Reversible Jump MCMC To Estimate A Piecewise Constant Model With Exponential Additive Noise

Suparman

Abstract— Piecewise constant is a mathematical model that is often used to model data in various fields. Exponential multiplicative noise or exponential additive noise can be added in a constant piecewise model. This study aims to estimate a constant piecewise model that has exponential additive noise. The estimation of the constant piecewise model is carried out in the Bayesian framework. The prior distribution for the number of constant models, the location of the change in the constant model, the height of the constant model, and the noise variance selected. This prior distribution is combined with the probability function of the data to get the posterior distribution. The Bayes estimator for the number of constant models, the location of the change in the constant model, the height of the constant model, and the noise variance are estimated based on the posterior distribution. The Bayes estimator cannot be formulated explicitly because the number of constant models is a parameter. The reversible jump method of the Monte Carlo Markov Chain (MCMC) is proposed to determine the Bayes estimator. This study resulted in estimating the parameters of a constant piecewise model with exponential additive noise. This method can be used to estimate a constant piecewise model that has exponential noise even though the number of constant models is unknown.

Index Terms— Bayesian, Piecewise Constant, Exponential Additive Noise, Reversible Jump MCMC.

1 Introduction

Piecewise constant is a mathematical model used to model data in various fields of life, for example [1], [2], and [3]. The constant piecewise model is used for smoothing images of flowers [1]. The piecewise constant model is used for population size modeling [2], [3]. The mathematical model contains a noise. This noise is assumed to have a certain distribution.

The piecewise constant model can contain additive noise or multiplicative noise. Additive noise is used by various authors, for example [4], [5], and [6]. Additive noise is added to the spatial regression model [4]. Additive noise is used in partially linear functional [5]. This linear functional model is partly applied to the Tecator data. Additive noise is used in the log regression model [6]. On the other hand, multiplicative noise is also used by several authors, for example [7] and [8]. Multiplicative noise is used in a constant piecewise model [8]. Noise is assumed to be G (L, L) distribution where the L value is assumed to be known.

In various applications of autoregressive models, noise mathematical models are often assumed to be exponentially distributed, for example [9], [10], and [11]. Genetic algorithms are used to estimate the exponential autoregressive model [9]. Exponential is used as noise in the 1st order autoregressive model [10]. Bayesian robustness method is used to get the optimal Bayes estimator for exponentially distributed autoregressive models [11]. In the studies above, the autoregressive model order is assumed to be known. The autoregressive model has exponential noise but the order of unknown models is examined in [12].

However, a constant piecewise model that has additive exponential noise has not been studied. In some applications, data is often modeled following a piecewise constant model with exponential noise. If the piecewise model is constant with exponential additive noise used to model the data, the model parameters are unknown. The model parameters include the number of constant models, the location of constant model changes, the constant model height, and noise variance. This study proposes an estimation method of a constant piecewise model that has additive noise where the number of constant models is unknown.

• *Suparman is a lecturer in mathematics education, University of Ahmad Dahlan, Indonesia, PH-081328201198. E-mail: suparman@pmat.uad.ac.id*

2 Method

As in [13], the Bayesian framework is adopted to estimate parameters. The prior distribution for the number of constant models, the location of changes in the constant model, the constant height of the model, and the noise variance are selected. Then this prior distribution is combined with the likelihood function of the data to get the posterior distribution. Based on this posterior distribution, the Bayes estimator for the number of constant models, the location of changes in the constant model, the constant height of the model, and the noise variance are estimated.

The reversible jump Monte Carlo Markov Chain (MCMC) method [14] was proposed to determine the Bayes estimator. The basic idea of the MCMC reversible jump method is the creation of a Markov chain that is recurrent and irreducible such that the limit distribution of the Markov chain will be the same as the posterior distribution. Furthermore, the resulting Markov chain is used to calculate estimators for parameters.

3 Result and Discussion

Suppose that n represents the number of data and $y_1, \ldots, y_n$ represents the data set. This data follows the piecewise constant model if this data satisfies the following mathematical equations:

$$y_t = m_t + z_t, \qquad t = 1, \dots, n \tag{1}$$

where

$$m_t = \begin{cases} h_1, & \tau_1 < t \le \tau_2 \\ h_2, & \tau_2 < t \le \tau_3 \\ \dots & \dots \\ h_{k+1}, & \tau_{k+1} < t \le \tau_{k+2} \end{cases} \tag{2}$$

with $\tau_1 = 0$ and $\tau_{k+2} = n$. The value of k denotes the number of constant models. The values of $\tau = (\tau_1, \dots, \tau_k)$ state the location of the change in the constant model. The value of $h = (h_1, \dots, h_{k+1})$ expresses the height of the constant model. Here, z_t is assumed to be mutually independent and exponentially distributed with parameter $\sigma > 0$.

3.1 Likelihood Function

The random variable z_t is distributed exponentially so that

$$g(z_t|\sigma) = \sigma exp - \sigma z_t \tag{3}$$

Suppose that $y = (y_1, \dots, y_n)$.. The likelihood function for data y is

$$f(y|k,\tau,h,\sigma) = \prod_{i=1}^{k+1} \prod_{t=\tau_i+1}^{\tau_{i+1}} \sigma exp - \sigma(y_t - m_t) = \prod_{i=1}^{k+1} \sigma^{n_i} exp - \sigma s_i \tag{4}$$

where $s_i = \sum_{t=\tau_i+1}^{\tau_{i+1}} (y_t - m_t)$ and $n_i = \tau_{i+1} - \tau_i$ for $i = 1, \dots, k+1$.

3.2 Prior Distribution

To obtain a posterior distribution, a prior distribution must be determined. As in [8], the prior distribution for k is chosen of the Binomial distribution with parameter $0 < \lambda < 1$. For $k = 0, 1, \dots, k_{max}$

$$\pi(k|\lambda) = \binom{k_{max}}{k} \lambda^k (1-\lambda)^{k_{max}-k} \tag{5}$$

where k_{max} states the maximum value of k. The hyperprior distribution for λ is chosen the uniform distribution. Prior distribution for $\tau_1, \dots, \tau_k$ according to the ordered statistics.

$$\pi(\tau_1, \dots, \tau_k|k) = \frac{(2k+1)!}{n^{2k}} \frac{1}{2^k} \prod_{i=1}^{k+1} n_i \tag{6}$$

Prior distribution for $h_1, \dots, h_{k+1}$ exponential distribution is chosen with parameter $\nu > 0$.

$$\pi(h_1, \dots, h_{k+1}|k, \nu) = \prod_{i=1}^{k+1} \nu \, exp - \nu h_i = \nu^{k+1} exp - \nu \sum_{i=1}^{k+1} h_i \tag{7}$$

and Jeffreys prior is chosen as a hyperprior distribution for ν.

$$\pi(\nu) \propto \nu^{-1}$$

Similarly, Jeffreys prior is also chosen as a hyperprior distribution for σ.

$$\pi(\sigma) \propto \sigma^{-1}$$

So the prior distribution for the parameters $(k, \tau, h, \lambda, \nu, \sigma)$ can be written as

$$\pi(k,\tau,h,\lambda,\nu,\sigma) = \binom{k_{max}}{k} \lambda^k (1-\lambda)^{k_{max}-k} \frac{(2k+1)!}{n^{2k}} \frac{1}{2^k} \prod_{i=1}^{k+1} n_i \, \nu^{k+1} exp - \nu \sum_{i=1}^{k+1} h_i \, \sigma^{-1} \nu^{-1} \tag{8}$$

3.3 Posterior Distribution

Let $H_1 = (k, \tau, h, \sigma)$ and $H_2 = (\lambda, \nu)$. Posterior distribution can be written as

$$\pi(H_1, H_2|y) \propto \prod_{i=1}^{k+1} \sigma^{n_i} exp - \sigma s_i \binom{k_{max}}{k} \lambda^k (1-\lambda)^{k_{max}-k} \frac{(2k+1)!}{n^{2k}} \frac{1}{2^k} \prod_{i=1}^{k+1} n_i \, \nu^{k+1} exp - \nu \sum_{i=1}^{k+1} h_i \, \sigma^{-1} \nu^{-1}$$
$$\propto \prod_{i=1}^{k+1} \sigma^{n_i - 1} exp - \sigma s_i \binom{k_{max}}{k} \lambda^k (1-\lambda)^{k_{max}-k} \frac{(2k+1)!}{n^{2k}} \frac{1}{2^k} \prod_{i=1}^{k+1} n_i \prod_{i=1}^{k+1} \nu^k exp - \nu \sum_{i=1}^{k+1} h_i \tag{9}$$

3.4 Reversible Jump MCMC

Parameter estimation (H_1, H_2) are carried out using the Gibbs algorithm which consists of two stages, namely: distribution simulation $\pi(H_2|H_1, y)$ and distribution simulation $\pi(H_1|H_2, y)$. Distribution simulation $\pi(H_2|H_1, y)$ can be done using the Gamma distribution.

$$\pi(H_2|H_1, y) \propto \otimes_{i=1}^{k+1} G(n_i, s_i) \otimes G(k+1, \sum_{i=1}^{k+1} h_i) \otimes B(k+1, k_{max} - k + 1) \tag{10}$$

Therefore $\int_0^\infty \sigma^{n_i - 1} exp - \sigma s_i \, d\sigma = \frac{\Gamma(n_i)}{s_i^{n_i}}$,

$\int_0^1 \lambda^k (1-\lambda)^{k_{max}-k}\, d\lambda = \frac{\Gamma(\lambda+1)\Gamma(k_{max}+1)}{\Gamma(k_{max}+2)}$, and $\int_0^\infty \nu^k \exp - \nu \sum_{i=1}^{k+1} h_i \; d\nu = \frac{\Gamma(a)}{b^a}$ then

$$\pi(H_1|H_2,y) \propto \prod_{i=1}^{k+1} \frac{\Gamma(n_i)}{{s_i}^{n_i}} \binom{k_{max}}{k} \frac{\Gamma(\lambda+1)\Gamma(k_{max}+1)}{\Gamma(k_{max}+2)} \frac{(2k+1)!}{n^{2k}} \frac{1}{2^k} \prod_{i=1}^{k+1} n_i \frac{\Gamma(a)}{b^a} \quad (11)$$

where $a = k+1$ and $b = \sum_{i=1}^{k+1} h_i$. Distribution simulation $\pi(H_1|H_2,y)$ is done by using the MCMC reversible jump algorithm. As in [14], this algorithm uses four transformations, namely: changes in the height of the constant model, changes in the location of the constant model, the birth of the constant model, and the death of the constant model.

Changes in the high are as follows. Select a height randomly between $h_1, \ldots, h_{k+1}$. If h_j is selected, then the height of h_j is deleted and replaced by the height h_j^*. The height h_j^* is specified so that $log\left(\frac{h_j^*}{h_j}\right) = u$ where $u \sim U(-\frac{1}{2}, \frac{1}{2}$. So $h_j^* = h_j \exp u$.

Suppose that $x = (\tau_1, \ldots, \tau_k, h_1, \ldots, h_j, \ldots,\ h_{k+1,})$ and $x^* = (\tau_1, \ldots, \tau_{k,} h_1, \ldots, h_j^*, \ldots,\ h_{k+1,})$. Point x^* will replace x with probability

$$\rho(x,x^*) = min\left\{1, \frac{f(y|x^*)}{f(y|x)} \frac{\pi(x^*|k)}{\pi(x|k)} \frac{q(x^*,x)}{q(x,x^*)}\right\} \quad (12)$$

where

$$\frac{f(y|x^*)}{f(y|x)} = \frac{\prod_{i=1}^{k+1} \frac{\Gamma(n_i^*)}{{s_i}^{n_i^*}}}{\prod_{i=1}^{k+1} \frac{\Gamma(n_i)}{{s_i}^{n_i}}} \quad (13)$$

$$\frac{\pi(x^*|k)}{\pi(x|k)} = \frac{\Gamma(a^*)}{(b^*)^{a^*}} \frac{b^a}{\Gamma(a)} \quad (14)$$

$$\frac{q(x^*,x)}{q(x,x^*)} = \frac{h_j^*}{h_j} \quad (15)$$

The change in the location of the constant model is as follows. Take a location randomly between $\tau_1, \ldots, \tau_k$. If τ_j is selected, the location τ_j is deleted and replaced by location τ_j^*. Take u randomly according to $U(\tau_{j-1}, \tau_{j+1})$. So that $\tau_j^* = u$.

Suppose that $x = (\tau_1, \ldots, \tau_j, \ldots, \tau_k, h_1, \ldots,\ h_{k+1,})$ and $x^* = (\tau_1, \ldots, \tau_j^*, \ldots, \tau_{k,} h_1, \ldots,\ h_{k+1,})$. Point x^* will replace x with probability

$$\rho(x,x^*) = min\left\{1, \frac{f(y|x^*)}{f(y|x)} \frac{\pi(x^*|k)}{\pi(x|k)} \frac{q(x^*,x)}{q(x,x^*)}\right\} \quad (16)$$

where

$$\frac{f(y|x^*)}{f(y|x)} = \frac{\prod_{i=1}^{k+1} \frac{\Gamma(n_i^*)}{{s_i}^{n_i^*}}}{\prod_{i=1}^{k+1} \frac{\Gamma(n_i)}{{s_i}^{n_i}}} \quad (17)$$

$$\frac{\pi(x^*|k)}{\pi(x|k)} = \frac{(\tau_{j+1} - \tau_j^*)}{(\tau_{j+1} - \tau_j)} \quad (18)$$

$$\frac{q(x^*,x)}{q(x,x^*)} = \frac{(\tau_j^* - \tau_{j-1})}{(\tau_j - \tau_{j-1})} \quad (19)$$

Birth of the constant model is as follows. Take location τ^* randomly between $2, \ldots, n-1$. If $\tau^* \in (\tau_j, \tau_{j+1})$ then the height of h_j is deleted and replaced by the height h_j^* and h_{j+1}^* such that

$$(\tau^* - \tau_j)\log(h_j^*) + (\tau_{j+1} - \tau^*)\log(h_{j+1}^*) = (\tau_{j+1} - \tau_j)\log(h_j) \quad (20)$$

Suppose that $x = (\tau_1, \ldots, \tau_j, \ldots, \tau_k, h_1, \ldots, h_j, \ldots, h_{k+1,})$ and $x^* = (\tau_1, \ldots, \tau_j, \tau_j^*, \tau_{j+1} \ldots, \tau_{k,} h_1, \ldots,, h_{j-1}, h_j^*,\ h_{j+1}^*, h_{j+1}, \ldots,\ h_{k+1,})$. Point x^* will replace x with probability

$$\rho(x,x^*) = min\left\{1, \frac{f(y|x^*)}{f(y|x)} \frac{\pi(x^*|k)}{\pi(x|k)} \frac{q(x^*,x)}{q(x,x^*)}\right\} \quad (21)$$

where

$$\frac{f(y|x^*)}{f(y|x)} = \frac{\prod_{i=1}^{k+1} \frac{\Gamma(n_i^*)}{{s_i}^{n_i^*}}}{\prod_{i=1}^{k+1} \frac{\Gamma(n_i)}{{s_i}^{n_i}}} \quad (22)$$

$$\frac{\pi(x^*|k)}{\pi(x|k)} = \frac{n-1-k}{k} \frac{\lambda}{1-\lambda} \frac{(2k+3)(2k+2)}{(n-1)^2} \frac{1}{2} \frac{(\tau^* - \tau_j)(\tau_{j+1} - \tau^*)}{(\tau_{j+1} - \tau_j)} \quad (23)$$

$$\frac{q(x^*,x)}{q(x,x^*)} = \frac{n-1}{k+1} \frac{(h_j^* + h_{j+1}^*)^2}{h_j} \quad (24)$$

Death of the constant model is as follows. Take a location randomly between $\tau_1, \ldots, \tau_k$. If location τ_{j+1} is selected, then location τ_{j+1} is deleted. The height of h_j and h_{j+1} is also deleted and replaced by the height h_j^* such that

$$(\tau_{j+1} - \tau_j)\log(h_j) + (\tau_{j+2} - \tau_{j+1})\log(h_{j+1}) = (\tau_{j+2} - \tau_j)\log(h_j^*) \quad (25)$$

Suppose that $x = (\tau_1, \ldots, \tau_j, \tau_{j+1}, \tau_{j+2}, \ldots, \tau_k, h_1, \ldots, h_j, h_{j+1}, \ldots, h_{k+1})$ and

$x^* = (\tau_1, \dots, \tau_j, \tau_{j+2}, \dots, \tau_k, h_1, \dots, h_{j-1}, h_j^*, h_{j+2}, \dots, h_{k+1})$. Point x^* will replace x with probability

$$\rho(x, x^*) = min\left\{1, \frac{f(y|x^*)}{f(y|x)} \frac{\pi(x^*|k)}{\pi(x|k)} \frac{q(x^*, x)}{q(x, x^*)}\right\} \quad (26)$$

where

$$\frac{f(y|x^*)}{f(y|x)} = \frac{\prod_{i=1}^{k+1} \frac{\Gamma(n_i^*)}{{s_i}^{n_i^*}}}{\prod_{i=1}^{k+1} \frac{\Gamma(n_i)}{{s_i}^{n_i}}} \quad (27)$$

$$\frac{\pi(x^*|k)}{\pi(x|k)} = \frac{k}{n-k} \frac{1-\lambda}{\lambda} \frac{(n-1)^2}{(2k+1)(2k)} 2 \frac{(\tau_{j+1} - \tau_j)}{(\tau_{j+2} - \tau_{j+1})(\tau_{j+1} - \tau_j)} \frac{\Gamma(a^*)}{(b^*)^{a^*}} \frac{b^a}{\Gamma(a)} \quad (28)$$

$$\frac{q(x^*, x)}{q(x, x^*)} = \frac{k}{n-1} \frac{h_j^*}{(h_j + h_{j+1})^2} \quad (29)$$

This simulation produces a Markov chain that has a limit distribution to the posterior distribution. The resulting Markov chain can be used to estimate the piecewise constant model parameter.

4 Conclusion

The Bayes estimator cannot be formulated explicitly because the number of constant models is a parameter. The piecewise constant model parameter includes the number of constant models, the location of changes in the constant model, the constant model height, and noise variance.

The reversible jump MCMC algorithm can estimate the piecewise constant model parameter that has exponential noise simultaneously. This algorithm can also estimate the hyperparameters present in the prior distribution.

Acknowledgment

The author would like to thank to the University of Ahmad Dahlan for giving a research grant. Finally, thank you to the reviewers who have given suggestions.

References

[1] X. Pang, S. Zhang, J. Gu, L. Li, B.Liu, and H. Wang, Improved L0 Gradient Minimization Alt L1 Fidelity for Image Smoothing, PLOS ONE, 2015, 1-10.

[2] D. Zivkovic, M. Steinrucken, Y.S. Song, and W. Stephan, Transition Densities and Sample Frequency Spectra of Diffusion Processes with Selection and Variable Population Size, Genetics, 2015, 601-617.

[3] J.A. Kamm, J.P. Spence, J. Chan, and Y.S. Song, Two-Locus Likelihoods Under Variable Population Size and Fine-Scale Recombination Rate Estimation, Genetics, 2016, 1381-1399.

[4] S. Nandy, C.Y. Lim, and T. Maiti, Additive model building for spatial regression, J.R.Statist. Soc.B, 2017, 779-800.

[5] Y. Hu, S. Feng, and L. Xue, Automatic Variable Selection for Partially Linear Functional Additive Model and Its Application to the Tecator Data Set, Mathematical Problems in Engineering, 2018, 1-9.

[6] R. Richardson, H.D. Tolley, W.E. Evenson, and B.M. Lunt, Accounting for measurement error in log regression models with applications to accelerated testing, PLOS ONE, 2018, 1-13.

[7] T.A. Marques, Predicting and Correcting Bias Caused by Measurement Error in Line Transect Sampling Using Multiplicative Error Models, Biometrics, 2004, 757-763.

[8] Suparman and M. Doisy, Bayesian Segmentation in Signal with Multiplicative Noise Using Reversible Jump MCMC, Telkomnika, 2018, 673-680.

[9] Z. Shi and H. Aoyama, Estimation of the Exponential Autoregressive time series model by using the genetic algorithm, Journal of Sound and Vibration, 1997, 309-321.

[10] N. Sad, On Exponential Autoregressive Time Series Models, J. Math, 1999, 97-101.

[11] L. Larbi and H. Fellag, Robust Bayesian Analysis of an Autoregressive Model with Exponential Innovations, Afr. Stat., 2016, 955-964.

[12] Suparman, A New Estimation Procedure Using A Reversible Jump MCMC Algorithm for AR Models of Exponential White Noise, International Journal of GEOMATE, 2018, 85-91.

[13] Suparman and M.S. Rusiman, Hierarchical Bayesian Estimation for Stationary Autoregressive Models Using Reversible Jump MCMC Algorithm International Journal of Engineering & Technology, 2018, 64-67.

[14] P.J. Green, Reversible Jump MCMC Computation and Bayesian Model Determination. Biometrika, 1995, 711-732.