talenta PUBLISHER

Journal of Research in Mathematics Trends and Technology

Estimation of Heteroskedasticity Semiparametric Regression Curve Using Fourier Series Approach

Rahmawati Pane[1*], Sutarman[1], Andi Tenri Ampa[2]

[1]*Department of Mathematics, Universitas Sumatera Utara, Medan, 20155, Indonesia*
[2]*Department of Statistics, Institut Teknologi Sepuluh Nopember, Surabaya, 60111, Indonesia*

Abstract. A heteroskedastic semiparametric regression model consists of two main components, i.e. parametric component and nonparametric component. The model assumes that any data $(\underset{\sim}{x}_i', t_i, y_i)$ follows $y_i = \underset{\sim}{x}_i'\underset{\sim}{\beta} + f(t_i) + \sigma_i\varepsilon_i$, where $i = 1,2,\dots,n$, $\underset{\sim}{x}_i' = (1, x_{i1}, x_{i2}, \dots, x_{ir})$ and t_i is the predictor variable. Parameter vector $\underset{\sim}{\beta} = (\beta_1, \beta_2, \dots, \beta_r)' \in \Re^r$ is unknown and $f(t_i)$ is also unknown and is assumed to be in interval of $C[0,\pi]$. Random error ε_i is independent on zero mean and variance σ^2. *Estimation of* the heteroskedastic semiparametric regression model was conducted to evaluate the parametric and nonparametric components. The nonparametric component $f(t_i)$ regression was approximated by Fourier series $F(t) = bt + \frac{1}{2}\alpha_0 + \sum_{k=1}^{K}\alpha_k \cos kt$. The estimation was obtained by means of Weighted Penalized Least Square (WPLS): $\min_{f\in C(0,\pi)}\left\{n^{-1}\left(\underset{\sim}{y} - X\underset{\sim}{\beta} - \underset{\sim}{f}\right)'W^{-1}\left(\underset{\sim}{y} - X\underset{\sim}{\beta} - \underset{\sim}{f}\right) + \lambda\int_0^{\pi}\frac{2}{\pi}[f''(t)]^2dt\right\}$. The WPLS solution provided nonparametric component $\underset{\sim}{\hat{f}}_\lambda(t) = M(\lambda)\underset{\sim}{y}^*$ for a matrix $M(\lambda)$ and parametric component $\underset{\sim}{\hat{\beta}} = [X'T(\lambda)X]^{-1}X'T(\lambda)\underset{\sim}{y}$.

Keyword: Fourier Series, Heteroskedastic Semiparametric Regression, Bandwith Parameter, WPLS

Abstrak. *Suatu model regresi semiparametrik heteroskedastis terdiri atas dua komponen utama, yaitu komponen parametrik dan komponen nonparametrik. Model ini mengasumsikan bahwa sebarang data* $(\underset{\sim}{x}_i', t_i, y_i)$ *dengan* $y_i = \underset{\sim}{x}_i'\underset{\sim}{\beta} + f(t_i) + \sigma_i\varepsilon_i$, *untuk* $i = 1,2,\dots,n$, $\underset{\sim}{x}_i' = (1, x_{i1}, x_{i2}, \dots, x_{ir})$ *dan* t_i *merupakan variabel prediksi. Vektor parameter* $\underset{\sim}{\beta} = (\beta_1, \beta_2, \dots, \beta_r)' \in \Re^r$ *tidak diketahui dan* $f(t_i)$ *adalah fungsi yang tidak diketahui bentuknya dan diasumsikan termuat dalam ruang fungsi kontinu* $C[0,\pi]$. *Error random* ε_i *independen dengan mean nol dan variansi* σ^2. *Estimasi dari model regresi semiparametric heteroskedastik dilakukan untuk mengevaluasi komponen parametrik dan nonparametrik. Kurva regresi komponen nonparametrik* $f(t_i)$ *dihampiri dengan deret Fourier* $F(t) = bt + \frac{1}{2}\alpha_0 + \sum_{k=1}^{K}\alpha_k \cos kt$. *Estimasi kurva regresi semiparametric heteroskedastik diperoleh dari menyelesaikan optimasi Weighted Penalized Least Square (WPLS):* $\min_{f\in C(0,\pi)}\left\{n^{-1}\left(\underset{\sim}{y} - X\underset{\sim}{\beta} - \underset{\sim}{f}\right)'W^{-1}\left(\underset{\sim}{y} - X\underset{\sim}{\beta} - \underset{\sim}{f}\right) + \lambda\int_0^{\pi}\frac{2}{\pi}[f''(t)]^2dt\right\}$. *Solusi dari WPLS di atas memberikan estimator komponen nonparametric dalam bentuk* $\underset{\sim}{\hat{f}}_\lambda(t) = M(\lambda)\underset{\sim}{y}^*$ *untuk suatu matriks* $M(\lambda)$ *dan komponen parametrik* $\underset{\sim}{\hat{\beta}} = [X'T(\lambda)X]^{-1}X'T(\lambda)\underset{\sim}{y}$.

Kata Kunci: *Deret Fourier, Regresi Semiparametrik Heteroskedastik, Parameter Bandwidth, WPLS*

*Corresponding author at: Department of Mathematics, Universitas Sumatera Utara, Medan, 20155, Indonesia

E-mail address: rahmawatipane@usu.ac.id

 e-ISSN: 2656-1514, DOI: 10.32734/jormtt.v2i1.3744
Journal Homepage: http://talenta.usu.ac.id/jormtt

Received 5 December 2019 | Revised 14 January 2020 | Accepted 17 February 2020

1. Introduction

Regression analysis is an important method in statistics. The objectives are to investigate the relationship between the predictor variables and the response variables and to evaluate the contribution of the predictor variable $x_i' = (1, x_{i1}, x_{i2}, \dots, x_{ir})$ and $\mathrm{t_i}$ to the response variable y_i. When the relationship of the variables is known, the regression is called parametric regression [1]. Otherwise, the regression is identified as nonparametric regression.

In many cases, the response variable has linear relationship with one of the predictor variables, but with other predictor variables, it remains unknown. If this happens, Wahba [2] suggested the use of semiparametric regression approach. Semiparametric regression is a combination of parametric with nonparametric regressions. According to [3] and [4], semiparametric regression model assumed that data (x_i', t_i, y_i) follows model $y_i = x_i'\beta + f(t_i) + \sigma_i\varepsilon_i$ where $i = 1,2, \dots, n$, and x_i' and t_i are the predictor variables. $x_i'\beta$ is the parametric component and $f(t_i)$ is the nonparametric component. The semiparametric regression has been widely used with various smoothing approaches. Some researches applied Spline function [5], [2], Wavelet [6], Kernel [7], [8], [9], [10], or local polynomial on the nonparametric component.

In fact, some data have periodic properties which are in a good agreement with Fourier series function. The Fourier series is known as a trigonometric polynomial that is very "flexible" to effectively approximate complicated functions. The Fourier series is best applied to describe any function whose terms is sines and cosines. The Fourier series estimator, in general, is used when the data in a question is periodic and unknown [11], [12]. A nonparametric regression by means of Fourier series has been developed by Bilodeau [12] to represent periodic functions with the same trend and variance (homoskedastic). However, a serious problem arises when it is applied to a heteroskedastic function; a periodic function with different random error variance (heteroskedastic). Therefore, it is crucial to develop an estimator based on Fourier series which includes elements of the trend, and the variance of the error in heteroskedastic semiparametric regression model.

In this report, a Fourier series estimator was derived to estimate the parametric and nonparametric components of a heteroskedastic function by means of semiparametric regression. The result could be an alternative model to the data pattern recurring/ periodic with trends in semiparametric regression heteroskedastic.

2. Semiparametric Regression Model

A semiparametric regression model is a combination of parametric with nonparametric regression models. The model is employed when the relationship between the predictor variable and the response variable can be described as certain curves. Given that paired data $(\underset{\sim}{t}_i, \underset{\sim}{x}_i', y_i)$ and $\underset{\sim}{x}_i' = (x_{i1}, x_{i2}, \dots, x_{ir})$ is the predictor variables that the relationship with the response variable y_i is known, and $\underset{\sim}{t}_i = \left(t_{\mathrm{i}1}, t_{\mathrm{i}2}, \dots, t_{pi}\right)'$ is the predictor variable that the relationship with the response variable is unknown. The relationship of the variables $(\underset{\sim}{t}_i, \underset{\sim}{x}_i', y_i)$ is assumed to follow a heteroskedastic semiparametric regression model:

$$y_i = \underset{\sim}{x}_i'\underset{\sim}{\beta} + f(t_i) + \sigma_i\varepsilon_i \ , \quad i = 1,2, \dots, n \tag{1}$$

The parameter $\underset{\sim}{\beta} = (\beta_1, \beta_2, \dots, \beta_r)' \in \Re^r$ is a vector parameter of the unknown parametric component with size of $r \times 1$. The random error ε_i, $i = 1,2, \dots, n$ is mutually independent with zero mean and variance σ^2. $f(t_i)$ is the nonparamteric component with unknown function and is assumed to be defined in $C[0,\pi]$, where $C[0,\pi] = \{f, f$ is a continuous function in the interval of $[0,\pi]\}$.

2.1. Estimation of Fourier Series-Based Heteroskedastic Semiparametric Regression Model

Estimating the semiparametric regression model by Fourier series is to define the parametric component $\underset{\sim}{x}_i'\underset{\sim}{\beta}$ as well as the nonparametric component $f(t_i)$. The nonparametric component of the heteroskedastic semiparametric regression curve can be approached using Fourier series as follows:

$$F(t) = bt + \frac{1}{2}\alpha_0 + \sum_{k=1}^{K} \alpha_k \cos kt \tag{2}$$

The estimators for regression curve f and parameter $\underset{\sim}{\beta}$ can be obtained by solving Weighted Penalized Least Square (WPLS) optimization:

$$\min_{f \in C(0,\pi)} \left\{ n^{-1}\left(\underset{\sim}{y} - X\underset{\sim}{\beta} - \underset{\sim}{f}\right)' W^{-1}\left(\underset{\sim}{y} - X\underset{\sim}{\beta} - \underset{\sim}{f}\right) + \lambda \int_0^{\pi} \frac{2}{\pi}[f''(t)]^2 dt \right\} \tag{3}$$

where the weighted matrix $W = \mathrm{diag}(\sigma_1^2, \sigma_2^2, \dots, \sigma_n^2)$, $\underset{\sim}{y} = (y_1, y_2, \dots, y_n)'$ and $\underset{\sim}{f}$ have size of $(n \times 1)$ and λ is the bandwith parameter. The WPLS optimization was executed by taking goodness of fit $G(f)$ and penalty $P(f)$ of the Eq. 3.

Where
$$G\left(\underset{\sim}{f}\right) = n^{-1}\left(\underset{\sim}{y} - X\underset{\sim}{\beta} - \underset{\sim}{f}\right)' W^{-1}\left(\underset{\sim}{y} - X\underset{\sim}{\beta} - \underset{\sim}{f}\right) \tag{4}$$

and
$$P(f) = \int_0^{\pi} \frac{2}{\pi}[f''(t)]^2 dt \tag{5}$$

the heteroskedastic semiparametric regression model from Eq. 1 can be rewritten in the form of:

$$\underset{\sim}{y} = X\underset{\sim}{\beta} + B(\underset{\sim}{t})\underset{\sim}{\delta} + W^{1/2}\underset{\sim}{\varepsilon}$$

Where

$$\mathrm{X} = \begin{pmatrix} 1 & x_{11} & x_{21} & L & x_{r1} \\ 1 & x_{12} & x_{22} & L & x_{r2} \\ M & M & M & M & M \\ 1 & x_{1n} & x_{2n} & L & x_{rn} \end{pmatrix}, \mathrm{B}(\underset{\sim}{t}) = \begin{pmatrix} t_1 & 1 & cos\, t_1 & cos\, 2t_1 & L & cos\, Kt_1 \\ t_2 & 1 & cos\, t_2 & cos\, 2t_2 & L & cos\, Kt_2 \\ M & M & M & M & M & M \\ t_n & 1 & cos\, t_n & cos\, 2t_n & L & cos\, Kt_n \end{pmatrix},$$

$W^{1/2} = \mathrm{diag}(\sigma_1, \sigma_2, \dots, \sigma_n)$ and $\underset{\sim}{\delta} = \left(b, \frac{1}{2}\alpha_0, \alpha_1, \alpha_2, \dots, \alpha_K\right)' \in \Re^{(K+2)}$.

The goodness of fit from Eq. 4 can be defined as:

$$G(\underset{\sim}{\delta}) = n^{-1}\left(\underset{\sim}{y} - X\underset{\sim}{\beta} - B(\underset{\sim}{t})\underset{\sim}{\delta}\right)' W^{-1}\left(\underset{\sim}{y} - X\underset{\sim}{\beta} - B(\underset{\sim}{t})\underset{\sim}{\delta}\right) \tag{6}$$

Since $F(t) = bt + \frac{1}{2}\alpha_0 + \sum_{k=1}^{K}\alpha_k \cos kt$, then:

$$\int_0^{\pi}\frac{2}{\pi}[f''(t)]^2 dt = \int_0^{\pi}\frac{2}{\pi}\left[bt + \frac{1}{2}\alpha_0 + \sum_{k=1}^{K}\alpha_k \cos kt\right]^2 dt.$$

In other expression, the last equation can be rewritten as follows:

$$\int_0^{\pi}\frac{2}{\pi}[f''(t)]^2 dt = \sum_{k=1}^{K} k^4\alpha_k^2.$$

Therefore:

$$\lambda\int_0^{\pi}\frac{2}{\pi}[f''(t)]^2 dt = \lambda\underset{\sim}{\delta}'\begin{pmatrix} 0 & 0 & 0 & L & 0 \\ 0 & 0 & 0 & L & 0 \\ 0 & 0 & 1^4 & L & 0 \\ M & M & M & 0 & M \\ 0 & 0 & 0 & L & K^4 \end{pmatrix}\underset{\sim}{\delta} = \lambda\underset{\sim}{\delta}' D\underset{\sim}{\delta} \tag{7}$$

with matrix $D = \mathrm{diag}(0,0,1^4, 2^4, \dots, K^4)$. If the goodness of fit (4) and penalty (5) are combined, then the WPLS optimization (3) can be expressed as:

$$\begin{aligned} &\min_{\underset{\sim}{\delta}\in\Re^{(K+2)}}\{G(\underset{\sim}{\delta}) + \lambda P(\underset{\sim}{\delta})\} \\ &\quad = \min_{\underset{\sim}{\delta}\in\Re^{(K+2)}}\left\{n^{-1}\left(\underset{\sim}{y} - X\underset{\sim}{\beta} - B(\underset{\sim}{t})\underset{\sim}{\delta}\right)' W^{-1}\left(\underset{\sim}{y} - X\underset{\sim}{\beta} - B(\underset{\sim}{t})\underset{\sim}{\delta}\right) + \lambda\underset{\sim}{\delta}' D\underset{\sim}{\delta}\right\} \\ &\quad = \min_{\underset{\sim}{\delta}\in\Re^{(K+2)}}\{Q(\underset{\sim}{\delta})\} \end{aligned} \tag{8}$$

Taking the partial derivative of the Eq. (8) yields:

$$\begin{aligned} \frac{\partial\{Q(\underset{\sim}{\delta})\}}{\partial\underset{\sim}{\delta}} = \frac{\partial}{\partial\underset{\sim}{\delta}}\{&n^{-1}\underset{\sim}{y}^{*\prime}W^{-1}\underset{\sim}{y}^{*} - 2n^{-1}\underset{\sim}{\delta}'B'(\underset{\sim}{t})W^{-1}\underset{\sim}{y}^{*} \\ &+ \underset{\sim}{\delta}'(n^{-1}B'(\underset{\sim}{t})W^{-1}B(\underset{\sim}{t}) + \lambda D)\underset{\sim}{\delta}\} \end{aligned} \tag{9}$$

The normal solution of the Eq. (9) is

$$-2n^{-1}B'(\underset{\sim}{t})W^{-1}\underset{\sim}{y}^{*} + 2(n^{-1}B'(\underset{\sim}{t})W^{-1}B(\underset{\sim}{t}) + \lambda D)\underset{\sim}{\delta} = \underset{\sim}{0}$$

$$-2(n^{-1}B'(\underset{\sim}{t})W^{-1}B(\underset{\sim}{t}) + \lambda D)\underset{\sim}{\delta} = -2n^{-1}B'(\underset{\sim}{t})W^{-1}\underset{\sim}{y}^{*}$$

The estimator for $\underset{\sim}{\delta}$ is given as:

$$\underset{\sim}{\hat{\delta}}_{\lambda} = (n^{-1}B'(\underset{\sim}{t})W^{-1}B(\underset{\sim}{t}) + \lambda D)^{-1}n^{-1}B'(\underset{\sim}{t})W^{-1}\left(\underset{\sim}{y} - X\underset{\sim}{\hat{\beta}}\right).$$

The Fourier series estimator for the estimation curve of the nonparametric component in the heteroskedastic semiparametric regression is define as:

$$\begin{aligned}\hat{\underset{\sim}{f}}_{\lambda}(t) &= B(t)\hat{\underset{\sim}{\delta}}_{\lambda} \\ &= B(t)(n^{-1}B'(t)W^{-1}B(t)+\lambda D)^{-1}n^{-1}B'(t)W^{-1}\left(\underset{\sim}{y}-X\hat{\underset{\sim}{\beta}}\right) \\ &= H(\lambda)\underset{\sim}{y}^{*}\end{aligned} \tag{10}$$

where $\underset{\sim}{y}^{*} = \underset{\sim}{y} - X\hat{\underset{\sim}{\beta}}$ and $H(\lambda) = B(t)(n^{-1}B'(t)W^{-1}B(t)+\lambda D)^{-1}n^{-1}B'(t)W^{-1}$.

The nonparametric component $\hat{\underset{\sim}{f}}_{\lambda}(t)$ is somewhat depending on the parametric component $\hat{\underset{\sim}{\beta}}$. To search for the parametric component estimator $\underset{\sim}{\beta}$, a least mean square method was used:

$$\underset{\sim}{y} = X\underset{\sim}{\beta} + B\underset{\sim}{\alpha} + \underset{\sim}{\varepsilon}$$

$$\underset{\sim}{y} = X\underset{\sim}{\beta} + \hat{\underset{\sim}{f}}_{\lambda}(t) + \underset{\sim}{\varepsilon}$$

$$\underset{\sim}{\varepsilon} = y - X\underset{\sim}{\beta} - \hat{\underset{\sim}{f}}_{\lambda}(t)$$

$$\varepsilon_i = y_i - \underset{\sim}{x}_i'\underset{\sim}{\beta} - \hat{\underset{\sim}{f}}_{\lambda}(t)$$

$$\sum_{i=1}^{n}\varepsilon_i^2 = \sum_{i=1}^{n}\left(y_i - \underset{\sim}{x}_i'\underset{\sim}{\beta} - \hat{\underset{\sim}{f}}_{\lambda}(t)\right)^2$$

$$\begin{aligned}\varepsilon'\varepsilon &= \left(y_i - \underset{\sim}{x}_i'\underset{\sim}{\beta} - \hat{\underset{\sim}{f}}_{\lambda}(t)\right)'\left(\left(y_i - \underset{\sim}{x}_i'\underset{\sim}{\beta} - \hat{\underset{\sim}{f}}_{\lambda}(t)\right)\right) \\ &= \left[\underset{\sim}{y} - X\underset{\sim}{\beta} - \left(H(\lambda)\left(y - X\underset{\sim}{\beta}\right)\right)\right]'\left[\underset{\sim}{y} - X\underset{\sim}{\beta} - \left(H(\lambda)\left(y - X\underset{\sim}{\beta}\right)\right)\right] \\ &= \underset{\sim}{y}'\left(I-H(\lambda)\right)'\left(I-H(\lambda)\right)\underset{\sim}{y} - 2\underset{\sim}{\beta}'X'\left(I-H(\lambda)\right)'\left(I-H(\lambda)\right)\underset{\sim}{y} \\ &\qquad + \underset{\sim}{\beta}'X'\left(I-H(\lambda)\right)'\left(I-H(\lambda)\right)X\underset{\sim}{\beta}\end{aligned}$$

Assume that

$$\begin{aligned}&\underset{\sim}{y}'\left(I-H(\lambda)\right)'\left(I-H(\lambda)\right)\underset{\sim}{y} - 2\underset{\sim}{\beta}'X'\left(I-H(\lambda)\right)'\left(I-H(\lambda)\right)\underset{\sim}{y} \\ &\qquad + \underset{\sim}{\beta}'X'\left(I-H(\lambda)\right)'\left(I-H(\lambda)\right)X\underset{\sim}{\beta} = Q\left(\underset{\sim}{\beta}\right)\end{aligned}$$

Deriving $Q\left(\underset{\sim}{\beta}\right)$ on $\underset{\sim}{\beta}$ and taking its zero value will yield:

$$\begin{aligned}\frac{\partial\{Q(\underset{\sim}{\beta})\}}{\partial\underset{\sim}{\beta}} &= \frac{\partial}{\partial\underset{\sim}{\beta}}\Big\{\underset{\sim}{y}'\left(I-H(\lambda)\right)'\left(I-H(\lambda)\right)\underset{\sim}{y} - 2\underset{\sim}{\beta}'X'\left(I-H(\lambda)\right)'\left(I-H(\lambda)\right)\underset{\sim}{y} \\ &\qquad + \underset{\sim}{\beta}'X'\left(I-H(\lambda)\right)'\left(I-H(\lambda)\right)X\underset{\sim}{\beta}\Big\} \\ &= -2\underset{\sim}{\beta}'X'\left(I-H(\lambda)\right)'\left(I-H(\lambda)\right)\underset{\sim}{y} + 2X'\left(I-H(\lambda)\right)'\left(I-H(\lambda)\right)X\underset{\sim}{\beta} \\ -2\underset{\sim}{\beta}'X'\left(I-H(\lambda)\right)'\left(I-H(\lambda)\right)\underset{\sim}{y} &+ 2X'\left(I-H(\lambda)\right)'\left(I-H(\lambda)\right)X\underset{\sim}{\beta} = 0 \\ -2X'\left(I-H(\lambda)\right)'\left(I-H(\lambda)\right)X\underset{\sim}{\beta} &= -2\underset{\sim}{\beta}'X'\left(I-H(\lambda)\right)'\left(I-H(\lambda)\right)\underset{\sim}{y} \\ \hat{\underset{\sim}{\beta}} &= \left[X'\left(I-H(\lambda)\right)'\left(I-H(\lambda)\right)X\right]^{-1}X'\left(I-H(\lambda)\right)'\left(I-H(\lambda)\right)\underset{\sim}{y} \\ \hat{\underset{\sim}{\beta}} &= [X'T(\lambda)X]^{-1}X'T(\lambda)\underset{\sim}{y}\end{aligned}$$

which is the parametric component estimator, where $T(\lambda) = \big(I - H(\lambda)\big)'\big(I - H(\lambda)\big)$. Furthermore

$$\begin{aligned}\hat{\underset{\sim}{f}}_{\lambda}(t) &= H(\lambda)\big(\underset{\sim}{y} - X\underset{\sim}{\beta}\big)\\ &= H(\lambda)\underset{\sim}{y} - H(\lambda)X[X'T(\lambda)X]^{-1}X'T(\lambda)\underset{\sim}{y}\\ &= H(\lambda)\underset{\sim}{y}\{I - X[X'T(\lambda)X]^{-1}X'T(\lambda)\}\\ &= M(\lambda)\underset{\sim}{y} \qquad (11)\end{aligned}$$

is the nonparametric component estimator with $M(\lambda) = H(\lambda)\{I - X[X'T(\lambda)X]^{-1}X'T(\lambda)\}$. Hence, the heteroskedastic semiparametric regression model for the Eq. (1) is:

$$\begin{aligned}\hat{\underset{\sim}{y}} &= X\hat{\underset{\sim}{\beta}} + \hat{\underset{\sim}{f}}_{\lambda}(t)\\ &= X[X'T(\lambda)X]^{-1}X'T(\lambda)\underset{\sim}{y} + M(\lambda)\underset{\sim}{y}\\ &= \{X[X'T(\lambda)X]^{-1}X'T(\lambda) + M(\lambda)\}\underset{\sim}{y}\\ &= N(\lambda)\underset{\sim}{y} \qquad (12)\end{aligned}$$

where: $N(\lambda) = X[X'T(\lambda)X]^{-1}X'T(\lambda) + M(\lambda)$.

3. Conclusions

Given that a heteroskedastic semiparametric regression model $y_i = \underset{\sim}{x}_i'\underset{\sim}{\beta} + f(t_i) + \sigma_i\varepsilon_i,\ i = 1,2,\dots,n$. The nonparametric regression component f form is unknown and is assumed to be smooth; it is defined in a continuous function space $C[0,\pi]$. The random error ε_i is mutually independent on the zero mean and variance σ^2. Fourier series from Eq. 2 is used to approace the nonparametric component regression curve $f(t)$. The estimation curve of the nonparametric component of the heteroskedastic semiparametric regression was obtained by solving WPLS optimization:

$$\min_{f\in C(0,\pi)}\left\{n^{-1}\big(\underset{\sim}{y} - X\underset{\sim}{\beta} - \underset{\sim}{f}\big)'W^{-1}\big(\underset{\sim}{y} - X\underset{\sim}{\beta} - \underset{\sim}{f}\big) + \lambda\int_0^{\pi}\frac{2}{\pi}[f''(t)]^2dt\right\}$$

a). The solution of the WPLS optimization gave the Fourier series-based nonparametric component estimator, which was described in $\hat{\underset{\sim}{f}}_{\lambda}(t) = M(\lambda)\underset{\sim}{y}$, for matrix $M(\lambda)$ and parametric component $\hat{\underset{\sim}{\beta}} = [X'T(\lambda)X]^{-1}X'T(\lambda)\underset{\sim}{y}$.

b). The finest estimator for nonparametric component by Fourier series in the heteroskedasticity semiparametric regression model is strongly dependent on the optimum values of bandwith parameter λ and K that can be obtained by various methods, for instance by means of generalized cross validation (GCV).

REFERENCES

[1] I. N. Budiantara, "Spline Model with Optimal Knot", *Journal of Basic Science FMIPA Jember University* vol. 7, no. 1, pp. 77-85, 2006.

[2] G. Wahba, *Spline Models for Observational Data*, Philadelphia: Society for Industrial and Applied Mathematics, 1990.

[3] W. Hardle, H. Liang and J. Gao, *Partially Linear Models*, Physica-Verlag Heidelberg, 2003.

[4] W. Hardle, Y. Mori and P. Vieu, *Statistical Methods for Biostatistics and Related Fields*, New York: Springer-Verlag Berlin Heidelberg, 2007.

[5] H. Becher, G. Kauermann, P. Khomski and B. Kouyaté, "Using Penalized Splines to Model Age- and Season-of-birth Dependent Effects of Childhood Mortality Risk Factors in Rural Burkina Faso", *Biometrical Journal*, vol. 51, no. 1, pp. 110-122, 2009.

[6] A. Antoniadis, J. Bigot and T. Sapatinas, "Wavelet Estimators in Nonparametric Regression: A Comparative Simulation Study", *Journal of Statistical Software*, vol. 6, pp. 1-83, 2001.

[7] M. Y. Cheng, R. L. Paige, S. Sun and K. Yan, "Variance Reduction for Kernel Estimatiors in Clustered/Longitudinal Data Analysis", *Journal of Statistical Planning and Inference*, vol. 140, no. 6, pp. 1389-1397, 2010.

[8] N. Wang, "Marginal Nonparametric Kernel Regession Accounting for Within-Subject Correlation", *Biometrika*, vol. 90, no. 1, pp. 43-52, 2003.

[9] M. Kayri and G. Zirhlioğlu, "Kernel Smoothing Function and Choosing Bandwitdh for Non-parametric Regression Methods", *Ozean Journal of Applied Sciences*., vol. 2, no. 1, pp. 49-60, 2009.

[10] C. O. Wu and C. T. Chiang, "Kernel Smoothing on Varying Coefficient Models with Longitudinal Dependent Variable", *Statistica Sinica*, vol. 10, no. 2, pp. 433-456, 2000.

[11] A. Tripena and I. N. Budiantara, *Fourier Estimator in Nonparametric Regression*, International Conference on Natural Sciences and Applied Natural Sciences, Ahmad Dahlan University, Yogyakarta, 2007.

[12] M. Bilodeau, "Fourier Smoother and Additive Models", *The Canadian Journal of Statistics*, vol. 20, no. 3, pp. 257-269, 1992.

[13] I. N. Budiantara, "Inferensi Statistik untuk Model Spline", *Jurnal Mat-Stat*, vol. 7, no. 1, pp. 1-14, 2007.