International Journal of Control Theory and Applications

ISSN : 0974-5572

Volume 10 • Number 4 • 2017

The Efficiency of Curve Regression in Semiparametric Smoothing Spline Regression Models with and Without Penalty

Solimun*, Umu Sa'adah, Suci Astutik, Adji Achmad Rinaldo Fernandes

Statistics Department, Faculty of Mathematics and Natural Sciences, University of Brawijaya, Indonesia
**Corresponding Author: solimun@ub.ac.id*

Abstract: The aim of this study are (1) obtain the estimation of smoothing spline-based Penalized Least Square (PLS) semiparametric regression (with a penalty), (2) obtain the estimation of smoothing spline-based Ordinary Least Square (OLS) semiparametric regression (without a penalty), (3) obtain the properties of the estimation of smoothing spline semiparametric regression function, (4) test the efficiency of curve regression estimation in the model of semiparametric regression in application data; i.e. the factors affecting poverty in East Java, Indonesia. The result show: (1) The estimation of spline function estimation in semiparametric regression with penalty (using Penalized Least Square or PLS) is $\hat{\underset{\sim}{f}}_\lambda = \mathbf{A}_\lambda \underset{\sim}{y}$,with

$$\mathbf{A}_\lambda = \mathbf{T}\left(\mathbf{T}^{\mathrm{T}}\hat{\mathbf{U}}^{-1}\hat{\boldsymbol{\Sigma}}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^{\mathrm{T}}\hat{\mathbf{U}}^{-1}\hat{\boldsymbol{\Sigma}}^{-1} + \mathbf{V}\hat{\mathbf{U}}^{-1}\hat{\boldsymbol{\Sigma}}^{-1}[\mathbf{I} - \mathbf{T}\left(\mathbf{T}^{\mathrm{T}}\hat{\mathbf{U}}^{-1}\hat{\boldsymbol{\Sigma}}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^{\mathrm{T}}\hat{\mathbf{U}}^{-1}\hat{\boldsymbol{\Sigma}}^{-1}].$$

The estimation of spline function estimation in semiparametric regression without penalty (using Ordinary Least Square or OLS) is $\hat{\underset{\sim}{f}} = \mathbf{A}y$, with $\mathbf{A} = \mathbf{T}\left(\mathbf{T}^{\mathrm{T}}\hat{\boldsymbol{\Sigma}}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^{\mathrm{T}}\hat{\boldsymbol{\Sigma}}^{-1}$. The properties of Semiparametric Regression Using Smoothing Spline is linear and bias characteristic of the estimator of nonparametric regression curve. The comparing of smoothing spline regression analysis with and without the penalty that is applied to the data of the rate of poverty in East Java. Research Data used is secondary data obtained from in Indonesian Badan Pusat Statistik / the National Statistical Bureau of East Java Province, with the rate of poverty (y), Labor Force Participation Rate (LFPR) and the Human Development Index (HDI) as predictors in 38 cities / districts in East Java in 2013, The efficient of semiparametric regression curve without penalty (OLS) amounted to 24.35% of semiparametric regression curve with a penalty (PLS). In other words, smoothing spline semiparametric regression model with a more efficient with penalty than without penalty.

Keywords: Smoothing Spline, Semiparametric, With Penalty, Without Penalty

1. INTRODUCTION

Regression analysis was first discovered by Galton [1] who conducted a study on the effect of height boy with his father's height. The study found that the height of the boy's father had a tendency to decrease high (Regress) approaching the midpoint population after several generations. That is, the son of the father of a very high tend to be shorter than his father and the son of a father who was very short cenderng taller than his father. Currently, the term regression applied to all kinds of forecasting, and not necessarily implies a regression approaches the middle of the population [1]

Regression analysis is an analysis method that analyzes the influence of predictors to response. In general regression analysis explained by the model

$$y_i = f(x_i) + \varepsilon_i$$

The curve $f(x_i)$ is the regression function and μ_i is assumed to be normally independent distributed errors with mean of 0 and the variance of $\tilde{A}^2$. In their use, the regression analysis is divided into three approaches, namely parametric, nonparametric, and semiparametric.

A regression analysis when the form of the regression curve is known then the regression model approach is called parametric regression models. Likewise, Nonparametric approach is used because the form does not follow a particular curve which is provided in parametric approaches, such as liner, cubic, quadratic, exponential and others. In the nonparametric regression, the form of regression curve is assumed to be unknown. A nonparametric regression curve is simply assumed to be smooth within the meaning contained in certain function spaces. Data are expected to seek their own estimation, thus the nonparametric regression approach has high flexibility [2].

In addition to parametric and nonparametric approach, regression analysis also has a semiparametric approach. Semiparametric regression is a combination of regression parametric and nonparametric regression. The combination in this case meant that the same semiparametric regression contains parametric regression model and nonparametric regression model. Semiparametric regression emerged due to cases of modeling which its relationships between variables in addition are linear and there are also unknown shapes [3].

To obtain the estimation of semiparametric regression function, especially for nonparametric components, can be done with optimization of Penalized Least Square PLS which adds a penalty for surface roughness. This study focused on the effectiveness test of the use of penalties in seminparametric regression. This study therefore aims to (1) obtain the estimation of smoothing spline-based PLS semiparametric regression (with a penalty), (2) obtain the estimation of smoothing spline-based OLS semiparametric regression (without a penalty), (3) obtain the properties of the estimation of smoothing spline semiparametric regression function, (4) test the efficiency of curve regression estimation in the model of semiparametric regression in application data; i.e. the factors affecting poverty in East Java, Indonesia.

East Java is a province that has the largest area among 6 provinces in Java with an area of 47 922 km^2, inside of which there are 38 districts / cities [4]. Besides East Java province is ranked second most populous province with 37,476,757 inhabitants, or it can be said that the population density of 784 inhabitants per square kilometers [5]. The objectives listed in the Sustainable Development Goals (SDGs) are an end to poverty. SDGs is a form of world action against a variety of social, economic, and environmental problems. Poverty appointed as the main purpose is certainly not without reason. The increase of welfare measurued in the reduction of poverty is a reflection of the success of the development of every country, including Indonesia [6]. To measure poverty, BPS (National Statistic Bureau) has used the concept of meeting basic needs. With this approach, poverty is seen as an economic inability to meet the basic needs of food and non-food which is measured from the expenditure side.

2. THEORETICAL REVIEW

2.1. Spline Theory in Nonparametric Regression

Regression analysisis used to predict relationship between predictor and response through estimation of regression curve *f*. Nonparametric regression approach is used when the regression curve *f* is not or has yet been identified, so that the shape of the curve is determined based on the existing data [1]. Some studies have developed spline estimator in nonparametric regression model for cross-section data, for instance Budiantara (2000) [9], Huang (2003) [10], Crainiceanu, Ruppert & Wand (2004) [11], Kim & Gu (2004) [12], Lee (2004) [13], and Howell (2007) [14].

Nonparametric regression model for cross-sectiondata that establish relationship between predictor and response for cross-section data that involves N is formulated as follow:

$$y_i = f(x_i) + \varepsilon_i; \; i = 1, 2, ..., N. \tag{1}$$

Notes:

y_i : response in ith observation,

x_i : predictor in ith observation,

f : regression curve for the relationship between predictor and response,

N : number of observation/ subject,

ε_i : random error in ith observation..

ε_i is random error that is independent to each other, normally distributed with the average of 0 and constant variance of σ^2 [15].

Spline approach in general specifies f in equation (2) into unknown regression curve. The assumption is that regression curve f is smooth that is included in particular function especially Sobolevor written as $f \in W_2^m[a,b]$ with:

$$W_2^m[a,b] = \left\{ f : f, f^{(1)}, ..., f^{(m-1)} \; absolutely\, continuous; \int_a^b [f^{(m)}(x_i)]^2 dx_i < \infty \right\}, \tag{2}$$

for one constant m that declares spline polynomial order. In order to obtain the estimation of regression curve f using optimization:

$$\underset{f \in W_2{}^m[a,b]}{\text{Min}} \left[N^{-1} \sum_{i=1}^{N} \left(y_i - f(x_i) \right)^2 \right],$$

with the requirement of $\int_a^b [f^{(m)}(x_i)]^2 dx_i \leq \gamma$,for $\gamma \geq 0.$

Spline is function obtained by minimizing Penalized Least Square PLS that refers to estimation criteria that combine goodness of fit and penalty function [9,16]. Using model (2) above, estimator of regression curve f is obtained by minimizing the following PLS:

$$\underset{f \in W_2{}^m[a,b]}{\text{Min}} \left[N^{-1} \sum_{i=1}^{N} \left(y_i - f(x_i) \right)^2 + \lambda \int_a^b (f^{(m)}(x_i))^2 dx_i \right], \tag{3}$$

for each f in Sobolev $W_2^m[a,b]$.

The first segment of equation (2) is function that measures goodness of fit, while the second segment is measuring roughness penalty) with λ as smoothing parameter *p* controlling goodness of fit and roughness penalty.The optimization above does not take roughness penalty into account so it is equivalent to Ordinary Least Square OLS [17].

PLS optimizationcan be done using RKHS approach. Hilbert space presents particular entity in RKHS linear operation where function *f* is unidentified and assumed smooth in Hilbert space. Then, Hilbert space is decompositioned into the addition of two spaces, H_0 and H_1, where $H = H_0 \oplus H_1$, with $H_0 = H_1^{\perp}$. As an example, basis for H_0 space is $\{\phi_{i1}, \phi_{i1}, \ldots, \phi_{im}\}$ and basis for H_1 space is $\{\xi_{i1}, \xi_{i1}, \ldots, \xi_{im}\}$, then for every function $f_i \in H$ it can be presented separately such as:

$$f_i = g_i + h_i,$$

Therefore, for every function $f_i \in H$ it can be presented separately such as:

$$f_i = g_i + h_i = \sum_{j=1}^{m} d_{ij}\phi_{ij} + \sum_{j=1}^{m} c_{ij}\xi_{ij} \text{ [18].} \tag{4}$$

2.2. Semiparametric Regression for Smoothing Spline

Problems often arose in the regression are that not all the explanatory variables can be approximated by a parametric approach, the absence of information about the relationship form between the predictor and the response, so it must be done a nonparametric approach. By combining parametric and nonparametric approaches in a regression approach, we will get a semiparametric model [8]. Simple linear semiparametric equation is as follows:

$$y_i = \beta_0 + \beta_1 x_i + f + \varepsilon_i \tag{5}$$

Where x_i : a predictor for parametric components

t_i : a predictor for nonparametric component

y_i : a response

f : nonparametric function with unknown shape

ε_i : An error which has an average 0 and variance of σ^2

So the equation (7) yields the equation in matrix form as follows:

$$\underset{\sim}{y} = \mathbf{x}\underset{\sim}{\beta} + \underset{\sim}{f} + \underset{\sim}{\varepsilon} \tag{6}$$

3. RESULT AND DISCUSSION

3.1. Estimation of Spline Function in Semiparametric Regression with Penalty

The first goal of the study is to obtain spline function estimation in semiparametric regression with penalty using Penalized Least Square or PLS. It is presented in Theorem 1 as follow:

Theorem 1: Penalized Least Square

When given a pair of data following nonparametric regression model that meets the shape of nonparametric regression function described in equation (4), with assumptions $\mathrm{E}(\underset{\sim}{\varepsilon}) = \underset{\sim}{0}$ and $\mathrm{Var}(\underset{\sim}{\varepsilon}) = \boldsymbol{\Sigma}$, then spline estimator minimizing PLS (when homoskedasticity is fulfilled then $\boldsymbol{\Sigma} = \mathbf{I}$)

$$\min_{f \in W_2^m[a,b]} \left\{ M^{-1}\left(\underset{\sim}{y} - \underset{\sim}{f}\right)^{\mathrm{T}} \boldsymbol{\Sigma} \left(\underset{\sim}{y} - \underset{\sim}{f}\right) + \lambda \int_a^b \left[f^{(m)}(x_t) \right]^2 dx_t \right\} \tag{7}$$

The spline estimator is $\hat{\underset{\sim}{f}}_\lambda = \mathbf{A}_\lambda \underset{\sim}{y}$, where:

$$\mathbf{A}_\lambda = \mathbf{T}\left(\mathbf{T}^{\mathrm{T}}\hat{\mathbf{U}}^{-1}\hat{\boldsymbol{\Sigma}}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^{\mathrm{T}}\hat{\mathbf{U}}^{-1}\hat{\boldsymbol{\Sigma}}^{-1} + \mathbf{V}\hat{\mathbf{U}}^{-1}\hat{\boldsymbol{\Sigma}}^{-1}[\mathbf{I} - \mathbf{T}\left(\mathbf{T}^{\mathrm{T}}\hat{\mathbf{U}}^{-1}\hat{\boldsymbol{\Sigma}}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^{\mathrm{T}}\hat{\mathbf{U}}^{-1}\hat{\boldsymbol{\Sigma}}^{-1}].$$

$$\hat{\mathbf{U}} = \hat{\boldsymbol{\Sigma}}^{-1}\mathbf{V} + M\boldsymbol{\Lambda}.$$

Proof: Considering the equation that is function $\underset{\sim}{f} = \mathbf{T}\underset{\sim}{d} + \mathbf{V}\underset{\sim}{c}$ by Fernandes (2015) [17], then the nonparametric regression model (1) can be stated as:

$$\underset{\sim}{y} = \underset{\sim}{f} + \underset{\sim}{\varepsilon} = \mathbf{T}\underset{\sim}{d} + \mathbf{V}\underset{\sim}{c} + \underset{\sim}{\varepsilon}.$$

where **T** is (*T*)´(*m*) matrix as follow:

$$\mathbf{T} = \begin{pmatrix} \langle \eta_1, \phi_1 \rangle & \langle \eta_1, \phi_2 \rangle & \cdots & \langle \eta_1, \phi_m \rangle \\ \langle \eta_2, \phi_1 \rangle & \langle \eta_2, \phi_2 \rangle & \cdots & \langle \eta_2, \phi_m \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle \eta_T, \phi_1 \rangle & \langle \eta_T, \phi_2 \rangle & \cdots & \langle \eta_T, \phi_m \rangle \end{pmatrix},$$

$$\langle \eta_t, \phi_j \rangle = \frac{x_t^{\,j-1}}{(j-1)!}, \text{with } t = 1, 2, ..., T;\ j = 1, 2, ..., m$$

$\underset{\sim}{d}$ is *m*-sized vector, from:

$$\underset{\sim}{d}' = (d_1, d_2, ..., d_m),$$

Vis (*T*)´(*T*)-sized matrix as follow:

$$\mathbf{V} = \begin{pmatrix} \langle \xi_1, \xi_1 \rangle & 0 & \cdots & 0 \\ \langle \xi_2, \xi_1 \rangle & \langle \xi_2, \xi_2 \rangle & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \langle \xi_T, \xi_1 \rangle & \langle \xi_T, \xi_2 \rangle & \cdots & \langle \xi_T, \xi_T \rangle \end{pmatrix},$$

$$\left\langle \xi_t, \xi_s \right\rangle = \int_a^b \frac{(x_t - u)_+^{m-1}(x_s - u)_+^{m-1}}{\left((m-1)!\right)^2} du, t = 1,2,...,T; s = 1,2,...,T \tag{8}$$

$\underset{\sim}{c}$ is T-sized vector, from:

$$\underset{\sim}{c}' = (c_1, c_2, ..., c_T).$$

Nonparametric regression analysis is conducted to get estimator of regression curve $\underset{\sim}{f}$. To get the estimation, Reproducing Kernel Hilbert Space (RKHS) is used. The purpose is to obtain the estimation of $\underset{\sim}{f}$ that meets PLS optimization:

$$\min_{f\in\mathcal{H}} \left\{ \left\| \boldsymbol{\Sigma}^{-\frac{1}{2}} \underset{\sim}{\varepsilon} \right\|^2 \right\} = \min_{f\in\mathcal{H}} \left\{ \left\| \boldsymbol{\Sigma}^{-\frac{1}{2}} (\underset{\sim}{y} - \underset{\sim}{f}) \right\|^2 \right\}, \tag{9}$$

with constrain:

$$\| f \|^2 < \gamma, \quad \gamma \geq 0. \tag{10}$$

Then, space function $\mathcal{H} = \mathrm{W}_2^m[a,b]$ used is order-2 Sobolev space defined as follow:

$$\mathrm{W}_2^m[a,b] = \left\{ f\colon \int_a^b [f^{(m)}(x_t)]^2 \, dx_t < \infty \right\},$$

Where $a \leq x_t \leq b$ and k = 1,2. Based on the space, nornm of every $f \in \mathrm{W}_2^m[a,b]$ is described as follow:

$$\| f \|^2 = \int_a^b [f^{(m)}(x_t)]^2 dx_t .$$

Optimization with constrain in equation (8) can be stated as:

$$\min_{f\in \mathrm{W}_2^m[a,b]} \left\{ \left\| \boldsymbol{\Sigma}^{-\frac{1}{2}} \underset{\sim}{\varepsilon} \right\|^2 \right\} = \min_{f\in \mathrm{W}_2^m[a,b]} \left\{ \left\| \boldsymbol{\Sigma}^{-\frac{1}{2}} (\underset{\sim}{y} - \underset{\sim}{f}) \right\|^2 \right\}, \tag{11}$$

With constrains in equation (11) into:

$$\int_a^b [f^{(m)}(x_t)]^2 dx_t < \gamma, \quad \gamma \geq 0. \tag{12}$$

Optimization (9) with equivalent constrain (10) by solving Penalized Least Square (PLS) optimization:

$$\min_{f\in \mathrm{W}_2^m[a,b]} \left\{ M^{-1} \left(\underset{\sim}{y} - \underset{\sim}{f} \right)^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} \left(\underset{\sim}{y} - \underset{\sim}{f} \right) + \lambda \int_a^b \left[f^{(m)}\left(x_t\right) \right]^2 dx_t \right\}, \tag{12}$$

where λ_k is smoothing parameter controlling between Goodness of fit:

$$T^{-1}\left(\underset{\sim}{y}-\underset{\sim}{f}\right)^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\left(\underset{\sim}{y}-\underset{\sim}{f}\right),$$

and penalty:

$$\lambda\int_a^b\left[f^{(m)}\left(x_t\right)\right]^2 dx_t\,.$$

To solve optimization in equation (11) with penalty component:

$$\lambda\int_a^b\left[f^{(m)}\left(x_t\right)\right]^2 dx_t=\underset{\sim}{c}^{\mathrm{T}}\lambda\mathbf{I}_T\mathbf{V}\underset{\sim}{c}\,, \tag{13}$$

Using $\underset{\sim}{f}=\mathbf{T}\underset{\sim}{d}+\mathbf{V}\underset{\sim}{c}$ as reference, Goodness of fit in PLS optimization (12) can be stated as:

$$T^{-1}\left(\underset{\sim}{y}-\underset{\sim}{f}\right)^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\left(\underset{\sim}{y}-\underset{\sim}{f}\right)=T^{-1}\left(\underset{\sim}{y}-\mathbf{T}\underset{\sim}{d}-\mathbf{V}\underset{\sim}{c}\right)^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\left(\underset{\sim}{y}-\mathbf{T}\underset{\sim}{d}-\mathbf{V}\underset{\sim}{c}\right). \tag{14}$$

Solving PLS optimization by combining goodness of fit (14) and penalty (13), can be described as:

$$\begin{aligned}
&\min_{f\in W_2^m[a,b]}\left\{M^{-1}\left(\underset{\sim}{y}-\underset{\sim}{f}\right)^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\left(\underset{\sim}{y}-\underset{\sim}{f}\right)+\lambda\int_a^b\left[f^{(m)}\left(x_t\right)\right]^2 dx_t\right\}\\
&=\min_{\substack{\underset{\sim}{c}\in\Re^T\\ \underset{\sim}{d}\in\Re^m}}\left\{T^{-1}\left(\underset{\sim}{y}-\mathbf{T}\underset{\sim}{d}-\mathbf{V}\underset{\sim}{c}\right)^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\left(\underset{\sim}{y}-\mathbf{T}\underset{\sim}{d}-\mathbf{V}\underset{\sim}{c}\right)+\underset{\sim}{c}^{\mathrm{T}}\lambda\mathbf{I}_T\mathbf{V}\underset{\sim}{c}\right\}\\
&=\min_{\substack{\underset{\sim}{c}\in\Re^T\\ \underset{\sim}{d}\in\Re^m}}\left\{\left(\left(\underset{\sim}{y}-\mathbf{T}\underset{\sim}{d}-\mathbf{V}\underset{\sim}{c}\right)^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\left(\underset{\sim}{y}-\mathbf{T}\underset{\sim}{d}-\mathbf{V}\underset{\sim}{c}\right)+\underset{\sim}{c}^{\mathrm{T}}T\lambda\mathbf{I}_T\mathbf{V}\underset{\sim}{c}\right)T^{-1}\right\}\\
&=\min_{\substack{\underset{\sim}{c}\in\Re^T\\ \underset{\sim}{d}\in\Re^m}}\left\{\left[(\underset{\sim}{y}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\underset{\sim}{y}-\underset{\sim}{y}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\mathbf{T}\underset{\sim}{d}-\underset{\sim}{y}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\mathbf{V}\underset{\sim}{c}-\underset{\sim}{d}^{\mathrm{T}}\mathbf{T}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\underset{\sim}{y}+\underset{\sim}{d}^{\mathrm{T}}\mathbf{T}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\mathbf{T}\underset{\sim}{d}+\right.\right.\\
&\qquad\left.\left.+\underset{\sim}{d}^{\mathrm{T}}\mathbf{T}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\mathbf{V}\underset{\sim}{c}-\underset{\sim}{c}^{\mathrm{T}}\mathbf{V}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\underset{\sim}{y}+\underset{\sim}{c}^{\mathrm{T}}\mathbf{V}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\mathbf{T}\underset{\sim}{d}+\underset{\sim}{c}^{\mathrm{T}}\mathbf{V}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\mathbf{V}\underset{\sim}{c}+\underset{\sim}{c}^{\mathrm{T}}T\lambda\mathbf{I}_T\mathbf{V}\underset{\sim}{c}\right]T^{-1}\right\}\\
&=\min_{\substack{\underset{\sim}{c}\in\Re^T\\ \underset{\sim}{d}\in\Re^m}}\left\{\left[(\underset{\sim}{y}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\underset{\sim}{y}-2\underset{\sim}{d}^{\mathrm{T}}\mathbf{T}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\underset{\sim}{y}-2\underset{\sim}{c}^{\mathrm{T}}\mathbf{V}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\underset{\sim}{y}+\underset{\sim}{d}^{\mathrm{T}}\mathbf{T}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\mathbf{T}\underset{\sim}{d}+\underset{\sim}{d}^{\mathrm{T}}\mathbf{T}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\mathbf{V}\underset{\sim}{c}+\right.\right.\\
&\qquad\left.\left.+\underset{\sim}{c}^{\mathrm{T}}\mathbf{V}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\mathbf{T}\underset{\sim}{d}+\underset{\sim}{c}^{\mathrm{T}}\left(\mathbf{V}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\mathbf{V}+T\lambda\mathbf{I}_T\mathbf{V}\right)\underset{\sim}{c}\right]T^{-1}\right\}\\
&=\min_{\substack{\underset{\sim}{c}\in\Re^T\\ \underset{\sim}{d}\in\Re^m}}\left\{\ \mathrm{Q}(\underset{\sim}{c},\underset{\sim}{d})\right\}.
\end{aligned} \tag{15}$$

Solving optimization (15) is obtained by conducting partial derivative $\mathrm{Q}(\underset{\sim}{c},\underset{\sim}{d})$ gradually towards $\underset{\sim}{c}$ and $\underset{\sim}{d}$, then the result equals to zero. The partial derivative is presented as follow:

$$\frac{\partial\,\mathrm{Q}(\underset{\sim}{c},\underset{\sim}{d})}{\partial \underset{\sim}{c}}=\underset{\sim}{0},$$

and the result is:

$$-2\mathbf{V}^{\mathrm{T}}\mathbf{\Sigma}^{-1}\underset{\sim}{y}+2\mathbf{V}^{\mathrm{T}}\mathbf{\Sigma}^{-1}\mathbf{T}\underset{\sim}{d}+2\left(\mathbf{V}^{\mathrm{T}}\mathbf{\Sigma}^{-1}\mathbf{V}+T\lambda\mathbf{I}_{T}\mathbf{V}\right)\hat{\underset{\sim}{c}}=\underset{\sim}{0}.$$

$$\mathbf{V}^{\mathrm{T}}\left\{-\mathbf{\Sigma}^{-1}\underset{\sim}{y}+\mathbf{\Sigma}^{-1}\mathbf{T}\underset{\sim}{d}+[\mathbf{\Sigma}^{-1}\mathbf{V}+T\lambda\mathbf{I}_{T}\mathbf{I}]\hat{\underset{\sim}{c}}\right\}=\underset{\sim}{0}.$$

$$-\mathbf{\Sigma}^{-1}\underset{\sim}{y}+\mathbf{\Sigma}^{-1}\mathbf{T}\underset{\sim}{d}+[\mathbf{\Sigma}^{-1}\mathbf{V}+T\lambda\mathbf{I}_{T}\mathbf{I}]\hat{\underset{\sim}{c}}=\underset{\sim}{0}. \qquad (16)$$

When matrix **U** is presented as:

$$\mathbf{U}=\mathbf{\Sigma}^{-1}\mathbf{V}+T\lambda\mathbf{I}_{T}.$$

equation (16) can be stated as:

$$-\mathbf{\Sigma}^{-1}\underset{\sim}{y}+\mathbf{\Sigma}^{-1}\mathbf{T}\underset{\sim}{d}+\mathbf{U}\hat{\underset{\sim}{c}}=\underset{\sim}{0}.$$

$$\mathbf{U}\hat{\underset{\sim}{c}}=\mathbf{\Sigma}^{-1}(\underset{\sim}{y}-\mathbf{T}\underset{\sim}{d}) \qquad (17)$$

Equation (17) is doubled from the left with $\mathbf{U}^{-1}$ and the following equation is ontained:

$$\hat{\underset{\sim}{c}}=\mathbf{U}^{-1}\mathbf{\Sigma}^{-1}(\underset{\sim}{y}-\mathbf{T}\underset{\sim}{d}) \qquad (18)$$

Furthermore, partial derivative:

$$\frac{\partial\,\mathrm{Q}(\underset{\sim}{c},\underset{\sim}{d})}{\partial \underset{\sim}{d}}=\underset{\sim}{0},$$

results in:

$$-\mathbf{T}^{\mathrm{T}}\mathbf{\Sigma}^{-1}\underset{\sim}{y}+\mathbf{T}^{\mathrm{T}}\mathbf{\Sigma}^{-1}\mathbf{T}\hat{\underset{\sim}{d}}+\mathbf{T}^{\mathrm{T}}\mathbf{\Sigma}^{-1}\mathbf{V}\hat{\underset{\sim}{c}}=\underset{\sim}{0}$$

Elaboration of eqaution (17) results in the following equations:

$$-\mathbf{T}^{\mathrm{T}}\mathbf{\Sigma}^{-1}\underset{\sim}{y}+\mathbf{T}^{\mathrm{T}}\mathbf{\Sigma}^{-1}\mathbf{T}\hat{\underset{\sim}{d}}+\mathbf{T}^{\mathrm{T}}\mathbf{\Sigma}^{-1}\mathbf{V}\{\mathbf{U}^{-1}\mathbf{\Sigma}^{-1}(\underset{\sim}{y}-\mathbf{T}\hat{\underset{\sim}{d}})\}=\underset{\sim}{0}$$

$$-\mathbf{T}^{\mathrm{T}}\mathbf{\Sigma}^{-1}\underset{\sim}{y}+\mathbf{T}^{\mathrm{T}}\mathbf{\Sigma}^{-1}\mathbf{T}\hat{\underset{\sim}{d}}+\mathbf{T}^{\mathrm{T}}[\mathbf{\Sigma}^{-1}\mathbf{V}\mathbf{U}^{-1}]\mathbf{\Sigma}^{-1}(\underset{\sim}{y}-\mathbf{T}\hat{\underset{\sim}{d}})=\underset{\sim}{0}. \qquad (19)$$

Considering $\mathbf{U}=\mathbf{\Sigma}^{-1}\mathbf{V}+T\lambda\mathbf{I}_{T}\mathbf{I}$, then $\mathbf{V}=\mathbf{\Sigma}(\mathbf{U}-T\lambda\mathbf{I}_{T}\mathbf{I})$, as the consequence, the result is the foillowing equations:

$$\mathbf{V}\mathbf{U}^{-1}=\mathbf{\Sigma}(\mathbf{U}-T\lambda\mathbf{I}_{T}\mathbf{I})\mathbf{U}^{-1}$$

$$\mathbf{V}\mathbf{U}^{-1} = \mathbf{\Sigma}(\mathbf{I} - T\lambda\mathbf{I}_T\mathbf{U}^{-1}).$$

Reduplicating the equation above with $\mathbf{\Sigma}^{-1}$resulting in:

$$\mathbf{\Sigma}^{-1}\mathbf{V}\mathbf{U}^{-1} = \mathbf{I} - T\lambda\mathbf{I}_T\mathbf{U}^{-1}.$$

The equation is substituted in equation (19) resulting in:

$$-\mathbf{T}^{\mathrm{T}}\mathbf{\Sigma}^{-1}\underset{\sim}{y} + \mathbf{T}^{\mathrm{T}}\mathbf{\Sigma}^{-1}\mathbf{T}\hat{\underset{\sim}{d}} + \mathbf{T}^{\mathrm{T}}[\mathbf{I} - T\lambda\mathbf{I}_T\mathbf{U}^{-1}]\mathbf{\Sigma}^{-1}(\underset{\sim}{y} - \mathbf{T}\hat{\underset{\sim}{d}}) = \underset{\sim}{0}$$

When the equation above is elaborated further, the result is:

$$-T\lambda\mathbf{I}_T\mathbf{T}^{\mathrm{T}}\mathbf{U}^{-1}\mathbf{\Sigma}^{-1}\underset{\sim}{y} + T\lambda\mathbf{I}_T\mathbf{T}^{\mathrm{T}}\mathbf{U}^{-1}\mathbf{\Sigma}^{-1}\mathbf{T}\hat{\underset{\sim}{d}} = \underset{\sim}{0}.$$

$$T\lambda\mathbf{I}_T\mathbf{T}^{\mathrm{T}}\mathbf{U}^{-1}\mathbf{\Sigma}^{-1}\mathbf{T}\hat{\underset{\sim}{d}} = T\lambda\mathbf{I}_T\mathbf{T}^{\mathrm{T}}\mathbf{U}^{-1}\mathbf{\Sigma}^{-1}\underset{\sim}{y}.$$

Both segments of the equation are reduplicated with $(T\lambda\mathbf{I}_T)^{-1}$ and then simplified resulting in:

$$\hat{\underset{\sim}{d}} = \left(\mathbf{T}^{\mathrm{T}}\mathbf{U}^{-1}\mathbf{\Sigma}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^{\mathrm{T}}\mathbf{U}^{-1}\mathbf{\Sigma}^{-1}\underset{\sim}{y}. \tag{20}$$

Equation (18) is substituted into equation (20) resulting in:

$$\begin{aligned}\hat{\underset{\sim}{c}} &= \mathbf{U}^{-1}\mathbf{\Sigma}^{-1}(\underset{\sim}{y} - \mathbf{T}[\left(\mathbf{T}^{\mathrm{T}}\mathbf{U}^{-1}\mathbf{\Sigma}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^{\mathrm{T}}\mathbf{U}^{-1}\mathbf{\Sigma}^{-1}\underset{\sim}{y}]) \\ &= \mathbf{U}^{-1}\mathbf{\Sigma}^{-1}[\mathbf{I} - \mathbf{T}\left(\mathbf{T}^{\mathrm{T}}\mathbf{U}^{-1}\mathbf{\Sigma}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^{\mathrm{T}}\mathbf{U}^{-1}\mathbf{\Sigma}^{-1}]\underset{\sim}{y}.\end{aligned} \tag{21}$$

Based on equation (20) and (21), estimator for nonparametric regressioncurve is as follow:

$$\begin{aligned}\hat{\underset{\sim}{f}}_{\lambda} &= \mathbf{T}\hat{\underset{\sim}{d}} + \mathbf{V}\hat{\underset{\sim}{c}} \\ \hat{\underset{\sim}{f}}_{\lambda} &= \mathbf{T}\left(\mathbf{T}^{\mathrm{T}}\mathbf{U}^{-1}\mathbf{\Sigma}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^{\mathrm{T}}\mathbf{U}^{-1}\mathbf{\Sigma}^{-1}\underset{\sim}{y} + \mathbf{V}\mathbf{U}^{-1}\mathbf{\Sigma}^{-1}[\mathbf{I} - \mathbf{T}\left(\mathbf{T}^{\mathrm{T}}\mathbf{U}^{-1}\mathbf{\Sigma}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^{\mathrm{T}}\mathbf{U}^{-1}\mathbf{\Sigma}^{-1}]\underset{\sim}{y} \\ \hat{\underset{\sim}{f}}_{\lambda} &= \{\mathbf{T}\left(\mathbf{T}^{\mathrm{T}}\mathbf{U}^{-1}\mathbf{\Sigma}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^{\mathrm{T}}\mathbf{U}^{-1}\mathbf{\Sigma}^{-1} + \mathbf{V}\mathbf{U}^{-1}\mathbf{\Sigma}^{-1}[\mathbf{I} - \mathbf{T}\left(\mathbf{T}^{\mathrm{T}}\mathbf{U}^{-1}\mathbf{\Sigma}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^{\mathrm{T}}\mathbf{U}^{-1}\mathbf{\Sigma}^{-1}]\}\underset{\sim}{y} \\ \hat{\underset{\sim}{f}}_{\lambda} &= \mathbf{A}_{\lambda}\underset{\sim}{y},\end{aligned} \tag{22}$$

where

$$\mathbf{A}_{\lambda} = \mathbf{T}\left(\mathbf{T}^{\mathrm{T}}\mathbf{U}^{-1}\mathbf{\Sigma}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^{\mathrm{T}}\mathbf{U}^{-1}\mathbf{\Sigma}^{-1} + \mathbf{V}\mathbf{U}^{-1}\mathbf{\Sigma}^{-1}[\mathbf{I} - \mathbf{T}\left(\mathbf{T}^{\mathrm{T}}\mathbf{U}^{-1}\mathbf{\Sigma}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^{\mathrm{T}}\mathbf{U}^{-1}\mathbf{\Sigma}^{-1}].$$

Error variance-covariance matrix $\hat{\mathbf{\Sigma}}$ described in [7], so Theorem 1 uses $\hat{\mathbf{\Sigma}}$ as well as $\hat{\mathbf{U}} = \hat{\mathbf{\Sigma}}^{-1}\mathbf{V} + T\lambda\mathbf{I}_T$. resulting in:

$$\hat{\underset{\sim}{d}} = \left(\mathbf{T}^{\mathrm{T}}\hat{\mathbf{U}}^{-1}\hat{\mathbf{\Sigma}}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^{\mathrm{T}}\hat{\mathbf{U}}^{-1}\hat{\mathbf{\Sigma}}^{-1}\underset{\sim}{y}.$$

$$\hat{\underset{\sim}{c}} = \hat{\mathbf{U}}^{-1}\hat{\mathbf{\Sigma}}^{-1}[\mathbf{I}-\mathbf{T}\left(\mathbf{T}^{\mathrm{T}}\hat{\mathbf{U}}^{-1}\hat{\mathbf{\Sigma}}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^{\mathrm{T}}\hat{\mathbf{U}}^{-1}\hat{\mathbf{\Sigma}}^{-1}]\underset{\sim}{y}.$$

$\hat{\underset{\sim}{f}}_{\underset{\sim}{\lambda}} = \mathbf{A}_{\underset{\sim}{\lambda}}\underset{\sim}{y}$ where

$$\mathbf{A}_{\underset{\sim}{\lambda}} = \mathbf{T}\left(\mathbf{T}^{\mathrm{T}}\hat{\mathbf{U}}^{-1}\hat{\mathbf{\Sigma}}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^{\mathrm{T}}\hat{\mathbf{U}}^{-1}\hat{\mathbf{\Sigma}}^{-1} + \mathbf{V}\hat{\mathbf{U}}^{-1}\hat{\mathbf{\Sigma}}^{-1}[\mathbf{I}-\mathbf{T}\left(\mathbf{T}^{\mathrm{T}}\hat{\mathbf{U}}^{-1}\hat{\mathbf{\Sigma}}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^{\mathrm{T}}\hat{\mathbf{U}}^{-1}\hat{\mathbf{\Sigma}}^{-1}].$$

Based on the the theorem above, the equation $\underset{\sim}{f} = \mathbf{T}\underset{\sim}{d} + \mathbf{V}\underset{\sim}{c}$ with spline function estimation for heteroscedasticity case is $\hat{\underset{\sim}{f}}_{\underset{\sim}{\lambda}} = \mathbf{A}_{\underset{\sim}{\lambda}}\underset{\sim}{y}$ dengan

$$\mathbf{A}_{\underset{\sim}{\lambda}} = \mathbf{T}\left(\mathbf{T}^{\mathrm{T}}\hat{\mathbf{U}}^{-1}\hat{\mathbf{\Sigma}}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^{\mathrm{T}}\hat{\mathbf{U}}^{-1}\hat{\mathbf{\Sigma}}^{-1} + \mathbf{V}\hat{\mathbf{U}}^{-1}\hat{\mathbf{\Sigma}}^{-1}[\mathbf{I}-\mathbf{T}\left(\mathbf{T}^{\mathrm{T}}\hat{\mathbf{U}}^{-1}\hat{\mathbf{\Sigma}}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^{\mathrm{T}}\hat{\mathbf{U}}^{-1}\hat{\mathbf{\Sigma}}^{-1}],$$

And one for homoskedasticityis $\hat{\underset{\sim}{f}}_{\underset{\sim}{\lambda}} = \mathbf{A}_{\underset{\sim}{\lambda}}\underset{\sim}{y}$ where

$$\mathbf{A}_{\underset{\sim}{\lambda}} = \mathbf{T}\left(\mathbf{T}^{\mathrm{T}}\hat{\mathbf{U}}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^{\mathrm{T}}\hat{\mathbf{U}}^{-1} + \mathbf{V}\hat{\mathbf{U}}^{-1}[\mathbf{I}-\mathbf{T}\left(\mathbf{T}^{\mathrm{T}}\hat{\mathbf{U}}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^{\mathrm{T}}\hat{\mathbf{U}}^{-1}]. \quad (23)$$

3.2. Estimation of Spline Function in Semiparametric Regression without Penalty

The second goal of the study is to obtain spline function estimation in semiparametric regression without penalty (using Ordinary Least Square or OLS). It is presented in Theorem 2 as follow:

Theorem 2: Ordinary Least Square OLS

When given a pair of data following nonparametric component in semiparametric regression model that meets the shape of nonparametric component in semiparametric regression function described in equation (4), with assumptions $\mathrm{E}(\underset{\sim}{\varepsilon}) = \underset{\sim}{0}$ and $\mathrm{Var}(\underset{\sim}{\varepsilon}) = \mathbf{\Sigma}$, then spline estimator minimizing PLS (when homoskedasticity is fulfilled then $\mathbf{\Sigma} = \mathbf{I}$)

$$\min_{f \in W_2^m[a,b], i=1,2}\left\{M^{-1}\left(\underset{\sim}{y} - \underset{\sim}{f}\right)^{\mathrm{T}}\mathbf{\Sigma}^{-1}\left(\underset{\sim}{y} - \underset{\sim}{f}\right)\right\} \quad (24)$$

The spline estimator is $\hat{\underset{\sim}{f}} = \mathbf{A}y$, where:

$$\mathbf{A} = \mathbf{T}\left(\mathbf{T}^{\mathrm{T}}\hat{\mathbf{\Sigma}}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^{\mathrm{T}}\hat{\mathbf{\Sigma}}^{-1}.$$

Proof: Considering the equation that is function $\underset{\sim}{f} = \mathbf{T}\underset{\sim}{d}$ by Fernandes [17], then the nonparametric component in semiparametric regression model (1) can be stated as:

$$\underset{\sim}{y} = \underset{\sim}{f} + \underset{\sim}{\varepsilon} = \mathbf{T}\underset{\sim}{d} + \underset{\sim}{\varepsilon}.$$

Nonparametric component in semiparametric regression analysis is conducted to get estimator of regression curve $\underset{\sim}{f}$. To get the estimation, Reproducing Kernel Hilbert Space RKHS is used. The purpose is to obtain estimation of $\underset{\sim}{f}$ that meets PLS optimization:

$$\min_{f\in\mathcal{H}}\left\{\left\|\boldsymbol{\Sigma}^{-\frac{1}{2}}\underset{\sim}{\varepsilon}\right\|^2\right\}=\min_{f\in\mathcal{H}}\left\{\left\|\boldsymbol{\Sigma}^{-\frac{1}{2}}(\underset{\sim}{y}-\underset{\sim}{f})\right\|^2\right\}, \tag{25}$$

Or

$$T^{-1}\left(\underset{\sim}{y}-\underset{\sim}{f}\right)^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\left(\underset{\sim}{y}-\underset{\sim}{f}\right)=T^{-1}\left(\underset{\sim}{y}-\mathbf{T}\underset{\sim}{d}\right)^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\left(\underset{\sim}{y}-\mathbf{T}\underset{\sim}{d}\right). \tag{26}$$

Solving PLS optimization by goodness of fit (25) without penalty, can be described as:

$$\begin{aligned}
&\min_{\underset{\sim}{d}\in\Re^{2m}}\left\{T^{-1}\left(\underset{\sim}{y}-\mathbf{T}\underset{\sim}{d}\right)^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\left(\underset{\sim}{y}-\mathbf{T}\underset{\sim}{d}\right)\right\}\\
&=\min_{\underset{\sim}{d}\in\Re^{m}}\left\{\left(\left(\underset{\sim}{y}-\mathbf{T}\underset{\sim}{d}\right)^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\left(\underset{\sim}{y}-\mathbf{T}\underset{\sim}{d}\right)\right)\mathbf{T}^{-1}\right\}\\
&=\min_{\underset{\sim}{d}\in\Re^{m}}\left\{\left[(\underset{\sim}{y}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\underset{\sim}{y}-\underset{\sim}{y}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\mathbf{T}\underset{\sim}{d}-\underset{\sim}{d}^{\mathrm{T}}\mathbf{T}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\underset{\sim}{y}+\underset{\sim}{d}^{\mathrm{T}}\mathbf{T}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\mathbf{T}\underset{\sim}{d}\right]\mathbf{T}^{-1}\right\}\\
&=\min_{\underset{\sim}{d}\in\Re^{m}}\left\{\left[(\underset{\sim}{y}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\underset{\sim}{y}-2\underset{\sim}{d}^{\mathrm{T}}\mathbf{T}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\underset{\sim}{y}+\underset{\sim}{d}^{\mathrm{T}}\mathbf{T}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\mathbf{T}\underset{\sim}{d}\right]\mathbf{T}^{-1}\right\}\\
&=\min_{\underset{\sim}{d}\in\Re^{m}}\left\{\mathrm{Q}(\underset{\sim}{d})\right\}.
\end{aligned} \tag{27}$$

Solving optimization (27) is obtained by conducting derivative $\mathrm{Q}(\underset{\sim}{d})$ by $\underset{\sim}{d}$, then the result equals to zero. The partial derivative is presented as follow:

$$\frac{\partial\,\mathrm{Q}(\underset{\sim}{d})}{\partial\underset{\sim}{d}}=\underset{\sim}{0},$$

and the result is:

$$-\mathbf{T}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\underset{\sim}{y}+\mathbf{T}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\mathbf{T}\hat{\underset{\sim}{d}}=\underset{\sim}{0}$$

$$\hat{\underset{\sim}{d}}=\left(\mathbf{T}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\underset{\sim}{y}. \tag{28}$$

Estimator for nonparametric component in semiparametric regressioncurve is as follow:

$$\hat{\underset{\sim}{f}}=\mathbf{T}\hat{\underset{\sim}{d}}$$

$$\hat{\underset{\sim}{f}}=\mathbf{T}\left(\mathbf{T}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\underset{\sim}{y}$$

$$\hat{\underset{\sim}{f}}=\{\mathbf{T}\left(\mathbf{T}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\}\underset{\sim}{y}$$

$$\hat{\underset{\sim}{f}}=\mathbf{A}\underset{\sim}{y}, \tag{29}$$

where

$$\mathbf{A} = \mathbf{T}\left(\mathbf{T}^{\mathrm{T}}\hat{\boldsymbol{\Sigma}}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^{\mathrm{T}}\hat{\boldsymbol{\Sigma}}^{-1}. \quad (30)$$

3.3. Properties of Semiparametric Regression Using Smoothing Spline

The third goal of the study is to obtain the properties of Semiparametric Regression Using Smoothing Spline. A semiparametric regression analysis (8) generates the equation in matrix form as follows:

$$\underset{\sim}{y} = \mathbf{x}\underset{\sim}{\beta} + \underset{\sim}{f} + \underset{\sim}{\varepsilon} \quad (31)$$

by replacing $\underset{\sim}{\beta}$ dan $\underset{\sim}{f}$ with each estimation, it is obtained

$$\hat{\underset{\sim}{y}} = \mathbf{x}\hat{\underset{\sim}{\beta}} + \hat{\underset{\sim}{f}} + \underset{\sim}{\varepsilon}$$

if $\underset{\sim}{z} = \underset{\sim}{y} - \mathbf{x}\underset{\sim}{\beta}$ hence $\hat{\underset{\sim}{f}}$ has the equation

$$\hat{\underset{\sim}{f}}_{\underset{\sim}{\lambda}} = \mathbf{A}_{\underset{\sim}{\lambda}}\underset{\sim}{z} \quad (32)$$

With the modification of the equation (22) to equation (23) thus the estimation of regression curve with a penalty (PLS) obtained is

$$\mathbf{A}_{\underset{\sim}{\lambda}} = \mathbf{T}\left(\mathbf{T}^{\mathrm{T}}\hat{\mathbf{U}}^{-1}\hat{\boldsymbol{\Sigma}}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^{\mathrm{T}}\hat{\mathbf{U}}^{-1}\hat{\boldsymbol{\Sigma}}^{-1}\,\underset{\sim}{z} + \mathbf{V}\hat{\mathbf{U}}^{-1}\hat{\boldsymbol{\Sigma}}^{-1}[\mathbf{I} - \mathbf{T}\left(\mathbf{T}^{\mathrm{T}}\hat{\mathbf{U}}^{-1}\hat{\boldsymbol{\Sigma}}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^{\mathrm{T}}\hat{\mathbf{U}}^{-1}\hat{\boldsymbol{\Sigma}}^{-1}]\underset{\sim}{z} \quad (33)$$

As well as the estimatipn of regression curves without penalty (OLS)

$$\mathbf{A} = \mathbf{T}\left(\mathbf{T}^{\mathrm{T}}\hat{\boldsymbol{\Sigma}}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^{\mathrm{T}}\hat{\boldsymbol{\Sigma}}^{-1}\underset{\sim}{z} \quad (33)$$

After obtaining the regression curve estimator, the following section will explain some linear and bias characteristics of the estimator of nonparametric regression curve $\hat{\underset{\sim}{f}}_{\underset{\sim}{\lambda}}$.

a) The estimator of nonparametric curve $\hat{\underset{\sim}{f}}_{\underset{\sim}{\lambda}}$ regression is linear. Seen from Equation (33) that estimator of nonparametric regression curve $\hat{\underset{\sim}{f}}_{\underset{\sim}{\lambda}}$, is a linear estimator in the observations.

b) The estimator of nonparametric curve $\hat{\underset{\sim}{f}}_{\underset{\sim}{\lambda}}$ regression is biased. Although the estimator of nonparametric curve $\hat{\underset{\sim}{f}}_{\underset{\sim}{\lambda}}$ regression is a linear estimator in the observations, but this is a biased estimator for the curve $\underset{\sim}{f}$ regression This statement can be proved as follows:

   1) The equation (22) generates:

   2) $\hat{\underset{\sim}{f}}_{\underset{\sim}{\lambda}} = \mathbf{A}_{\underset{\sim}{\lambda}}\underset{\sim}{y}$,

   3) The equation above is taken the expectations, an equation obtained is:

4) $E[\hat{f}_{\lambda}] = E\left[\hat{f}_{1,\lambda_1}(x_1)\right]$

$E[\hat{f}_{\lambda}]$5) $= E\left[\mathbf{A}_{\lambda} y\right] = E\left[\mathbf{A}_{\lambda}(y)\right]$

$E[\hat{f}_{\lambda}] = \mathbf{A}_{\lambda}\left[E(y)\right] = \mathbf{A}_{\lambda}\left[f(x)\right]$. 6)

Seeing as $\mathbf{A}_{\lambda} \neq \mathbf{I}$, hence

$$E[\hat{f}_{\lambda}] \neq \left[f(x)\right]$$

3.4. The efficiency of a estimator spline curve in semiparametric regression model on the data application

The fourth objective of this research was to compare the results of smoothing spline regression analysis with and without the penalty that is applied to the data of the rate of poverty in East Java. Research Data used is secondary data obtained from Badan Pusat Statistik the National Statistical Bureau of East Java Province. The taken data are the rate of poverty (y), Labor Force Participation Rate LFPR and the Human Development Index HDI as predictors in 38 cities / districts in East Java in 2013. To separate between parametric and nonparametric predictors, Harvey courier test was used. The results of Harvey courier showed the poverty which has more appropriate p-value > ± when processed with parametric regression analysis, in contrast to the HDI with p-value < ± more appropriately processed by nonparametric regression analysis. Thus, it can be concluded that the poverty rate is a paremetrik component predictor (x) and IPM is a nonparametric component predictor (t). Figure 1 shows the Poverty rate (y) with LFPR (t) along with the value of $\hat{y}$ with and without smoothing parameters (red line with penalty, green line without penalty):

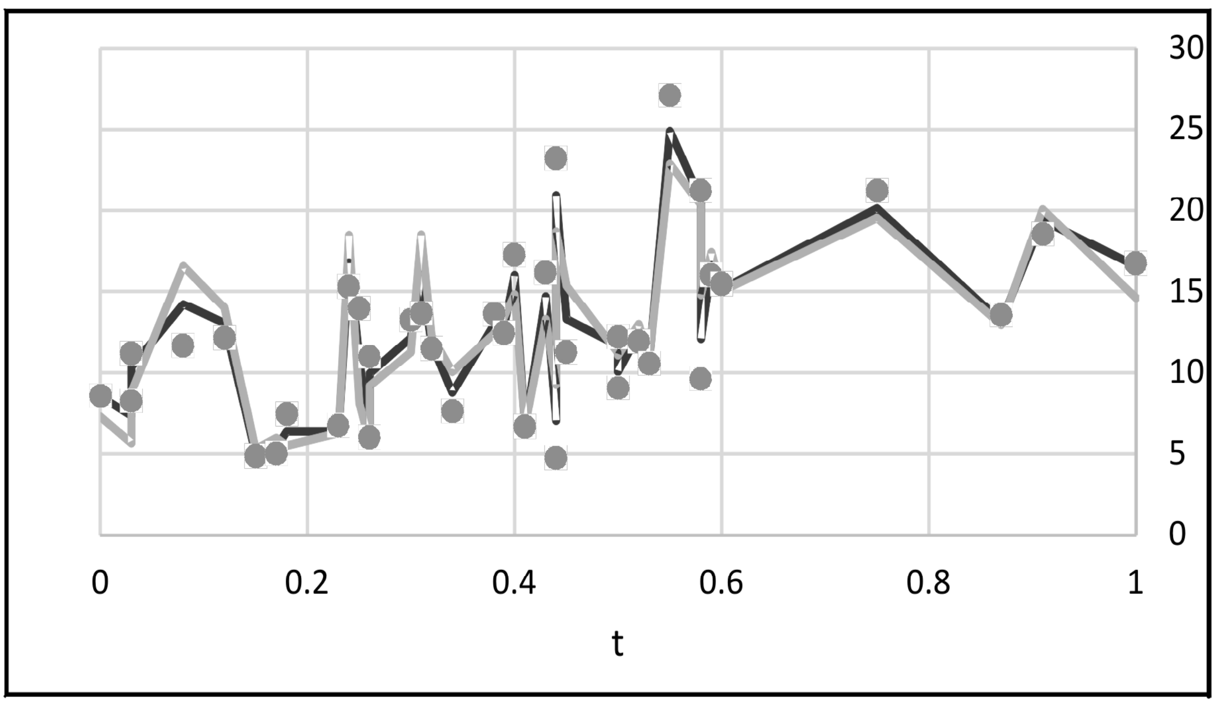

Figure 1: Estimation Curve of Semiparametric Regression

From Figure 1 (redline indicate with penalty, green line indicate without penality), we can see that the probe value for the respon variable ($\hat{y}$) produced is not so different, but it appears that the value of ($\hat{y}$) generated from the semiparametric equation that uses smoothing parameter has a smaller error (closer to the actual value), as at the first point when *t* is 0 and the point of *t* is 1. Also, the result of the R program by disclosing that the semiparametric model uses a penalty with MSE value of 1.7019 and R^2 *93.5584%* better than semiparametric model without using a penalty that has a value of MSE 6.9889 and R^2 *73.5483%*.

The probe is said to be efficient if the probe has the minimum error of variance. The efficiency of the estimator is the ratio of the minimum error variance of estimators. Meanwhile, the relative efficiency is the ratio of error variance of both compared estimators. For instance, $\hat{g}_{DP}(x)$ and $\hat{g}_{TP}(x)$ are two estimators of smoothing spline nonparametric regression of $g(\theta)$. If the estimator follows the general condition of the Cramer-Rao hence the Relative Efficiency (ER) from and $\hat{g}_{DP}(x)$ as the ratio of the variance error is defined as follows.

$$ER_\theta(\hat{g}_{PLS},\hat{g}_{LS}) = \frac{MSE_{PLS}(\theta)}{MSE_{LS}(\theta)}$$

If $ER_\theta(\hat{g}_{PLS},\hat{g}_{LS}) < 1$ hence $\hat{g}_{PLS}(x)$ is more efficient than $\hat{g}_{LS}(x)$

$$ER_\theta(\hat{g}_{PLS},\hat{g}_{LS}) = \frac{MSE_{PLS}(\theta)}{MSE_{LS}(\theta)}$$

$$= \frac{1.7019}{6.9889}$$

$$=0.2435$$

The efficient of semiparametric regression curve without penalty (OLS) amounted to 24.35% of semiparametric regression curve with a penalty (PLS). In other words, smoothing spline semiparametric regression model with a more efficient with penalty than without penalty. In its application (see in Figure 1), it appears that the curve generated by the penalty is more smooth and easier to capture the data patterns compared without penalty.

4. CONCLUSION AND RECCOMENDATION

Based on the analysis result and discussion, the conclusion of this research show: (1) The estimation of spline function estimation in semiparametric regression with penalty (using Penalized Least Square or PLS) is $\hat{\underset{\sim}{f}}_\lambda = \mathbf{A}_\lambda \underset{\sim}{y}$, with

$$\mathbf{A}_\lambda = \mathbf{T}\left(\mathbf{T}^{\mathrm{T}}\hat{\mathbf{U}}^{-1}\hat{\mathbf{\Sigma}}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^{\mathrm{T}}\hat{\mathbf{U}}^{-1}\hat{\mathbf{\Sigma}}^{-1} + \mathbf{V}\hat{\mathbf{U}}^{-1}\hat{\mathbf{\Sigma}}^{-1}[\mathbf{I} - \mathbf{T}\left(\mathbf{T}^{\mathrm{T}}\hat{\mathbf{U}}^{-1}\hat{\mathbf{\Sigma}}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^{\mathrm{T}}\hat{\mathbf{U}}^{-1}\hat{\mathbf{\Sigma}}^{-1}].$$

(2) The estimation of spline function estimation in semiparametric regression without penalty (using Ordinary Least Square or OLS) is $\hat{\underset{\sim}{f}} = \mathbf{A}y$, with

$$\mathbf{A} = \mathbf{T}\left(\mathbf{T}^{\mathrm{T}}\hat{\mathbf{\Sigma}}^{-1}\mathbf{T}\right)^{-1}\mathbf{T}^{\mathrm{T}}\hat{\mathbf{\Sigma}}^{-1}.$$

(3) The properties of Semiparametric Regression Using Smoothing Spline is linear and bias characteristic of the estimator of nonparametric regression curve. (4) The comparing of smoothing spline regression analysis with and without the penalty that is applied to the data of the rate of poverty in East Java. Research Data used is secondary data obtained from Badan Pusat Statistik the National Statistical Bureau of East Java Province, with the rate of poverty (y), Labor Force Participation Rate LFPR and the Human Development Index HDI as predictors in 38 cities / districts in East Java in 2013, The efficient of semiparametric regression curve without penalty (OLS) amounted to 24.35% of semiparametric regression curve with a penalty (PLS). In other words, smoothing spline semiparametric regression model with a more efficient with penalty than without penalty.

Based on the conclusion that has been described, it can be suggested in the application of the smoothing spline semiparametric regression, suggestions which can be given for future research are to compare the effectiveness of the method of smoothing spline with the truncated spline on the same case and data, and do the same semiparametric regression analysis method application in another more difficult case, as in the multirespond case or longitudinal data.

REFERENCES

[1] Eubank, R.L. (1999), *Nonparametric Regression and Spline Smoothing*. Second Edition. New York: Marcel Dekker, Inc.

[2] Takezawa, K.. (2006), *Introduction to Nonparametric Regression*, John Wiley and Sons, Inc, USA.

[3] Gujarati, D. (2011), *Econometrics By Example*, Palgrave Macmillan, Britania.

[4] Kutner, M.H., Nachtsheim, C.J., Neter, J.& Li W. (2005), *Applied Linear Statistical Models*, Fifth Edition, McGraw-Hill International, Boston.

[5] Fissler, J.A., (1994), PWLS Image Reconstruction for Positron Emission Tomography, *IEEE Transactions on Medical Imaging* Vol.13(2).

[6] Fernandes, A.A.R, Budiantara, I.N, Otok, B.W., and Suhartono. (2014), Reproducing Kernel Hilbert Space and Penalized Weighted Least Square in Nonparametric Regression, *Applied Mathematical Science* Vol 8, 2014, No 146, pp. 7289-7300. *http://dx.doi.org/10.12988/ams.2014.49759.*

[7] Fernandes, A.A.R, Budiantara, I.N, Otok, B.W., and Suhartono. (2015), Spline Estimator for Bi-Responses and Multi-Predictors Nonparametric Regression Model in Case of Longitudinal Data, *Journal of Mathematics and Statistics*, Vol 11, No 2, 2015, pp. 61-69. *http://dx.doi.org/10.3844/jmssp.2015.61.69.*

[8] Brandimarte, P. (2011), *Quantitative Methods An Introduction for Business Management*, John Wiley and Sons, Inc, New Jersey.

[9] Budiantara, I.N. (2000), Method of U, GML, CV and GCV in Nonparametric Regression Spline (In Indonesian Metode U, GML, CV dan GCV Dalam Regresi Nonparametric Spline). *Majalah Ilmiah Himpunan Matematika Indonesia (MIHMI)*, 6, p.285-290.

[10] Huang, Z. (2003), Local Asymptotics for Polynomial Spline Regression. *The Annals of Statistics*, **31**(5), 1600-1635.

[11] Crainiceanu, C.M., Ruppert, D., and Wand, M.P. (2004), Bayesian Analysis for Penalized Spline Regression Using WinBugs. *Journal of Statistical Software*, **14**(14), 6-14.

[12] Kim, Y.J., and Gu, C. (2004), Smoothing Spline Gaussian Regression: More Scalable Computation via Efficient Approximation. *Royal Statistical Society: Series B*, **66**(2), 337-356.

[13] Lee, T.C.M. (2004), Improved Smoothing Spline Regression by Combining Estimates of Different Smoothness. *Statistics and Probability Letters*, **67**(1), 133-140.

[14] Howell, J.R. (2007), *Analysis Using Smoothing Splines As Implemented In LME() In R*. Thesis. Brigham Young University.

[15] Wang, Y. (2011), *Smoothing Spline Methods Applications,* CRC Press, New York.

[16] Wahba, G.(1990), Spline *Models for Observational Data*. Pensylvania. SIAM.

[17] Fernandes, A.A.R, Budiantara, I.N, Otok, B.W., and Suhartono. (2014), Reproducing Kernel Hilbert Space for Penalized Regression multi-predictors: case in longitudinal data. *International Journal of Mathemathical Analysis*, Vol. 8, no. 40, 2014, 1951-1961. *http://dx.doi.org/10.12988/ijma.2014.47212.*

[18] Fernandes, A.A.R, Budiantara, I.N, Otok, B.W., *and* Suhartono. 2014. *Spline estimator for bi-responses nonparametric regression model for* longitudinal *data*. **Applied Mathematical Sciences**, Vol. 8, no. 114, 2014, 5653-5665. *http://dx.doi.org/10.12988/ams.2014.47566.*

[19] Badan Pertahanan Nasional (2016). *http://jatim.bpn.go.id/.*

[20] Badan Pusat Statistik (2015). *http://jatim.bps.go.id/.*

[21] Badan Pusat Statistik (2015). *Statistik Mobilitas Penduduk dan Tenaga Kerja 2015*. Badan Pusat Statistik.