalized spline[7-9].

This paper came to shed light on the semiparametric regression model which has two parts, the parametric (first part) is assumed to be linear function of p-dimensional covariates and the nonparametric (second part) is assumed to be a smooth penalized spline, as well as the error term which has normal distribution with mean zero and variance σ_ε^2. We can represent semiparametric regression model as a mixed model by using a convenient connection between penalized splines and mixed model.

In this paper, Bayesian approach based on Markov chain Monte Carlo (MCMC) to (fully) Semiparametric regression problems is described as a mixed model using a convenient connection between penalized splines and mixed models. We investigate the inferences on the model coefficients under some conditions on the prior, as well as studying some properties of the posterior distribution and identifying the analytic form of the Bayes factors.

2 Description of the Problem and the Prior Distribution

Consider the model:

$$y_i = \sum_{j=0}^{p} \beta_j x_{ji} + g(x_{p+1,i}) + \varepsilon_i, i = 1, 2, ..., n \tag{1}$$

where $\sum_{j=1}^{p} \beta_j x_{ji}$ is the parametric part which assumed to be linear function of p-dimensional covariates, $g(x_{p+1,i})$ the nonparametric part and the unobserved errors $\varepsilon_1, \varepsilon_2, ..., \varepsilon_n$ are known to be i.i.d. normal with mean zero and covariance $\sigma_\varepsilon^2 I_n$ with σ_ε^2 unknown. By using penalized spline of degree q for the model (1) we get:

$$y_i = \sum_{j=0}^{p} \beta_j x_{ji} + \sum_{j=1}^{q} \beta_{p+j} x_{p+1,i}^{j} + \sum_{k=1}^{K} u_k (x_{p+1,i} - k_k)_+^q + \varepsilon_i, i = 1, 2, ..., n \tag{2}$$

where $k_1, ..., k_K$ are inner knots $a < k_1 < ... < k_K < b$. By using a convenient connection between penalized splines and mixed models, the model (2) is rewritten as follows:

$$Y = X\beta + Zu + \varepsilon \tag{3}$$

where Y has length n, X is a $n \times (p + q + 1)$ design matrix of pure polynomial component of the spline, Z is a $n \times K$ design matrix of spline truncated functions, β is a $(p + q + 1)$-vector of parameters of pure polynomial component of the spline, u is a K- vector with spline truncated functions, and the vector of error term ε has length n, $\varepsilon \overset{iid}{\sim} N(0, \sigma_\varepsilon^2 I)$.

Assume that u and ε are independent and the prior distribution on u, $\pi_0(u)$ is $N(0, \sigma_u^2 I)$, the prior distribution on the parameters vector β, $\pi_0(\beta)$ is $N(0, \sigma_\beta^2 I)$, and we will assume that the prior distribution on σ_ε^2, $\pi_0(\sigma_\varepsilon^2)$ is inverse gamma $IG(\alpha_\varepsilon, \beta_\varepsilon)$, also we assume that $\sigma_u^2 \sim IG(\alpha_u, \beta_u)$, where the hyperparameters $\alpha_\varepsilon, \beta_\varepsilon, \alpha_u, \beta_u$ that determine the priors and must be chosen by the statistician.

3 Posterior Distribution

From the model (3) we have

$$Y \mid \theta, \sigma_\varepsilon^2, \sigma_u^2 \sim N(C\theta, \sigma_\varepsilon^2 I_n) \tag{4}$$

where, $C = [X \quad Z]$ and $\theta = [\beta \quad u]^T$. Then, the likelihood function $L(Y \mid \theta, \sigma_\varepsilon^2, \sigma_u^2)$ is

$$L(Y \mid \theta, \sigma_\varepsilon^2, \sigma_u^2) \propto \left|\sigma_\varepsilon^2\right|^{-1/2} \exp\left\{-\tfrac{1}{2}(Y - C\theta)^T (\sigma_\varepsilon^2 I_n)^{-1} (Y - C\theta)\right\} \tag{5}$$

Then, the posterior distributions of the vector of coefficients θ and the error variance σ_ε^2 and σ_u^2 are

$$\begin{aligned} &\pi_1(\theta \mid Y, \sigma_\varepsilon^2, \sigma_u^2) \propto L(Y \mid \theta, \sigma_\varepsilon^2, \sigma_u^2)\pi_0(\theta) \\ &\Rightarrow \pi_1(\theta \mid Y, \sigma_\varepsilon^2, \sigma_u^2) \propto \exp\left\{-\tfrac{1}{2}(Y - C\theta)^T (\sigma_\varepsilon^2 I_n)^{-1} (Y - C\theta)\right\} \pi_0(\theta) \end{aligned} \tag{6}$$

and

$$\pi_1(\sigma_\varepsilon^2 \mid Y, \theta, \sigma_u^2) \propto \left|\sigma_\varepsilon^2\right|^{-1/2} \exp\left\{-\tfrac{1}{2}(Y - C\theta)^T (\sigma_\varepsilon^2 I_n)^{-1} (Y - C\theta)\right\} \pi_0(\sigma_\varepsilon^2) \tag{7}$$

$$\pi_1(\sigma_u^2 \mid Y, \theta, \sigma_\varepsilon^2) \propto \left|\sigma_\varepsilon^2\right|^{-1/2} \exp\left\{-\tfrac{1}{2}(Y - C\theta)^T (\sigma_\varepsilon^2 I_n)^{-1} (Y - C\theta)\right\} \pi_0(\sigma_u^2) \tag{8}$$

From (6) we can see

$$\theta \mid Y, \sigma_\varepsilon^2, \sigma_u^2 \sim N\left(\mu_{\theta|Y,\sigma_\varepsilon^2,\sigma_u^2}, \Sigma_{\theta|Y,\sigma_\varepsilon^2,\sigma_u^2}\right) \tag{9}$$

where

$$\mu_{\theta|Y,\sigma_\varepsilon^2,\sigma_u^2} = \left\{\Sigma \Lambda\right\}\left\{\sigma_\varepsilon^2 I_n + [\Sigma \Lambda C^T C]\right\}^{-1} C^T \tag{10}$$

$$\Sigma_{\theta|Y,\sigma_\varepsilon^2,\sigma_u^2} = \Sigma\Lambda - \Sigma^2\Lambda C^T\left\{\sigma_\varepsilon^2 I_n + [\Sigma\Lambda C^T C]\right\}^{-1}\left\{C\Lambda\right\} \quad (11)$$

and

$$\Sigma = \begin{bmatrix} \sigma_\beta^2 & 0 \\ 0 & \sigma_u^2 \end{bmatrix}, \Lambda = \begin{bmatrix} I_{p+q+1} & 0 \\ 0 & I_{n-(p+q+1)} \end{bmatrix}$$

then

$$\Sigma\Lambda = \begin{bmatrix} \sigma_\beta^2 I_{p+q+1} & 0 \\ 0 & \sigma_u^2 I_{n-(p+q+1)} \end{bmatrix}$$

Now, by spectral decomposition we obtain $C^T C = PDP^T$ [4], where $D = diag(d_1,...,d_n)$ is the matrix of eigenvalues and P is the orthogonal matrix of eigenvectors. Thus

$$\sigma_\varepsilon^2 I_n + \Sigma\Lambda C^T C = \sigma_\varepsilon^2 P\left\{I_n + \begin{pmatrix} \delta I_{p+q+1} & 0 \\ 0 & \gamma I_{n-(p+q+1)} \end{pmatrix} D\right\}P^T \quad (12)$$

where, $\delta = \dfrac{\sigma_\beta^2}{\sigma_\varepsilon^2}$ and $\gamma = \dfrac{\sigma_u^2}{\sigma_\varepsilon^2}$. Then, the conditional density of Y given σ_ε^2, δ and γ can be written as:

$$m(Y \mid \sigma_\varepsilon^2,\delta,\gamma) = \frac{1}{(2\pi\sigma_\varepsilon^2)^{n/2}\det\left[I_n + \begin{pmatrix} \delta I_{p+q+1} & 0 \\ 0 & \gamma I_{n-(p+q+1)} \end{pmatrix} D\right]^{1/2}}$$

$$\exp\left\{-\frac{1}{2\sigma_\varepsilon^2}Y^T P\left[I_n + \begin{pmatrix} \delta I_{p+q+1} & 0 \\ 0 & \gamma I_{n-(p+q+1)} \end{pmatrix} D\right]^{-1} P^{TY}\right\}$$

$$= \frac{1}{(2\pi\sigma_\varepsilon^2)^{n/2}}\frac{1}{\left[\prod_{i=1}^{p+q+1}[1+\delta d_i]\right]^{1/2}\left[\prod_{i=p+q+2}^{n}[1+\gamma d_i]\right]^{1/2}}$$

$$\exp\left\{-\frac{1}{2\sigma_\varepsilon^2}\left(\sum_{i=1}^{p+q+1}\frac{s_i^2}{1+\delta d_i} + \sum_{i=p+q+2}^{n}\frac{s_i^2}{1+\gamma d_i}\right)\right\}$$

where $s = (s_1,...,s_n)^T = P^T Y$.

Theorem 1. The joint posterior density of δ, γ given Y is

$$\pi_1(\delta,\gamma \mid Y) \propto \frac{\gamma^{(b/2)-1}\delta^{\alpha_\varepsilon - 1}e^{\frac{\sigma_\beta^2\delta}{\beta_\varepsilon}}}{(a+b\gamma)^{-(a+b)/2}}\left(\prod_{i=1}^{p+q+1}(1+\delta d_i)\right)^{-1/2}\left(\prod_{i=p+q+2}^{n}(1+\gamma d_i)\right)^{-1/2}$$
$$\left(2\beta_\varepsilon + \sum_{i=1}^{p+q+1}\frac{s_i^2}{1+\delta d_i} + \sum_{i=p+q+2}^{n}\frac{s_i^2}{1+\gamma d_i}\right) \quad (13)$$

Proof. Since $\delta \sim Gamma(\alpha_\varepsilon,\beta_\varepsilon)$, $\gamma \sim F(b,a)$ [5]

$$\therefore \pi_1(\delta,\gamma \mid Y) = \int m(Y \mid \sigma_\varepsilon^2,\delta,\gamma) f(\delta,\alpha_\varepsilon,\beta_\varepsilon) f(\gamma,b,a) f(\sigma_\varepsilon^2,\alpha_\varepsilon,\beta_\varepsilon) d\sigma_\varepsilon^2$$

$$= \int \frac{1}{(2\pi\sigma_\varepsilon^2)^{n/2}}\left(\prod_{i=1}^{p+q+1}(1+\delta\, d_i)\right)^{-1/2}\left(\prod_{i=p+q+2}^{n}(1+\gamma\, d_i)\right)^{-1/2}\frac{\sigma_\beta^2}{\beta_\varepsilon\Gamma(\alpha_\varepsilon)}\left(\frac{\sigma_\beta^2\,\delta}{\beta_\varepsilon}\right)^{\alpha_\varepsilon - 1}$$

$$\exp\left\{-\frac{1}{2\sigma_\varepsilon^2}\left(\sum_{i=1}^{p+q+1}\frac{s_i^2}{1+\delta d_i}+\sum_{i=p+q+2}^{n}\frac{s_i^2}{1+\gamma d_i}\right)\right\}\exp\left\{\frac{\sigma_\beta^2\delta}{\beta_\varepsilon}\frac{b^{b/2}a^{a/2}}{\beta(b,a)}\frac{\gamma^{(b/2)-1}}{(a+b\gamma)^{-(a+b)/2}}\right\}$$

$$\frac{\beta_\varepsilon^{\alpha_\varepsilon}}{\Gamma(\alpha_\varepsilon)}(\sigma_\varepsilon^2)^{-(\alpha_\varepsilon+1)}\exp\left\{-\frac{\beta_\varepsilon}{\sigma_\varepsilon^2}\right\}d\sigma_\varepsilon^2$$

$$=(2\pi)^{-n/2}\frac{(\sigma_\beta^2\delta)^{\alpha_\varepsilon-1}e^{\frac{\sigma_\beta^2\delta}{\beta_\varepsilon}}}{(\Gamma(\alpha_\varepsilon))^2}\frac{b^{b/2}a^{a/2}}{\beta(b,a)}\frac{\gamma^{(b/2)-1}}{(a+b\gamma)^{-(a+b)/2}}$$

$$\int\frac{1}{(2\pi\sigma_\varepsilon^2)^{n/2}}\left(\prod_{i=1}^{p+q+1}(1+\delta d_i)\right)^{-1/2}\left(\prod_{i=p+q+2}^{n}(1+\gamma d_i)\right)^{-1/2}$$

$$\exp\left\{-\frac{1}{2\sigma_\varepsilon^2}\left(\beta_\varepsilon+\sum_{i=1}^{p+q+1}\frac{s_i^2}{1+\delta d_i}+\sum_{i=p+q+2}^{n}\frac{s_i^2}{1+\gamma d_i}\right)\right\}(\sigma_\varepsilon^2)^{-(n+2\alpha_\varepsilon+2)/2}d\sigma_\varepsilon^2$$

$$=(2\pi)^{-n/2}\frac{(\sigma_\beta^2\delta)^{\alpha_\varepsilon-1}e^{\frac{\sigma_\beta^2\delta}{\beta_\varepsilon}}}{(\Gamma(\alpha_\varepsilon))^2}\frac{b^{b/2}a^{a/2}}{\beta(b,a)}\frac{\gamma^{(b/2)-1}}{(a+b\gamma)^{-(a+b)/2}}(2)^{(n+2\alpha_\varepsilon+2)/2}$$

$$\int\frac{1}{(2\pi\sigma_\varepsilon^2)^{n/2}}\left(\prod_{i=1}^{p+q+1}(1+\delta d_i)\right)^{-1/2}\left(\prod_{i=p+q+2}^{n}(1+\gamma d_i)\right)^{-1/2}$$

$$\exp\left\{-\frac{1}{2\sigma_\varepsilon^2}\left(2\beta_\varepsilon+\sum_{i=1}^{p+q+1}\frac{s_i^2}{1+\delta d_i}+\sum_{i=p+q+2}^{n}\frac{s_i^2}{1+\gamma d_i}\right)\right\}$$

$$\left(\frac{2\beta_\varepsilon+\sum_{i=1}^{p+q+1}\frac{s_i^2}{1+\delta d_i}+\sum_{i=p+q+2}^{n}\frac{s_i^2}{1+\gamma d_i}}{2\sigma_\varepsilon^2}\right)^{(n+2\alpha_\varepsilon+2)/2}$$

$$\left(2\beta_\varepsilon+\sum_{i=1}^{p+q+1}\frac{s_i^2}{1+\delta d_i}+\sum_{i=p+q+2}^{n}\frac{s_i^2}{1+\gamma d_i}\right)^{-(n+2\alpha_\varepsilon+2)/2}d\sigma_\varepsilon^2$$

$$\propto\frac{\gamma^{(b/2)-1}\delta^{\alpha_\varepsilon-1}e^{\frac{\sigma_\beta^2\delta}{\beta_\varepsilon}}}{(a+b\gamma)^{-(a+b)/2}}\int\frac{1}{(2\pi\sigma_\varepsilon^2)^{n/2}}\left(\prod_{i=1}^{p+q+1}(1+\delta d_i)\right)^{-1/2}\left(\prod_{i=p+q+2}^{n}(1+\gamma d_i)\right)^{-1/2}$$

$$\exp\left\{-\frac{1}{2\sigma_\varepsilon^2}\left(2\beta_\varepsilon+\sum_{i=1}^{p+q+1}\frac{s_i^2}{1+\delta d_i}+\sum_{i=p+q+2}^{n}\frac{s_i^2}{1+\gamma d_i}\right)\right\}$$

$$\left(\frac{2\beta_\varepsilon+\sum_{i=1}^{p+q+1}\frac{s_i^2}{1+\delta d_i}+\sum_{i=p+q+2}^{n}\frac{s_i^2}{1+\gamma d_i}}{2\sigma_\varepsilon^2}\right)^{[(n+2\alpha_\varepsilon+2)/2]-1}$$

$$\left(2\beta_\varepsilon+\sum_{i=1}^{p+q+1}\frac{s_i^2}{1+\delta d_i}+\sum_{i=p+q+2}^{n}\frac{s_i^2}{1+\gamma d_i}\right)^{-(n+2\alpha_\varepsilon+2)/2}d\sigma_\varepsilon^2$$

$$\propto\frac{\gamma^{(b/2)-1}\delta^{\alpha_\varepsilon-1}e^{\frac{\sigma_\beta^2\delta}{\beta_\varepsilon}}}{(a+b\gamma)^{-(a+b)/2}}\Gamma((n+2\alpha_\varepsilon+4)/2)\left(\prod_{i=1}^{p+q+1}(1+\delta d_i)\right)^{-1/2}\left(\prod_{i=p+q+2}^{n}(1+\gamma d_i)\right)^{-1/2}$$

$$\left(2\beta_\varepsilon + \sum_{i=1}^{p+q+1}\frac{s_i^2}{1+\delta d_i} + \sum_{i=p+q+2}^{n}\frac{s_i^2}{1+\gamma d_i}\right)^{-(n+2\alpha_\varepsilon+2)/2}$$

$$\therefore \pi_1(\delta,\gamma \mid Y) \propto \frac{\gamma^{(b/2)-1}\delta^{\alpha_\varepsilon-1}e^{\frac{\sigma_\beta^2\delta}{\beta_\varepsilon}}}{(a+b\gamma)^{-(a+b)/2}}\left(\prod_{i=1}^{p+q+1}(1+\delta d_i)\right)^{-1/2}\left(\prod_{i=p+q+2}^{n}(1+\gamma d_i)\right)^{-1/2}$$

$$\left(2\beta_\varepsilon + \sum_{i=1}^{p+q+1}\frac{s_i^2}{1+\delta d_i} + \sum_{i=p+q+2}^{n}\frac{s_i^2}{1+\gamma d_i}\right)^{-(n+2\alpha_\varepsilon+2)/2} \square$$

Theorem 2. The posterior mean and covariance matrix of θ are

$$E(\theta \mid Y) = \Lambda P \quad E\left\{\left[I_n + \begin{pmatrix}\delta I_{p+q+1} & 0\\ 0 & \gamma I_{n-(p+q+1)}\end{pmatrix}D\right]^{-1} \mid Y\right\}C^T s \tag{14}$$

and

$$Var(\theta \mid Y) = \frac{1}{n+2\alpha_\varepsilon+2}E\left[\frac{\left(2\beta_\varepsilon + \left(\sum_{i=1}^{p+q+1}\frac{s_i^2}{1+\delta d_i} + \sum_{i=p+q+2}^{n}\frac{s_i^2}{1+\gamma d_i}\right)\right)}{Y}\right]\Lambda -$$

$$\frac{1}{n+2\alpha_\varepsilon+2}\Lambda C^T P \quad E\left(2\beta_\varepsilon + \left(\sum_{i=1}^{p+q+1}\frac{s_i^2}{1+\delta d_i} + \sum_{i=p+q+2}^{n}\frac{s_i^2}{1+\gamma d_i}\right)\right) \tag{15}$$

$$\left[\left[I_n + \begin{pmatrix}\delta I_{p+q+1} & 0\\ 0 & \gamma I_{n-(p+q+1)}\end{pmatrix}D\right]^{-1} \mid Y\right]P'C\Lambda +$$

$$E\left\{\Lambda C^T P\left\{\left[I_n + \begin{pmatrix}\delta I_{p+q+1} & 0\\ 0 & \gamma I_{n-(p+q+1)}\end{pmatrix}D\right]^{-1}\right\}\mathrm{s}\right\}$$

Proof.

$$E(\theta \mid Y) = \mu_{\theta|Y} = \left\{\sum\Lambda\right\}\left\{\sigma_\varepsilon^2 I_n + \left[\sum\Lambda C^T C\right]\right\}^{-1} C^T Y$$

$$= \sum\Lambda\left\{\sigma_\varepsilon^2 P\left[I_n + \begin{pmatrix}\delta I_{p+q+1} & 0\\ 0 & \gamma I_{n-(p+q+1)}\end{pmatrix}D\right]P^T\right\}^{-1} C^T Y$$

$$= \frac{\sum}{\sigma_\varepsilon^2}\Lambda(P^T)^{-1}\left\{\left[I_n + \begin{pmatrix}\delta I_{p+q+1} & 0\\ 0 & \gamma I_{n-(p+q+1)}\end{pmatrix}D\right]\right\}^{-1} P^{-1}C^T Y$$

$\because P$ is the orthogonal matrix of eigenvectors, then $P^{-1} = P^T$ and $(P^T)^{-1} = P$.

Therefore

$$E(\theta \mid Y) = \Lambda P\begin{pmatrix}\delta I_{p+q+1} & 0\\ 0 & \gamma I_{n-(p+q+1)}\end{pmatrix}\left\{\left[I_n + \begin{pmatrix}\delta I_{p+q+1} & 0\\ 0 & \gamma I_{n-(p+q+1)}\end{pmatrix}D\right]\right\}^{-1} P^{-1}C^T Y$$

$$= \Lambda P \quad E\left\{\left[I_n + \begin{pmatrix} \delta I_{p+q+1} & 0 \\ 0 & \gamma I_{n-(p+q+1)} \end{pmatrix} D\right]^{-1} \mid Y\right\} C^T s$$

where the expectation $E\left\{\left[I_n + \begin{pmatrix} \delta I_{p+q+1} & 0 \\ 0 & \gamma I_{n-(p+q+1)} \end{pmatrix} D\right]^{-1} \mid Y\right\}$ taken with respect to $\pi_1(\delta, \gamma \mid Y)$. By following the same way we can prove the variance of $(\theta \mid Y)$ □

4 Model Checking and Bayes Factors

We would like to choose between a Bayesian penalized spline semiparametric regression model as a mixed model and a Bayesian penalized spline semiparametric regression model with known coefficients by using Bayes factors for two hypotheses

$$H_0 : y_i = \sum_{j=0}^{p} \beta_j^0 x_{ji} + \sum_{j=1}^{q} \beta_{j+1}^0 x_{p+1,i}^j + \sum_{k=1}^{K} u_k^0 (x_{p+1,i} - k_k)_+^q + \varepsilon$$

versus

$$H_1 : y_i = \sum_{j=0}^{p} \beta_j x_{ji} + \sum_{j=1}^{q} \beta_{j+1} x_{p+1,i}^j + \sum_{k=1}^{K} u_k (x_{p+1,i} - k_k)_+^q + \varepsilon$$

or

$$\left.\begin{matrix} H_0 : Y = X\beta^0 + Zu^0 + \varepsilon \\ versus \\ H_1 : Y = X\beta + Zu + \varepsilon \end{matrix}\right\} \tag{16}$$

where β^0 and u^0 are known. We compute the Bayes factor, B_{01}, of H_0 relative to H_1 for testing problem (16) as follows

$$B_{01}(Y) = \frac{m(Y \mid H_0)}{m(Y \mid H_1)} \tag{17}$$

where $m(Y \mid H_i)$ is the marginal density of Y under model $H_i, i = 0,1$. We have:

$$m(Y \mid H_0) = \int f\left(Y \mid \beta^0, u^0, \sigma_\varepsilon^2\right) \pi_0(\sigma_\varepsilon^2) d\sigma_\varepsilon^2$$

$$= (2\pi)^{-n/2} \frac{\beta_\varepsilon^{\alpha_\varepsilon}}{\Gamma(\alpha_\varepsilon)} \int (\sigma_\varepsilon^2)^{-n/2} \exp\left(\frac{\beta_\varepsilon}{\sigma_\varepsilon^2}\right) (\sigma_\varepsilon^2)^{-(\alpha_\varepsilon+1)} \exp\left(-\frac{(Y - X\beta^0 - Zu^0)^2}{2\sigma_\varepsilon^2}\right) d\sigma_\varepsilon^2$$

$$= (2\pi)^{-n/2} \frac{\beta_\varepsilon^{\alpha_\varepsilon}}{\Gamma(\alpha_\varepsilon)} \int (\sigma_\varepsilon^2)^{-\left(\frac{n}{2}+\alpha_\varepsilon+1\right)} \exp\left(\frac{-\beta_\varepsilon + \frac{1}{2}(Y - X\beta^0 - Zu^0)^2}{2\sigma_\varepsilon^2}\right) d\sigma_\varepsilon^2$$

$$= (2\pi)^{-n/2} \frac{\beta_\varepsilon^{\alpha_\varepsilon}}{\Gamma(\alpha_\varepsilon)} \int (\sigma_\varepsilon^2)^{-\left(\frac{n}{2}+\alpha_\varepsilon+1\right)} \left(\beta_\varepsilon + \tfrac{1}{2}(Y - X\beta^0 - Zu^0)^2\right)^{\left(\frac{n}{2}+\alpha_\varepsilon+1\right)}$$

$$\left(\beta_\varepsilon + \tfrac{1}{2}(Y - X\beta^0 - Zu^0)^2\right)^{-\left(\frac{n}{2}+\alpha_\varepsilon+1\right)} \exp\left(\frac{-\beta_\varepsilon + \frac{1}{2}(Y - X\beta^0 - Zu^0)^2}{2\sigma_\varepsilon^2}\right) d\sigma_\varepsilon^2$$

$$= (2\pi)^{-n/2} \frac{\beta_\varepsilon^{\alpha_\varepsilon}}{\Gamma(\alpha_\varepsilon)} \int \frac{\left(\beta_\varepsilon + \frac{1}{2}(Y - X\beta^0 - Zu^0)^2\right)^{\left(\frac{n}{2}+\alpha_\varepsilon+1\right)}}{(\sigma_\varepsilon^2)^{\left(\frac{n}{2}+\alpha_\varepsilon+1\right)}}$$

$$\exp\left(\frac{-\beta_\varepsilon + \frac{1}{2}(Y - X\beta^0 - Zu^0)^2}{2\sigma_\varepsilon^2}\right) \left(\beta_\varepsilon + \tfrac{1}{2}(Y - X\beta^0 - Zu^0)^2\right)^{-\left(\frac{n}{2}+\alpha_\varepsilon+1\right)} d\sigma_\varepsilon^2$$

$$= (2\pi)^{-n/2} \frac{\beta_\varepsilon^{\alpha_\varepsilon}}{\Gamma(\alpha_\varepsilon)} \int \left(\frac{\beta_\varepsilon + \frac{1}{2}(Y - X\beta^0 - Zu^0)^2}{\sigma_\varepsilon^2} \right)^{\left(\frac{n}{2}+\alpha_\varepsilon+2\right)-1}$$

$$\exp\left(\frac{-\beta_\varepsilon + \frac{1}{2}(Y - X\beta^0 - Zu^0)^2}{2\sigma_\varepsilon^2} \right) \left(\beta_\varepsilon + \tfrac{1}{2}(Y - X\beta^0 - Zu^0)^2 \right)^{-\left(\frac{n}{2}+\alpha_\varepsilon+1\right)} d\sigma_\varepsilon^2 \quad (18)$$

$$= (2\pi)^{-n/2} \frac{\beta_\varepsilon^{\alpha_\varepsilon}}{\Gamma(\alpha_\varepsilon)} \Gamma(\tfrac{n}{2} + \alpha_\varepsilon + 1) \left(\beta_\varepsilon + \tfrac{1}{2}(Y - X\beta^0 - Zu^0)^2 \right)^{-\left(\frac{n}{2}+\alpha_\varepsilon+1\right)}$$

and

$$m(Y \mid H_1, \sigma_\varepsilon^2, \delta, \gamma) = (2\pi\sigma_\varepsilon^2)^{-n/2} \left(\prod_{i=1}^{p+q+1} (1 + \delta d_i) \right)^{-1/2} \left(\prod_{i=p+q+2}^{n} (1 + \gamma d_i) \right)^{-1/2}$$
$$\exp\left\{ -\frac{1}{2\sigma_\varepsilon^2} \left(\sum_{i=1}^{p+q+1} \frac{s_i^2}{1 + \delta d_i} + \sum_{i=p+q+2}^{n} \frac{s_i^2}{1 + \gamma d_i} \right) \right\} \quad (19)$$

Therefore,

$$m(Y \mid H_1) = \int m(Y \mid H_1, \sigma_\varepsilon^2, \delta, \gamma) \pi_0(\sigma_\varepsilon^2, \delta, \gamma) d\sigma_\varepsilon^2 d\delta d\gamma$$

$$= \int \frac{\beta_\varepsilon^{\alpha_\varepsilon}}{\Gamma(\alpha_\varepsilon)} (\sigma_\varepsilon^2)^{-(\alpha_\varepsilon+1)} \exp\left(\frac{\beta_\varepsilon}{\sigma_\varepsilon^2} \right) (2\pi\sigma_\varepsilon^2)^{-n/2} \left(\prod_{i=1}^{p+q+1} (1 + \delta d_i) \right)^{-1/2} \left(\prod_{i=p+q+2}^{n} (1 + \gamma d_i) \right)^{-1/2}$$

$$\exp\left\{ -\frac{1}{2\sigma_\varepsilon^2} \left(\sum_{i=1}^{p+q+1} \frac{s_i^2}{1 + \delta d_i} + \sum_{i=p+q+2}^{n} \frac{s_i^2}{1 + \gamma d_i} \right) \right\} \pi_0(\delta, \gamma) d\sigma_\varepsilon^2 d\delta d\gamma$$

$$= \frac{\beta_\varepsilon^{\alpha_\varepsilon}}{\Gamma(\alpha_\varepsilon)} (2\pi)^{-n/2} \int \left(\prod_{i=1}^{p+q+1} (1 + \delta d_i) \right)^{-1/2} \left(\prod_{i=p+q+2}^{n} (1 + \gamma d_i) \right)^{-1/2} \pi_0(\delta, \gamma)$$

$$\int \left\{ \exp\left\{ -\frac{1}{2\sigma_\varepsilon^2} \left(\beta_\varepsilon + \sum_{i=1}^{p+q+1} \frac{s_i^2}{1 + \delta d_i} + \sum_{i=p+q+2}^{n} \frac{s_i^2}{1 + \gamma d_i} \right) \right\} d\sigma_\varepsilon^2 \right\} d\delta d\gamma$$

$$\therefore m(Y \mid H_1) = \frac{\beta_\varepsilon^{\alpha_\varepsilon}}{\Gamma(\alpha_\varepsilon)} (2\pi)^{-n/2} \Gamma\left(\frac{n}{2} + \alpha_\varepsilon \right)$$
$$\int \left(\prod_{i=1}^{p+q+1} (1 + \delta d_i) \right)^{-1/2} \left(\prod_{i=p+q+2}^{n} (1 + \gamma d_i) \right)^{-1/2} \pi_0(\delta, \gamma)$$
$$\left(\beta_\varepsilon + \frac{1}{2} \left(\sum_{i=1}^{p+q+1} \frac{s_i^2}{1 + \delta d_i} + \sum_{i=p+q+2}^{n} \frac{s_i^2}{1 + \gamma d_i} \right) \right) d\delta d\gamma$$

5 Simulation Results

In this section, we illustrate the effectiveness of our methodology, we generated observations from the model (1) with the following regression functions:

$$\text{(i)} \quad y_1 = 2 + 3x_1 + \exp\{(x_2 + 0.4)^2\},$$

$$\text{(ii)} \; y_2 = 3x_1 + \frac{\sin\{12(x_2 + 0.2)\}}{(x_2 + 0.2)} - x_2^3.$$

The observations for x are generated from uniform distribution on the interval [0,1]. The sample sizes taken are $n = 25, \; 50, \; 100, \; 150, \; 200$.

The goodness of fit of the estimated models quantified by computing the criterions average mean squared error $(AMSE)$ and average mean absolute error $(AMAE)$ are defined as:

$$AMSE = \frac{1}{N}\sum_{i=1}^{N} MSE(x_i),$$

$$AMAE = \frac{1}{N}\sum_{i=1}^{N} MAE(x_i),$$

where MSE and MAE are mean squared error and mean absolute error criterions respectively.

Table (1) presents summary values of the $(AMSE)$ and $(AMAE)$ for the estimation method. From this table we can see that the values of $(AMSE)$ and $(AMAE)$ when $(n = 200)$ are smaller than their values for the first test function, which were (0.0005306171) and (0.000164242) respectively. While the values of $(AMSE)$ and $(AMAE)$ are smaller when $(n = 200)$ for the second test function were (0.0001630011) and (0.000343007) respectively. Figures (1) and (3) below show the number of iterations of Gibbs sampler used in this paper, which was (10000) iterations for the first and second test functions respectively when $(n = 200)$. While figures (2) and (4) show the density estimates based on (10000) iterations of σ_ε^2 and σ_u^2 for the first and second test functions respectively when $(n = 200)$.

Table 1. Results of the $(AMSE)$ and $(AMAE)$ criterions for Bayesian semiparametric regression model.

Test functions	Sample size	$AMSE$	$AMAE$
y_1	25	0.0035217013	0.001642312
	50	0.0036214631	0.001656452
	100	0.0026621641	0.000731483
	150	0.0006316071	0.000185535
	200	0.0005306171	0.000164242
y_2	25	0.0030001601	0.004203166
	50	0.0025130211	0.002063413
	100	0.0002110123	0.001022111
	150	0.0002001465	0.000406561
	200	0.0001630011	0.000343007

The model checking approach based on Bayes factors has been tested on simulated examples. These Bayes factors are given in table (2). From this table, it can be seen that the model corresponding to the first test function obtains the largest Bayes factor when $(n = 25)$ followed by that the second test function when $(n = 25)$, and the Bayes factor favors H_1 with strong evidence with all samples sizes for two test functions.

Table 2. Values of Bayes factors

Test functions	**Sample size**	$B_{01}(y)$	Evidence
y_1	25	3.116634×10^{-4}	Strongly favors H_1
	50	3.028341×10^{-7}	Strongly favors H_1
	100	2.765311×10^{-9}	Strongly favors H_1
	150	2.067733×10^{-10}	Strongly favors H_1
	200	3.773121×10^{-13}	Strongly favors H_1
y_2	25	5.432133×10^{-6}	Strongly favors H_1
	50	7.865514×10^{-14}	Strongly favors H_1
	100	6.876765×10^{-18}	Strongly favors H_1
	150	9.145433×10^{-20}	Strongly favors H_1
	200	4.112464×10^{-28}	Strongly favors H_1

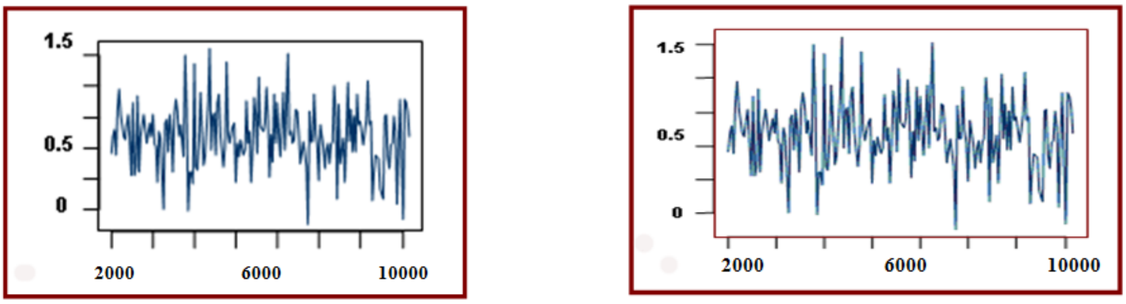


Figure 1. (10000) iterations of the Gibbs sampler for the first test function when $(n = 200)$.

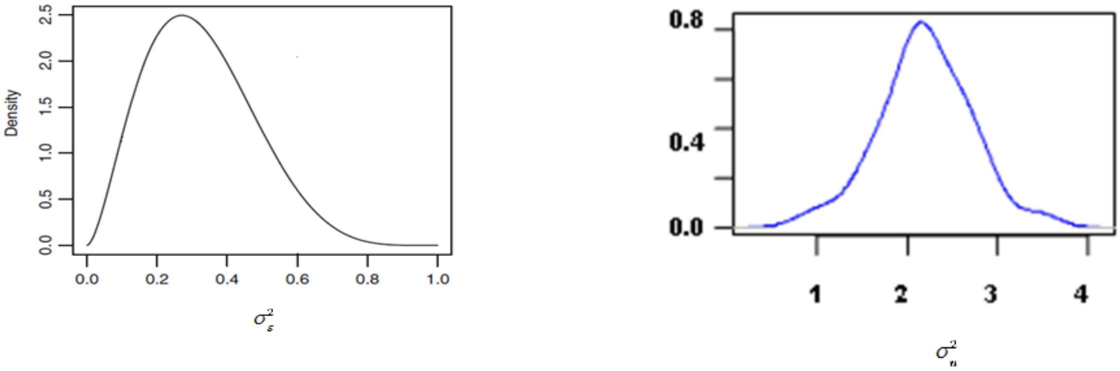


Figure 2. The density estimates based on (10000) iterations of σ_ε^2 and σ_u^2 for the first test function when $(n = 200)$.

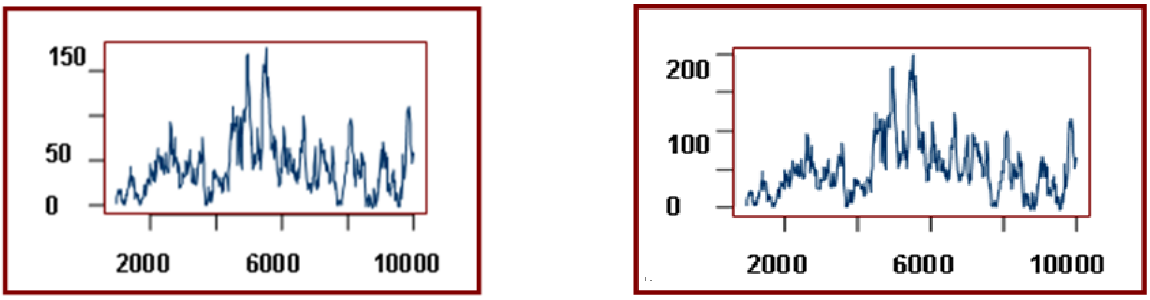
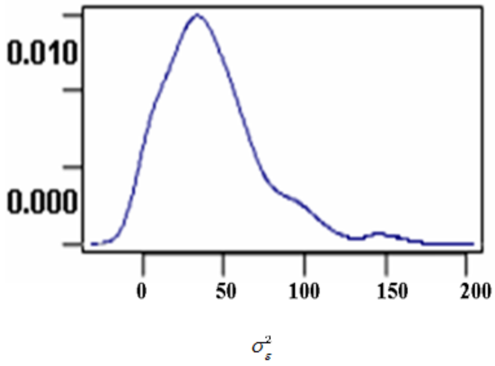


Figure 3. (10000) iterations of the Gibbs sampler for the second test function when $(n = 200)$.



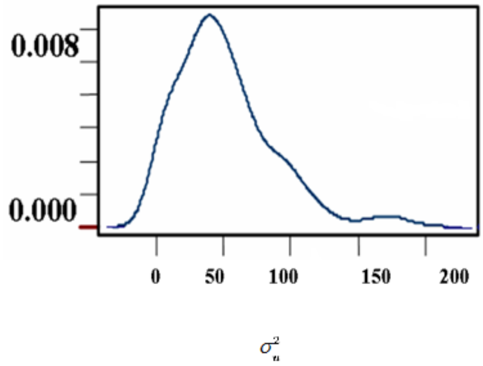


Figure 4. The density estimates based on (10000) iterations of σ_ε^2 and σ_u^2 for the second test function when $(n = 200)$.

6 Conclusions

The conclusions obtained throughout this paper are as follows:

(1) The joint posterior density of δ, γ given Y is

$$\pi_1(\delta,\gamma \mid Y) \propto \frac{\gamma^{(b/2)-1}\delta^{\alpha_\varepsilon-1}e^{\frac{\sigma_\beta^2\delta}{\beta_\varepsilon}}}{(a+b\gamma)^{-(a+b)/2}}\left(\prod_{i=1}^{p+q+1}(1+\delta d_i)\right)^{-1/2}\left(\prod_{i=p+q+2}^{n}(1+\gamma d_i)\right)^{-1/2}$$
$$\left(2\beta_\varepsilon+\sum_{i=1}^{p+q+1}\frac{s_i^2}{1+\delta d_i}+\sum_{i=p+q+2}^{n}\frac{s_i^2}{1+\gamma d_i}\right)^{-(n+2\alpha_\varepsilon+2)/2}$$

(2) The marginal density of Y under model $H_i, i = 0,1$ is:

$$m(Y \mid H_0) = (2\pi)^{-n/2}\frac{\beta_\varepsilon^{\alpha_\varepsilon}}{\Gamma(\alpha_\varepsilon)}\Gamma(\tfrac{n}{2}+\alpha_\varepsilon+1)\left(\beta_\varepsilon+\tfrac{1}{2}(Y-X\beta^0-Zu^0)^2\right)^{-\left(\frac{n}{2}+\alpha_\varepsilon+1\right)},$$

and

$$m(Y \mid H_1) = \frac{\beta_\varepsilon^{\alpha_\varepsilon}}{\Gamma(\alpha_\varepsilon)}(2\pi)^{-n/2}\Gamma\left(\frac{n}{2}+\alpha_\varepsilon\right)\int\left(\prod_{i=1}^{p+q+1}(1+\delta d_i)\right)^{-1/2}\left(\prod_{i=p+q+2}^{n}(1+\gamma d_i)\right)^{-1/2}\pi_0(\delta,\gamma)$$
$$\left(\beta_\varepsilon+\frac{1}{2}\left(\sum_{i=1}^{p+q+1}\frac{s_i^2}{1+\delta d_i}+\sum_{i=p+q+2}^{n}\frac{s_i^2}{1+\gamma d_i}\right)\right)d\delta d\gamma$$

(3) In the simulation results, we concluded the following:

(a) The values of $(AMSE)$ and $(AMAE)$ when $(n = 200)$ are smaller than their values for the first test function, which were (0.0005306171) and (0.000164242) respectively.

(b) The values of $(AMSE)$ and $(AMAE)$ are smaller when $(n = 200)$ for the second test function were (0.0001630011) and (0.000343007) respectively.

(c) The model corresponding to the first test function obtains the largest Bayes factor when $(n = 25)$ followed by that the second test function when $(n = 25)$.

(d) The Bayes factor favors H_1 with strong evidence with all samples sizes for two test functions.

Acknowledgments. We thank the editors and referees for providing critical comments which have brought significant improvements to our presentation.

References

1. Brezger, Andereas, Kneib, Thomas and Lang, Stefan, "Analyzing Bayesian semiparametric regression models", Work shop AG-Bayes, Universität München, 2002.
2. Choi, T., Lee, J. and Roy, A., "A note on the Bayes Factor in a Semiparametric regression model", Journal of Multivariate Analysis, 316-1327, 2008.
3. Jensen, M. J. and Maheu, J. M.,"Bayesian Semiparametric stochastic volatility modeling", The Rimini Center for Economic Analysis, Italy, 2007.
4. Jonsson, R. A. and Wichern, D. W. "Applied Multivariate Statistical Analysis" Prentice Hall, Englewood Cliffs, New Jersey 07632, 1988.
5. Jon, W., "Bayesian and frequentist regression methods, Springer New York Heidelberg Dordrecht London, (2013).
6. Lenk, P. J., "Bayesian inference for Semiparametric regression using a Fourier representation ", J.R. Statist. Soc. Ser. B61, part 4, 1999.
7. Mohaisen, A. J. and Abdulhussain, A. M., "A note on Bayes Semiparametric regression ", Mathematical Theory and Modeling www.iiste.org ISSN (Paper)2224-5804 ISSN (Online)2225-0522 Vol.3, No. 12, 2013.
8. Mohaisen, A. J. and Abdulhussain, A. M., "Bayesian Semiparametric regression with fuzzy sets ", International Journal of Pure and Applied Research in Engineering and Technology IJPRET, Vol. 2 (3):1-18 IJPRET, Research article ISSN: 2319-507X, 2013.
9. Mohaisen, A. J. and Abdulhussain, A. M., "Fuzzy sets and penalized spline in Bayesian Semiparametric regression ", LAP LAMBERT Academic Publishing, ISBN: 978-3-659-18439-0, 2014.
10. Natio, K., "Semiparametric regression with multiplicative adjustment, communications in statistics, theory and methods", 31 2289-2309, 2002.
11. Pelenis, J., "Bayesian Semiparametric Regression" Institute for Advanced Studies, Vienna 1-39, (2012).
12. Ruppert, D., Wand, M.P. and Carroll, R. J. "Semiparametric regression", Cambridge university press, (2003).
13. Temaratram, K., Robust estimation and model selection in Semiparametric regression models", Proefschrift voorgedragen tot het behalen van de grad van Doctor in de Toegepaste Economische Wetenschappen door, (2011).
14. Tsiatis, A. A. and MA, Y., "locally efficient Semiparametric estimators for functional measurement error models ", Biometrika, 91, 4, pp.835-848, 2004.
15. Wand, M.P. "Semiparametric regression and graphical models". Aust, N. Z. J. Stat., 51 (1), 2009.
16. Yuan, A. and DE Gooijer, J.G. "Semiparametric regression with kernel error model", Board of the Foundation of the Scandinavian Journal of statistics, Published by Blackwell Publishing Ltd, 2007.

