E-ISSN : 2278 – 179X

JECET; June 2016- August 2016; Sec. B; Vol.5. No.3.

Scientific and academic publication

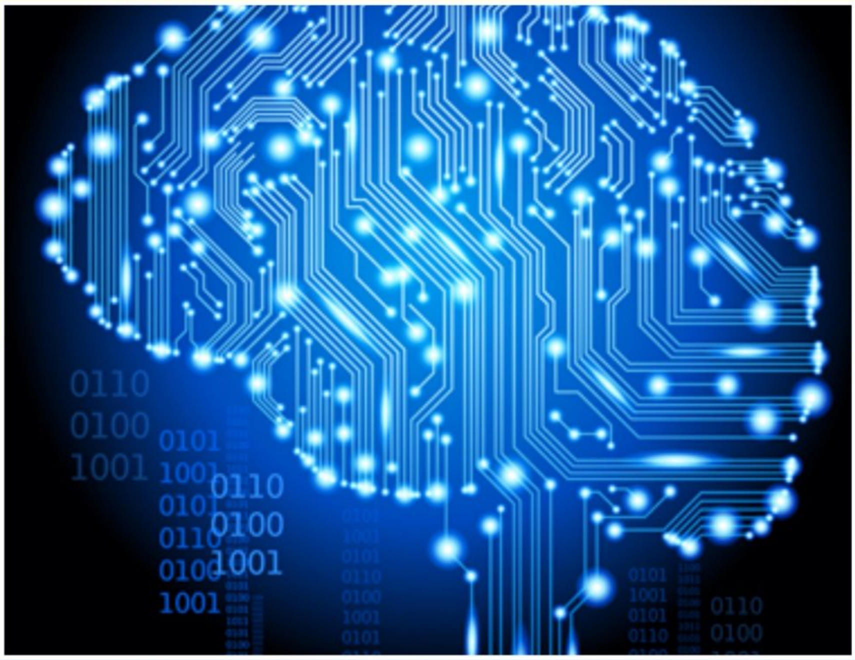
SJIF Impact Factor 2015 : 5.048

Journal of Environmental Science, Computer Science and Engineering & Technology

An International Peer Review E-3 Journal of Science

Section B: Computer Science

www. jecet.org



JECET : Journal of Environmental Science, Computer Science and Engineering & Technology

Indexing

CiteFactor — DOAJ Directory of Open Access Journals — EBSCO HOST — Genamics

Google Scholar — I2OR — Index Copernicus International — ISA Indian Science Abstracts

Open J-Gate — Journal Seek — New Jour — OCLC

SerialsSolutions — UlrichsWeb Global Serials Directory — OAJI Open Academic Journals Index

Editor-in-chief

Dr. (Mrs.) Divya Parashar
9, Geetanjali enclave, Malviya Nagar, New Delhi, 110017, India
9414245586, vkparashardholpur@gmail.com

Chief Editor

Reach us

Delhi office :
9, Geetanjali enclave, Malviya Nagar, New Delhi , 110017, India
Rajasthan office :
9, Ambika Dham, Dhoolkot chauraha, Pahada Road, Udaipur, 313001

JECET : Journal of Environmental Science, Computer Science and Engineering & Technology

Editorial Board

Dr. Syeda Azeem Unnisa
Research Officer,Osmania University,Hyderabad,A.P,India

Department of Mechanical Engineering, Faculty of Engineering, Rajasthan

JECET : Journal of Environmental Science, Computer Science and Engineering & Technology

Research Paper in Section B: Computer Science

Volume 5 and Issue 3

Below are the plagiarism and references report of MS# JECET: 1082 which you submitted for consideration for publication in the **Journal of Environmental Science, Computer Science and Engineering & Technology**. **We have checked your article in the plagiarism software and the following results observed/found:**

12.	**Plagiarism Report:**	We have checked your article in the plagiarism software and the following results observed/found:
	Abstract	**100% Unique Content and 00% copying from other journal/internet.**
	Introduction	**99% Unique Content and 01% copying from other journal/internet.**
	Likelihood Function	**100% Unique Content and 00% copying from other journal/internet.**
	BAYESIAN	**100% Unique Content and 00% copying from other journal/internet.**
	REVERSIBLE JUMP MCMC ALGORITHM	**100% Unique Content and 00% copying from other journal/internet.**
	Conclusion	**100% Unique Content and 00% copying from other journal/internet.**
	Overall	**99% Unique Content and 01% copying from other journal/internet.**
	Result of PlagLevel: the article having more than100% Unique Content and only 01% text of article match with the text of another journal/internet. The article is suitable for publication.	

References: we have checked the references in Google website and all the references are matched.

Below are the reviews of MS# JECET: 1082 which you submitted for consideration for publication in **Journal of Environmental Science, Computer Science and Engineering & Technology**. Here we are giving comments, suggestions and recommendations of the Editorial reviewer.

Editorial Recommendations:

Editorial General Comment

1 **Manuscript No** JECET 1082
Entitled: Selection in Subset Polynomial Regression Model by Using Reversible Jump MCMC Algorithm

2.	**Type of manuscript: Research Article/ Review article/ Short notes/Short communication**	**Research Article**
3.	**Is the length of the paper adequate?**	**Yes**
4.	**Are the title and the abstract pertinent, descriptive and concise?**	**Yes**
5.	**Does the paper contain new data or new ideas or both of them?**	**Yes**
6.	**Is the language fluent and precise?**	**Ok**
7.	**The figures are up to the mark in quality/ printing.**	**--**
8.	**Originality:**	**Yes**
9.	**References: I have checked all the references on Google and found all Correct.**	
10.	**Latest Reference.**	**2015**

Result of Plag Level: The article having more than 86% Unique Content and only 14% text of article match with the text of another journal/internet. The article is suitable for publication.

Object of research: developed the estimation of the parameter for subset polynomial regression model where the order is unknown with using reversible jump MCMC algorithm

- **Overall comments: Authors have reported the following major finding in the research work:**
- Reported that subset polynomial regression model is a model that is very flexible for modeling data.
- Subset polynomial regression is more flexible than full polynomial regression for modeling data. Subset polynomial is often used in many fields.
- Authors have proposes a method for selecting the subset polynomial regression where the order is unknown. The method used to estimate the parameters of the subset polynomial regression is the Bayesian method
- Authors suggested that If the subset polynomial regression model fitted to the data, the model parameter is generally unknown. For the order is known, many methods have proposed to estimate the model parameter.
- To overcome this problem, they have proposed the reversible jump MCMC method.
- Authors have described Likelihood Function and Bayesian approach.
- Described Reversible Jump MCMC Algorithm
- Authors suggested the advantage of this algorithm of this algorithm is both the order, coefficient and variance can simultaneously be estimated.

- They have also suggested that a comparison with other existing approach is currently under investigation.

12. **Recommendation: with consideration the above points, Article recommended for publication**

JECET; June 2016- August 2016; Sec. B; Vol.5. No.3, 214-219. E-ISSN: 2278–179X

Journal of Environmental Science, Computer Science and Engineering & Technology

An International Peer Review E-3 Journal of Sciences and Technology

Available online at www.jecet.org

Section B: Computer Science

Research Article

Selection in Subset Polynomial Regression Model by Using Reversible Jump MCMC Algorithm

Suparman

Department of Mathematics Education, Ahmad Dahlan University, Yogyakarta, Indonesia

Received: 29 July 2016; **Revised:** 13 August 2016; **Accepted:** 18 August 2016

Abstract: Subset polynomial regression is more flexible than full polynomial regression for modeling data. If the subset polynomial regression is fitted to the data, then the parameters are generally unknown. This paper proposes a method for selecting the subset polynomial regression where the order is unknown. The method used to estimate the parameters of the subset polynomial regression is the Bayesian method. However, the Bayesian estimator is analytically not able to be found. To solve these problems, the reversible jump MCMC algorithm is proposed. The key of this algorithm is the producing of the Markov Chain that converges to the limit distribution of the posterior distribution.

Keywords: Subset Polynomial Regression, Bayesian, Reversible Jump MCMC algorithm.

INTRODUCTION

Selection in subset polynomial regression is to find a subset of the available subset polynomial regressions that does a good predicting. Subset polynomial regression model is a model that is very flexible for modeling data. Subset polynomial is often used in many fields. For example, it is used in the health for diagnosing breast cancer[1] and for segmenting medical image[2].

If the subset polynomial regression model fitted to the data, the model parameter is generally unknown. For the order is known, many methods have proposed to estimate the model parameter, for example[3]. In fact, the order is unknown. Because the order of the subset polynomial regression is unknown, previous method cannot used to estimate the parameter of the subset polynomial regression. To overcome this problem, the reversible jump MCMC[4] is proposed. There are many subset polynomial regressions; here they are limited to the subset polynomial regression model with the mean 0 and variance unknown.

1. LIKELIHOOD FUNCTION

Let $y=(y_1,\cdots,y_n)$ be a dependent variable and let $x=(x_1,\cdots,x_n)$ be a independent variable where n is the number of observations. The subset polynomial regression can be written by:

$$y_t=\beta_0+\beta_{n_1}x_t^{n_1}+\cdots+\beta_{n_p}x_t^{n_p}+z_t$$

The set $\{n_1,\cdots,n_p\}$ is a subset of the set {1, ..., p} where p is the order. For a polynomial regression with p = 3, there are $2^4=16$ subset polynomial regressions. In this case, the residual z_t are normally distributed with mean 0 and variance σ^2, denoted by $z_t\sim N(0,\sigma^2)$. So for $t=1,2,\cdots,n$

$$f(z_t|\ \sigma^2)=\frac{1}{\sqrt{2\pi\sigma^2}}\exp-\frac{1}{2\sigma^2}z_t^2.$$

If the variable transformation $z_t=y_t-\beta_0-\beta_{n_1}x_t^{n_1}-\cdots-\beta_{n_p}x_t^{n_p}$ is used, then $\frac{dz_t}{dy_t}=1$. So that

$$f(y_t|\ p,\beta,\sigma^2,x_t)=\frac{1}{\sqrt{2\pi\sigma^2}}\exp-\frac{1}{2\sigma^2}(y_t-\beta_0-\beta_{n_1}x_t^{n_1}-\cdots-\beta_{n_p}x_t^{n_p})^2$$

If $\beta=(\beta_0,\beta_{n_1},\cdots,\beta_{n_p})$, then the likelihood function is

$$\begin{aligned}L(y|\ p,\beta,\sigma^2,x) &=\prod_{t=1}^{n}\frac{1}{\sqrt{2\pi\sigma^2}}\exp-\frac{1}{2\sigma^2}(y_t-\beta_0-\beta_{n_1}x_t^{n_1}-\cdots-\beta_{n_p}x_t^{n_p})^2\\ &=(2\pi\sigma^2)^{-\frac{n}{2}}\exp-\frac{1}{2\sigma^2}\sum_{t=1}^{n}(y_t-\beta_0-\beta_{n_1}x_t^{n_1}-\cdots-\beta_{n_p}x_t^{n_p})^2.\end{aligned}$$

2. BAYESIAN

To use the Bayesian approach[5], the prior distribution is selected. The same choice with[6], the prior distribution of p is a binomial distribution with hyper parameter λ, denoted by $p\sim BIN(\lambda,p_{max})$,

$$\pi(p|\lambda)=\binom{p_{maks}}{p}\lambda^p(1-\lambda)^{p_{maks}-p}$$

Where $p=0,1,2,\cdots,p_{maks}$.

The prior distribution of (β,σ^2) is a Jeffreys prior distribution,

$$\pi(\beta,\sigma^2|p,\lambda)\propto\sigma^{-2}.$$

As in[7], the hyper parameter λ is considered as a variable. So, hyper prior distribution of λ is also selected. The prior distribution of λ is a Jeffreys prior distribution,

$$\pi(\lambda)\propto[\lambda(1-\lambda)]^{-\frac{1}{2}}.$$

Therefore, the prior distribution of $(p,\beta,\sigma^2,\lambda)$ is

$$\begin{aligned}\pi(p,\beta,\sigma^2,\lambda) &= \pi(\beta,\sigma^2|p,\lambda)\ \pi(p|\lambda)\ \pi(\lambda)\\ &\propto \binom{p_{maks}}{p}\lambda^p\,(1-\lambda)^{p_{maks}-p}\ \sigma^{-2}\ [\lambda(1-\lambda)]^{-\frac{1}{2}}.\end{aligned}$$

Using Bayes theorem, the posterior distribution of $(p,\beta,\sigma^2,\lambda)$ is written in the following way:

$$\begin{aligned}\pi(p,\beta,\sigma^2,\lambda|y,x) &\propto L(y|\ p,\beta,\sigma^2,x)\ \pi(p,\beta,\sigma^2,\lambda)\\ &\propto (2\pi\sigma^2)^{-\frac{n}{2}}\exp-\frac{1}{2\sigma^2}\sum_{t=1}^{n}\ (y_t-\beta_0-\beta_{n_1}x_t^{n_1}-\cdots-\beta_{n_p}x_t^{n_p})^2\\ &\quad \binom{p_{maks}}{p}\lambda^p\,(1-\lambda)^{p_{maks}-p}\ \sigma^{-2}\ [\lambda(1-\lambda)]^{-\frac{1}{2}}\\ &\propto (2\pi)^{-\frac{n}{2}}\ (\sigma^2)^{-\frac{n}{2}-1}\exp-\frac{1}{2\sigma^2}\\ &\quad \sum_{t=1}^{n}\ (y_t-\beta_0-\beta_{n_1}x_t^{n_1}-\cdots-\beta_{n_p}x_t^{n_p})^2\\ &\quad \binom{p_{maks}}{p}\lambda^{p+\frac{1}{2}-1}(1-\lambda)^{p_{maks}-p+\frac{1}{2}-1}\end{aligned}$$

3. REVERSIBLE JUMP MCMC ALGORITHM

Suppose that $\xi=(p,\beta,\sigma^2,\lambda)$. The MCMC algorithm is a method of sampling to produce a homogeneous Markov Chain $\xi_1,\xi_2,\cdots,\xi_m$ that satisfies aperiodic and irreductibel[8] such that $\xi_1,\xi_2,\cdots,\xi_m$ can be considered as a random variable whose distribution $\pi(p,\beta,\sigma^2,\lambda|y,x)$. Thus $\xi_1,\xi_2,\cdots,\xi_m$ can be used to estimate parameter ξ. The algorithm consists of three steps:

1. Simulate $\sigma^2 \sim \pi(\sigma^2|\ p,\beta,\lambda,y,x)$
2. Simulate $\lambda \sim \pi(\lambda|\ p,\beta,\sigma^2,y,x)$
3. Simulate $(p,\beta) \sim \pi(p,\beta|\sigma^2,\lambda,y,x)$

Because the conditional distribution of σ^2 given (β,p,λ) is

$$\pi(\sigma^2|\ p,\beta,\lambda,y,x) = IG(\frac{n}{2},s^2)$$

where $s^2 = \frac{1}{2}\sum_{t=1}^{n}(y_t - \beta_0 - \beta_{n_1}x_t - \cdots - \beta_{n_p}x_t^{n_p})$, it is easy to simulate $\sigma^2 \sim IG(\frac{n}{2},s^2)$. In the same reason, since the conditional distribution of λ given (β,p,σ^2) is

$$\pi(\lambda|\ p,\beta,\sigma^2,y,x) = \mathrm{Beta}\,(p+\frac{1}{2},p_{max}-p+\frac{1}{2}),$$

It is also easy to generate $\lambda \sim \mathrm{Beta}\,(p+\frac{1}{2},p_{max}-p+\frac{1}{2})$.

The conditional distribution of (p,β) given (σ^2,λ) is

$$\pi(p,\beta|\sigma^2,\lambda,y,x) = \int_0^\infty \int_0^1 \pi(p,\beta,\sigma^2,\lambda|y,x)d\lambda d\sigma^2$$

$$\propto (2\pi)^{-\frac{n}{2}} \frac{\Gamma(1+\frac{1}{2})\Gamma(p_{maks}-p+\frac{1}{2})}{\Gamma(p_{maks}+1)}$$

$$\frac{\Gamma(\frac{n}{2})}{[\frac{1}{2}\sum_{t=1}^{n}(y_t-\beta_0-\beta_{n_p}x_t^{n_p}-\cdots-\beta_{n_1}x_t)]^{n/2}}.$$

But because the conditional distribution of (p,β) given (σ^2,λ) is unknown distribution, it is not easy to simulate $(p,\beta) \sim \pi(p,\beta|\sigma^2,\lambda,y,x)$. To simulate it, the reversible jump MCMC algorithm is used.

Let $\xi = (p,\beta)$ be the actual point of the Markov chain. There are three types of transformations are used, namely: the birth of the order, the death of the order and the change of the coefficient. Further suppose that N_p is the probability of transformation from p to p + 1, D_p is the probability of transformation from p + 1 to p, and C_p is the probability of transformation from p to p (the same order).

3.1. Change of the coefficient: The transformation of the change of coefficient does not change the order. Let $\xi=(p,\beta)$ be a current point and let $\xi^*=(p^*,\beta^*)$ be a updated point. If the change of the coefficient is selected, then $p^*=p$ and simulate $\beta^*\sim N(\beta,\sigma^2)$. Let α_p be the probability of acceptance for the change of the coefficient. This probability of acceptance is written by

$$\alpha_p=\min\left\{1,\exp-\frac{1}{\sigma^2}(s^{*2}-s^2)\right\}$$

where $s^2=\frac{1}{2}\sum_{t=1}^{n}(y_t-\beta_0^*-\beta_{n_1}^*x_t-\cdots-\beta_{n_p}^*x_t^{n_p})$.

3.2 Birth of order: Transformation of the birth of order changes from p to p+1. Let $\xi=(p,\beta)$ be a current point and let $\xi^*=(p^*,\beta^*)$ be a updated point. If the birth of the coefficient is selected, then $p^*=p+1$, $\beta_i^*=\beta_i-u$ and $\beta_{i+1}^*=\beta_i+u$ where $u\sim U(0,1)$. Let α_n be the probability of acceptance for the birth of order. This probability of acceptance is given by

$$\alpha_n=\min\left\{1,\ \frac{p_{max}-p}{p+1}\frac{\lambda}{1-\lambda}2\exp-\frac{1}{\sigma^2}(s^{*2}-s^2)\right\}.$$

3.3 Death of order: Transformation of the death of order changes from p+1 to p. Let $\xi=(p+1,\beta)$ be a current point and let $\xi^*=(p^*,\beta^*)$ be a updated point. If the death of the coefficient is selected, then $p^*=p$ and $\beta_i^*=\frac{(\beta_i+\beta_{i+1})}{2}$. Let α_d be the probability of acceptance for the death of order. This probability of acceptance is defined by

$$\alpha_d=\min\left\{1,\frac{p+1}{p_{max}-p}\frac{1-\lambda}{\lambda}\frac{1}{2}\exp-\frac{1}{\sigma^2}(s^{*2}-s^2)\right\}.$$

4. CONCLUSION

This paper developed the estimation of the parameter for subset polynomial regression model where the order is unknown with using reversible jump MCMC algorithm. The advantage of this algorithm of this algorithm is both the order, coefficient and variance can simultaneously be estimated. Because the polynomial regression is special case of this subset polynomial regression, then this algorithm can use to estimate the parameter of the polynomial regression. A comparison with other existing approach is currently under investigation.

ACKNOWLEDGMENT

I would like to thank the Ahmad Dahlan University in Yogyakarta who has supported the grant for this research.

REFERENCES

1. T. O'Neill, J. Penm, and J.S. Penm, A Subset Polynomial Neural Networks Approach for Breast Cancer Diagnosis, *International Journal of Electronic Healtheare*, 2007, 3, 3, 293-302.
2. C.H. Xie, Y.J. Liu and J.Y. Chang, Medical Image Segmentation Using Rough Set and Local Polynomial Regression. *Multimedia Tools and Aplplications*, 2015,74,6, 1885-1914.
3. M.J. Brusco, D. Steinley and J.D. Cradit, an Exact Algorithm for Hierarchically Well-Formulated Subset in Second-Order Polynomial Regression. *Technometrics*, 2009, 51, 3, 306-315.
4. P.J. Green, Reversible Jump MCMC Computation and Bayesian Model Determination. *Biometrika*, 1995, 82,711-732.
5. C.P. Robert, the Bayesian Choice: A Decision-Theory Motivation, Springer, New York, 2001.
6. Suparman and M. Doisy, Bayesian Segmentation of Piecewise Linear Regression Models Using Reversible Jump MCMC Algorithm. *Computer Technology and Application*, 2015, 6, 1, 14-18.
7. Suparman and M. Doisy, Hierarchical Bayesian of ARMA Models Using Simulated Algorithm, *Telkomnika*, 2014,12,1, 87-96.
8. C.P. Robert and G. Casella, Monte Carlo Statistical Methods, Springer, New York, 1999.

*** Corresponding author: Suparman;**

Department of Mathematics Education, Ahmad Dahlan University, Yogyakarta, Indonesia