Global Journal of Pure and Applied Mathematics.
ISSN 0973-1768 Volume 12, Number 6 (2016), pp. 5047-5057

http://www.ripublication.com/gjpam.htm

Estimation of Nonparametric Regression Curve using Mixed Estimator of Multivariable Truncated Spline and Multivariable Kernel

Vita Ratnasari, I Nyoman Budiantara, Madu Ratna and Ismaini Zain

Department of Statistics, Faculty of Mathematics and Natural Sciences, Institut Teknologi Sepuluh Nopember (ITS), Surabaya, Indonesia.

Abstract

Data given in pairs, $(t_{1i},...,t_{pi},z_{1i},...,z_{qi},y_i)$, i = 1,2,...,n which follows the nonparametric regression model multivariable predictors of additives:

$$y_i = \mu(t_{1i},...,t_{pi},z_{1i},...,z_{qi}) + \varepsilon_i = \sum_{r=1}^{p} g_r(t_{ri}) + \sum_{s=1}^{q} h_s(z_{si}) + \varepsilon_i .$$

The regression curve $g_r(t_r)$, r =1,2,...,p and $h_s(z_s)$, s = 1,2,...,q assumed smooth, and each approached using Spline function truncated and Kernel functions. Nonparametric regression curve estimation multivariable predictor truncated Spline and Kernel mixed obtained from optimization:

$$\underset{g,h}{Min}\left(\left\| y - \sum_{r=1}^{p} g_r(t_{ri}) - \sum_{s=1}^{q} h_s(z_{si}) \right\|^2 \right).$$

Truncated Spline component multivariable estimator, multivariable Kernel component, and a mixture of truncated Spline and Kernel are follow:

$$\sum_{r=1}^{p} \hat{g}_r(t_{ri},K,\alpha), \sum_{s=1}^{q} \hat{h}_s(z_{si},\alpha) \text{ and } \hat{\mu}(t_{1i},...,t_{pi},z_{1i},...,z_{qi},K,\alpha).$$

Truncated Spline components multivariable estimator, Kernel multivariable components, and mix Spline Kernel truncated and each is a biased estimator, but it is a linear estimator class under observation. Spline Estimator mixture truncated and multivariable Kernel is depend on the points of knot K and bandwidth parameters. The mix of Truncted Spline and multivariable Kernel Estimator associated with knot K optimal and optimal bandwidth parameters. Point knot and optimal bandwidth parameters derived from minimum of Generalized Cross Validation (GCV).

Keywords: MultivariableTruncated Spline, Multivariable Kernel, Mixed Estimator**,** Nonparametric Regression

1. INTRODUCTION

Nonparametric regression is one branch of a very flexible regression in estimating regression curve. If the shape of the regression curve is known, it can be used parametric regression approach [1], [2]. In some circumstances, common data pattern does not follow a specific pattern, such as pattern data associated with poverty [3], the production of paddy rice [4], the growth [5], [6], [7], [8], unemployment [9] and mortality [10]. In such conditions it is advisable to use a nonparametric regression approach. There are several methods of nonparametric regression which has been developed by researchers, such as the smoothing Spline [1], [2], [11], [12], [13], [14], [15], Truncated Spline [2], [3], [8], [16], Kernel [17], [18] and the Fourier series [19], [20]. If the pattern of response predictors, most follow certain patterns and some do not follow a specific pattern, it can use the approach of semiparametric regression [1], [4], [6], [7], [9], [15], [16], [20].

The results of the research mentioned above, both associated with the parametric regression, nonparametric regression, and semiparametric regression, all estimates the regression curve by using only one type of shape functions for all patterns of relationships between the multivariable predictors to repon. In some applications, the pattern of the relationship between the response predictors with each other predictors can be different from one another. Therefore, in this condition are advised to use a mixed of approaches estimator for the estimation of the nonparametric regression curve multivariable predictors. Budiantara, et. al. [21] has used a mixture of Spline and Kernel estimator to estimate multivariable predictors nonparametric regression curve. Estimator this mixture is best used if some pattern of response data with predictors do not have a specific pattern, and some of them have a certain pattern, but the change in the sub-intervals. Mixed estimator developed by Budiantara, et. al. [21], is still limited to using just one predictor for each component and component Spline kernel. As a result of this mixture estimator can not be implemented to resolve the issue of regression, if the nonparametric regression multivariable predictors contains more than one for each component. Therefore, in this paper will be given the mixed model of truncated Spline and Kernel nonparametric regression multivariable in which each component is either truncated or component parts Spline kernel contains more than one predictor variable (Multivariable).

2. ESTIMATIONOF MIXED MULTIVARIABLE REGRESSION NONPARAMETRIC CURVE

Given the data pairs $(t_{1i},...,t_{pi},z_{1i},...,z_{qi},y_i)$, i = 1,2,...,n modeled nonparametric regression multivariable predictors:

$$y_i = \mu(t_{1i},...,t_{pi},z_{1i},...,z_{qi}) + \varepsilon_i . \tag{1}$$

Random error assumed to be independent and normally distributed with $E(\varepsilon_i) = 0$ and $E(\varepsilon_i^2) = \sigma^2$. The regression curve is assumed to follow the model of additives:

$$\mu(t_{1i},...,t_{pi},z_{1i},...,z_{qi}) = \sum_{r=1}^{p} g_r(t_{ri}) + \sum_{s=1}^{q} h_s(z_{si}) . \tag{2}$$

Functions $g_r(t_r)$, $r=1,2,...,p$ and $h_s(z_s)$, $s=1,2,...,q$ are assumed smooth, in the sense of continuous and differentiabel. Function $g_r(t_r)$, $r=1,2,...,p$ approached using truncated spline function and the function $h_s(z_s)$, $s=1,2,...,q$ approached with Kernel Nadaraya-Watson. To get a nonparametric regression curve estimation using a mixture Truncated Spline estimator and Kernel multivariable, given some of the following lemma.

Lemma 1

Multivariable nonparametric regression model given in equation (1). If the function, $g_r(t_r)$, $r=1,2,...,p$ approached using Spline function truncated linear with knot $\tilde{k}_r = (K_{r1}, K_{r2}, ..., K_{rm})^T$, then

$$\sum_{r=1}^{p} \tilde{g}_r(t_r) = \mathbf{G}(\tilde{k})\tilde{\theta},$$

with $\tilde{g}_r(t_r) = \left(g_r(t_{r1}) \;\; g_r(t_{r2}) \;\; ... \;\; g_r(t_{rn}) \right)^T$, $\mathbf{G}(\tilde{k}) = \left(\mathbf{G}(k_1) \quad \mathbf{G}(k_2) \quad \cdots \quad \mathbf{G}(k_p) \right)$,

$$\mathbf{G}(k_r) = \begin{pmatrix} t_{r1} & (t_{r1} - K_{r1})_+ & \cdots & (t_{r1} - K_{rm})_+ \\ t_{r2} & (t_{r1} - K_{r1})_+ & \cdots & (t_{r1} - K_{rm})_+ \\ \vdots & \vdots & \vdots & \vdots \\ t_{rn} & (t_{rn} - K_{r1})_+ & \cdots & (t_{rn} - K_{rm})_+ \end{pmatrix},$$

$\tilde{\theta} = \left(\tilde{\theta}_1^T \; \tilde{\theta}_2^T \; ... \; \tilde{\theta}_p^T \right)^T$, and $\tilde{\theta}_r = \left(\theta_{r1} \; \lambda_{r1} \; ... \; \lambda_{rm} \right)$, $r = 1,2,...,p$.

Proof:

Regression curve $g_r(t_{ri})$ approached with truncated Spline linear function with knot $\tilde{k}_r = (K_{r1}, K_{r2}, ..., K_{rm})^T$, then it can be written as:

$$g_r(t_{ri}) = \theta_{r1} x_{ri} + \sum_{j=1}^{m} \lambda_{rj} (x_{ri} - K_{rj})_+ \text{ , } i = 1,2,...,n$$

with $\theta_{r1}, \lambda_{r1}, ..., \lambda_{rm}$ are unknown parameters, and truncated function as follow:

$$(x_{ri} - K_{rj})_+ = \begin{cases} (x_{ri} - K_{rj}), & x_{ri} \geq K_{rj} \\ 0, & x_{ri} < K_{rj} \end{cases}$$

For $i = 1$ model is obtained:

$$\begin{pmatrix} g_1(t_{11}) \\ g_1(t_{12}) \\ \vdots \\ g_1(t_{1n}) \end{pmatrix} = \begin{pmatrix} t_{11} & (t_{11} - K_{11})_+ & \cdots & (t_{11} - K_{1m})_+ \\ t_{12} & (t_{11} - K_{11})_+ & \cdots & (t_{11} - K_{1m})_+ \\ \vdots & \vdots & \cdots & \vdots \\ t_{1n} & (t_{1n} - K_{11})_+ & \cdots & (t_{1n} - K_{1m})_+ \end{pmatrix} \begin{pmatrix} \theta_{11} \\ \lambda_{11} \\ \vdots \\ \lambda_{1m} \end{pmatrix}$$

As a result, obtained the following models:

$$\tilde{g}_1(t_1) = \mathbf{G}(k_1)\tilde{\theta}_1$$

In a similar way to $r = 2$, ..., $r = p$, obtained the model:

$$\tilde{g}_2(t_2) = \mathbf{G}(k_2)\tilde{\theta}_2, \quad ..., \tilde{g}_p(t_p) = \mathbf{G}(k_p)\tilde{\theta}_p .$$

Subsequently obtained:

$$\sum_{r=1}^{p} \tilde{g}_r(t_r) = \mathbf{G}(k_1)\tilde{\theta}_1 + \mathbf{G}(k_2)\tilde{\theta}_2 + ... + \mathbf{G}(k_p)\tilde{\theta}_p$$

$$= \left(\mathbf{G}(k_1) \ \ \mathbf{G}(k_2) \cdots \ \mathbf{G}(k_r)\right)\left(\tilde{\theta}_1^T \ \ \tilde{\theta}_2^T \ \cdots \ \tilde{\theta}_p^T\right)^T = \mathbf{G}(\tilde{k})\tilde{\theta} \tag{3}$$

Lemma 2

Multivariable nonparametric regression model given in equation (1). If the function, $h_s(z_{si})$, s = 1,2, ..., q approached using Kernel Nadaraya-Watson then Kernel estimator for multivariable regression curve $\sum_{s=1}^{q} h_s(z_s)$ is given by:

$$\sum_{s=1}^{q} \hat{\hat{h}}_{\alpha_s}(z_s) = \mathbf{D}(\tilde{\alpha})\tilde{y}$$

$\mathbf{D}(\tilde{\alpha})$ matrix and $\tilde{y}$ vectors are follow:

$$\mathbf{D}(\tilde{\alpha}) = \begin{pmatrix} n^{-1}\sum_{k=1}^{q} W_{\alpha_k 1}(z_1) & n^{-1}\sum_{k=1}^{q} W_{\alpha_k 2}(z_1) & \cdots & n^{-1}\sum_{k=1}^{q} W_{\alpha_k n}(z_1) \\ n^{-1}\sum_{k=1}^{q} W_{\alpha_k 1}(z_2) & n^{-1}\sum_{k=1}^{q} W_{\alpha_k 2}(z_2) & \cdots & n^{-1}\sum_{k=1}^{q} W_{\alpha_k n}(z_2) \\ \vdots & \vdots & \vdots & \vdots \\ n^{-1}\sum_{k=1}^{q} W_{\alpha_k 1}(z_n) & n^{-1}\sum_{k=1}^{q} W_{\alpha_k 2}(z_n) & \cdots & n^{-1}\sum_{k=1}^{q} W_{\alpha_k n}(z_n) \end{pmatrix}, \text{ and } \tilde{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

Proof:

Because the regression curve $h_s(z_{si})$ was approached using Kernel Nadaraya and Watson then it can be written as:

$$\hat{h}_{\alpha_s}(z_{si}) = n^{-1}\sum_{j=1}^{n} W_{\alpha_{sj}}(z_{si}) y_j .$$

For s=1, obtain:

$$\hat{h}_{\alpha_1}(z_{1i}) = n^{-1}\sum_{j=1}^{n} W_{\alpha_1 j}(z_{1i}) y_j .$$

This is apply for $i = 1, 2, ..., n,$ then:

$$\begin{pmatrix} \hat{h}_{\alpha_1}(z_{11}) \\ \hat{h}_{\alpha_1}(z_{12}) \\ \vdots \\ \hat{h}_{\alpha_1}(z_{1n}) \end{pmatrix} = \begin{pmatrix} n^{-1}W_{\alpha_1 1}(z_{11}) & n^{-1}W_{\alpha_1 2}(z_{12}) & \cdots & n^{-1}W_{\alpha_1 n}(z_{11}) \\ n^{-1}W_{\alpha_1 1}(z_{12}) & n^{-1}W_{\alpha_1 2}(z_{12}) & \cdots & n^{-1}W_{\alpha_1 n}(z_{12}) \\ \vdots & \vdots & \vdots & \vdots \\ n^{-1}W_{\alpha_1 1}(z_{1n}) & n^{-1}W_{\alpha_1 2}(z_{1n}) & \cdots & n^{-1}W_{\alpha_1 n}(z_{1n}) \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

The above model can be written in the form:

$$\hat{\hat{h}}_{\alpha_1}(z_1) = \mathbf{D}(\alpha_1)\tilde{y}.$$

For $s = 2, ..., s = q$ given the equation:

$$\hat{\hat{h}}_{\alpha_2}(z_2) = \mathbf{D}(\alpha_2)\tilde{y}, \; ... \;, \; \hat{\hat{h}}_{\alpha_q}(z_q) = \mathbf{D}(\alpha_q)\tilde{y}$$

Consequently equation:

$$\sum_{s=1}^{q} \hat{\hat{h}}_s(z_s) = \mathbf{D}(\alpha_1)\tilde{y} + \mathbf{D}(\alpha_2)\tilde{y} + ... + \mathbf{D}(\alpha_q)\tilde{y} = \left[\mathbf{D}(\alpha_1) + \mathbf{D}(\alpha_2) + ... + \mathbf{D}(\alpha_q)\right]\tilde{y}$$

$$= \begin{pmatrix} n^{-1}\sum_{k=1}^{q} W_{\alpha_k 1}(z_1) & n^{-1}\sum_{k=1}^{q} W_{\alpha_k 2}(z_1) & \cdots & n^{-1}\sum_{k=1}^{q} W_{\alpha_k n}(z_1) \\ n^{-1}\sum_{k=1}^{q} W_{\alpha_k 1}(z_2) & n^{-1}\sum_{k=1}^{q} W_{\alpha_k 2}(z_2) & \cdots & n^{-1}\sum_{k=1}^{q} W_{\alpha_k n}(z_2) \\ \vdots & \vdots & \vdots & \vdots \\ n^{-1}\sum_{k=1}^{q} W_{\alpha_k 1}(z_n) & n^{-1}\sum_{k=1}^{q} W_{\alpha_k 2}(z_n) & \cdots & n^{-1}\sum_{k=1}^{q} W_{\alpha_k n}(z_n) \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

$$= \mathbf{D}(\tilde{\alpha})\tilde{y} \quad (4)$$

Then, using Lemma 1 and Lemma 2, can be obtained estimates for nonparametric regression curve Truncated Spline mixture and Kernel multivariable predictors using Least Square method. Shape mixture of Truncated Spline estimator and Kernel nonparametric regression multivariable predictors given by Theorem 1.

Theorem 1

Given the data pairs $(t_{1i}, ..., t_{pi}, z_{1i}, ..., z_{qi}, y_i)$, i = 1,2,...,n modeled nonparametric regression multivariable predictors of additives:

$$y_i = \mu(t_{1i}, ..., t_{pi}, z_{1i}, ..., z_{qi}) + \varepsilon_i = \sum_{r=1}^{p} g_r(t_{ri}) + \sum_{s=1}^{q} h_s(z_{si}) + \varepsilon_i \; .$$

If the regression curve $g_r(t_r)$, r =1,2,...,p and $h_s(z_s)$, s = 1,2,...,q assumed to be smooth and each approached using linear Truncated Spline function and Kernel Nadaraya-Watson, then:

a). Nonparametric Curve Regression Estimator for Multivariable Truncated Spline Component as follow:

$$\sum_{r=1}^{p}\hat{\tilde{g}}_r(t_r,\tilde{k},\tilde{\alpha}) = \mathbf{K}(\tilde{k},\tilde{\alpha})\tilde{y}$$

With $\mathbf{K}(\tilde{k},\tilde{\alpha}) = \mathbf{G}(\tilde{k})\left(\mathbf{G}(\tilde{k})^T\mathbf{G}(\tilde{k})\right)^{-1}\mathbf{G}(\tilde{k})^T\left(\mathbf{I}-\mathbf{D}(\tilde{\alpha})\right)$.

b). The mixed of Truncated Spline and Multivariable Kernel Estimator as follow:

$$\hat{\tilde{\mu}}(t_1,...,t_p,z_1,...,z_q,\tilde{k},\tilde{\alpha}) = \mathbf{M}(\tilde{k},\tilde{\alpha})\tilde{y}$$

With $\mathbf{M}(\tilde{k},\tilde{\alpha}) = \mathbf{K}(\tilde{k},\tilde{\alpha}) + \mathbf{D}(\tilde{\alpha})$

Proof :

Nonparametric regression model multivariable predictors of additives as follows:

$$y_i = \sum_{r=1}^{p} g_r(t_{ri}) + \sum_{s=1}^{q} h_s(z_{si}) + \varepsilon_i$$

If the regression curve $g_r(t_r)$, r =1,2,...,p approached using linear Truncated Spline function, and the regression curve, $h_s(z_s)$, s = 1,2,...,q approached with Kernel Nadaraya and Watson, then The Model of mixed Truncated Spline and multivariable predictor regression can be written into a matrix form:

$$\tilde{y} = \mathbf{G}(\tilde{k})\tilde{\theta} + \mathbf{D}(\tilde{\alpha})\tilde{y} + \tilde{\varepsilon}$$

Subsequently acquired norm for vector error:

$$\|\tilde{\varepsilon}\|^2 = \left\|\left(I-\mathbf{D}(\tilde{\alpha})\right)\tilde{y} - \mathbf{G}(\tilde{k})\tilde{\theta}\right\|^2$$

Nonparametric regression curve estimation of Mixed Spline truncated and Kernel multivariable predictors can be obtained using Least Square method. Shape estimator obtained from completing optimization:

$$\underset{\tilde{\theta}}{Min}\left(\|\tilde{\varepsilon}\|^2\right) = \underset{\tilde{\theta}}{Min}\left(\left\|\left(I-\mathbf{D}(\tilde{\alpha})\right)\tilde{y} - \mathbf{G}(\tilde{k})\tilde{\theta}\right\|^2\right)$$

With a calculation obtained by the equation:

$$\|\tilde{\varepsilon}\|^2 = \tilde{y}^T\left(\mathbf{I}-\mathbf{D}(\tilde{\alpha})\right)^T\left(\mathbf{I}-\mathbf{D}(\tilde{\alpha})\right)\tilde{y} - 2\tilde{\theta}^T\mathbf{G}(\tilde{k})^T\left(\mathbf{I}-\mathbf{D}(\tilde{\alpha})\right)\tilde{y} + \tilde{\theta}^T\mathbf{G}(\tilde{k})^T\mathbf{G}(\tilde{k})\tilde{\theta}$$

By using the obtained partial derivatives and the result is equated to zero obtained equation:

$$\frac{\partial \| \tilde{\varepsilon} \|^2}{\partial \tilde{\theta}} = -2\mathbf{G}(\tilde{k})^T \left(\mathbf{I} - \mathbf{D}\left(\tilde{\alpha}\right)\right) \tilde{y} + 2\mathbf{G}(\tilde{k})^T \mathbf{G}(\tilde{k}) \tilde{\theta} = 0$$

$$\mathbf{G}(\tilde{k})^T \mathbf{G}(\tilde{k}) \hat{\tilde{\theta}} = \mathbf{G}(\tilde{k})^T (\mathbf{I} - \mathbf{D}\left(\tilde{\alpha}\right)) \tilde{y}$$

The result obtained estimator:

$$\hat{\tilde{\theta}} = \left(\mathbf{G}(\tilde{k})^T \mathbf{G}(\tilde{k})\right)^{-1} \mathbf{G}(\tilde{k})^T \left(\mathbf{I} - \mathbf{D}\left(\tilde{\alpha}\right)\right) \tilde{y} = \mathbf{B}(\tilde{k}, \tilde{\alpha}) \tilde{y},$$

With matrix

$$\mathbf{B}(\tilde{k}, \tilde{\alpha}) = \left(\mathbf{G}(\tilde{k})^T \mathbf{G}(\tilde{k})\right)^{-1} \mathbf{G}(\tilde{k})^T \left(\mathbf{I} - \mathbf{D}\left(\tilde{\alpha}\right)\right).$$

Based on the obtained estimator $\hat{\tilde{\theta}}$ of multivariable truncated spline regression curve Spline components:

$$\sum_{r=1}^{p} \hat{\tilde{g}}_r(t_r, \tilde{k}, \tilde{\alpha}) = \mathbf{G}(\tilde{k}) \hat{\tilde{\theta}} = \mathbf{G}(\tilde{k}) \left(\mathbf{G}(\tilde{k})^T \mathbf{G}(\tilde{k})\right)^{-1} \mathbf{G}(\tilde{k})^T \left(\mathbf{I} - \mathbf{D}\left(\tilde{\alpha}\right)\right) \tilde{y} = \mathbf{K}(\tilde{k}, \tilde{\alpha}) \tilde{y}$$

Furthermore, estimator of mixed nonparametric regression curve Spline and Kernel multivariable predictor is given by:

$$\hat{\tilde{\mu}}(t_1, ..., t_p, z_1, ..., z_q, \tilde{k}, \tilde{\alpha}) = \mathbf{K}(\tilde{k}, \tilde{\alpha}) \tilde{y} + \mathbf{D}(\tilde{\alpha}) \tilde{y} = \left(\mathbf{K}(\tilde{k}, \tilde{\alpha}) + \mathbf{D}(\tilde{\alpha})\right) \tilde{y} = \mathbf{M}(\tilde{k}, \tilde{\alpha}) \tilde{y}.$$

Estimator of curve nonparametric regression using Mixed truncated Spline multivariable predictors Kernel $\hat{\tilde{\mu}}(t_1, ..., t_p, z_1, ..., z_q, \tilde{k}, \tilde{\alpha})$ depend on the point of knot and bandwidth parameters. To get the best mixed estimator, necessary to point selection parameter knot and optimal bandwidth. A method that can be used to select a point knot and optimal bandwidth parameter is the method of Generalized Cross Validation (GCV). GCV function is given by:

$$V\left(\tilde{k}, \tilde{\alpha}\right) = \frac{n^{-1} \| (\mathbf{I} - \mathbf{M}(\tilde{k}, \tilde{\alpha})) \tilde{y} \|^2}{\left(n^{-1} tr\left(\mathbf{I} - \mathbf{M}(\tilde{k}, \tilde{\alpha})\right)\right)^2}$$

Point knot and optimal bandwidth parameters obtained from completing optimization:

$$V\left(\tilde{k}_{opt}, \tilde{\alpha}_{opt}\right) = \underset{\tilde{k}, \tilde{\alpha}}{Min} \left\{ V\left(\tilde{k}, \tilde{\alpha}\right) \right\}.$$

3. PROPERTIES OF MIXED ESTIMATION

The properties of multivariable estimator truncated Spline component, multivariable Kernel component, and a mixed of truncated Spline and Kernel in nonparametric regression multivariable, is given by the following theorem.

Theorem 2

If the estimator

$$\sum_{r=1}^{p}\hat{\tilde{g}}_r(t_r,\tilde{k},\tilde{\alpha}),\ \sum_{s=1}^{q}\hat{\tilde{h}}_{\alpha_s}(z_s), \text{and}\ \ \hat{\tilde{\mu}}(t_1,...,t_p,z_1,...,z_q,\tilde{k},\tilde{\alpha})$$

given by Theorem 1 and Lemma 2, then estimator estimator is a biased estimator and a linear estimator class under observation.

Proof :

Theorem 1 given:

$$\sum_{r=1}^{p}\hat{\tilde{g}}_r(t_r,\tilde{k},\tilde{\alpha})=\mathbf{K}(\tilde{k},\tilde{\alpha})\tilde{y}$$

The above results show that the truncated multivariable estimator Spline component is a linear estimator in the observation. Subsequently obtained:

$$E\left[\sum_{r=1}^{p}\hat{\tilde{g}}_r(t_r,\tilde{k},\tilde{\alpha})\right]=\mathbf{K}(\tilde{k},\tilde{\alpha})E\left[\tilde{y}\right]=\mathbf{K}(\tilde{k},\tilde{\alpha})\left(I-\mathbf{D}\left(\tilde{\alpha}\right)\right)^{-1}\left(\mathbf{G}(\tilde{k})\tilde{\theta}+E\left[\tilde{\varepsilon}\right]\right)$$

$$=\mathbf{K}(\tilde{k},\tilde{\alpha})\left(I-\mathbf{D}\left(\tilde{\alpha}\right)\right)^{-1}\mathbf{G}(\tilde{k})\tilde{\theta}\neq\ \mathbf{G}(\tilde{k})\tilde{\theta}\ =\sum_{r=1}^{p}\hat{\tilde{g}}_r(t_r,\tilde{k},\tilde{\alpha})$$

Lemma 2given:

$$\sum_{s=1}^{q}\hat{\tilde{h}}_s(z_s)=\sum_{s=1}^{q}\mathbf{D}(\alpha_s)\tilde{y}$$

Then the Kernel multivariable estimator component is a linear estimator in the observation. Furthermore, the expected value of the estimator is as follows:

$$E\left[\sum_{s=1}^{q}\hat{\tilde{h}}_{\alpha_s}(z_s)\right]=\mathbf{D}(\tilde{\alpha})E\left[\tilde{y}\right]=\sum_{r=1}^{p}\tilde{g}_r(t_r)+\sum_{s=1}^{q}\tilde{h}_s(z_s)+E\left[\tilde{\varepsilon}\right]$$

$$=\sum_{r=1}^{p}\tilde{g}_r(t_r)+\sum_{s=1}^{q}\tilde{h}_s(z_s)\neq\sum_{s=1}^{q}\tilde{h}_s(z_s)$$

Theorem 1 gives the estimator a mixed Truncated Spline and multivariable Kernel:

$\hat{\tilde{\mu}}(t_1,...,t_p,z_1,...,z_q,\tilde{k},\tilde{\alpha})=\mathbf{M}(\tilde{k},\tilde{\alpha})\tilde{y}.$

The the mixed estimator is an linear estimator in observation. Subsequently obtained:

$$E\left[\hat{\tilde{\mu}}(t_1,...,t_p,z_1,...,z_q,\tilde{k},\tilde{\alpha})\right]=\mathbf{M}(\tilde{k},\tilde{\alpha})E\left[\tilde{y}\right]=\mathbf{M}(\tilde{k},\tilde{\alpha})\left(\sum_{r=1}^{p}\tilde{g}_r(t_r)+\sum_{s=1}^{q}\tilde{h}_s(z_s)\right)$$

$$\neq\left(\sum_{r=1}^{p}\tilde{g}_r(t_r)+\sum_{s=1}^{q}\tilde{h}_s(z_s)\right)=\tilde{\mu}(t_1,...,t_p,z_1,...,z_q,\tilde{k},\tilde{\alpha})$$

4. CONCLUSION

Given the data pairs $(t_{1i},...,t_{pi},z_{1i},...,z_{qi},y_i)$, i = 1,2,...,n modeled nonparametric regression multivariable predictors of additives:

$$y_i = \mu(t_{1i},...,t_{pi},z_{1i},...,z_{qi}) + \varepsilon_i = \sum_{r=1}^{p} g_r(t_{ri}) + \sum_{s=1}^{q} h_s(z_{si}) + \varepsilon_i \, .$$

If the regression curve $g_r(t_r)$, r =1,2,...,$p$and $h_s(z_s)$, s = 1,2,...,$q$assumed to be smooth and each approached using linear Truncated Spline function and Kernel Nadaraya-Watson, then:

a). Nonparametric Curve Regression Estimator for Multivariable Truncated Spline Component as follow:

$$\sum_{r=1}^{p} \hat{\tilde{g}}_r(t_r,\tilde{k},\tilde{\alpha}) = \mathbf{K}(\tilde{k},\tilde{\alpha})\tilde{y}$$

With $\mathbf{K}(\tilde{k},\tilde{\alpha}) = \mathbf{G}(\tilde{k})\left(\mathbf{G}(\tilde{k})^T\mathbf{G}(\tilde{k})\right)^{-1}\mathbf{G}(\tilde{k})^T\left(\mathbf{I}-\mathbf{D}\left(\tilde{\alpha}\right)\right).$

b). The mixed of Truncated Spline and Multivariable Kernel Estimator as follow:

$$\sum_{s=1}^{q} \hat{\tilde{h}}_{\alpha_s}(z_s) = \mathbf{D}(\tilde{\alpha})\tilde{y}$$

c). Estimates for nonparametric regression curve by mixed Truncated Spline and Kernel multivariable given by:

$$\hat{\tilde{\mu}}(t_1,...,t_p,z_1,...,z_q,\tilde{k},\tilde{\alpha}) = \left(\mathbf{K}(\tilde{k},\tilde{\alpha})+\mathbf{D}(\tilde{\alpha})\right)\tilde{y} = \mathbf{M}(\tilde{k},\tilde{\alpha})\tilde{y}.$$

d). Estimator of component of multivariable truncated Spline estimator and component of multivariable Kernel, and the mixed Truncated Spline and Multivariable Kernel Estimator are biased estimator, but belong to the class of linear estimator.

c). The Estimator of mixture Truncated Spline and Kernel in nonparametric regression multivariable depend on the parameters of bandwidth and knot. Estimator obtained the best mix of bandwidth parameters and the optimal point knot. Parameter bandwidth and optimal point of knot that can be obtained using the method of GCV.

ACKNOWLEDGEMENT

This article is a part of the research result of *Unggulan Perguruan Tinggi*funded by DP2M-DIKTI in 2016. The author will like to thank the General Director of DIKTI in Jakarta for the fund provided.

REFERENCES

[1] G. Wahba, *Spline Models For Observasion Data*, SIAM Pensylvania, 1990.

[2] R. L. Eubank, *Spline Smoothing and Nonparametric Regression*, Mercel Dekker, New York, 1988.

[3] I. N. Budiantara, M. Ratna, I. Zain and W. Wibowo, Modeling the Percentage of Poor People in Indonesia Using Spline Nonparametric Regression Approach, *International Journal Basic & Applied Sciences*, 12, (2012), 118-124.

[4] L. J. Asriniand I. N. Budiantara, Fourier Series Semiparametric Regression Models (Case Study : The Production of Lawland Rice Irrigation in Central Java), *ARPN Journal of Engineering and Applied Sciences*, 9 (2014), 1501-1506.

[5] N. Chamidah, I. N. Budiantara, S. Sunaryo and I. Zain,Disigning of Child Growth Chart Based on Multirespon Local Polynomial Modeling, *Journal of Mathematics and Statistics*,8 (2012),342-347.

[6] I. N. Budiantara,Estimator Spline Terbobot Dalam Regresi Semiparametrik, *Majalah Ilmu Pengetahuan dan Teknologi*, 10 (1999), 103-109.

[7]I. N. Budiantara,Model Keluarga Spline Polinomial Truncated dalam Regresi Semiparametrik,*Berkala Ilmiah MIPA*, 15 (2005), 55-61.

[8]I. N. Budiantara,Model Spline dengan Knot Optimal,*Jurnal Ilmu Dasar, FMIPA Universitas Jember*, 7 (2006), 77-85.

[9] R. Diana, I. N. Budiantara, Purhadi and S. Darmesto, Statistical Modeling for Unemployment Rate Using Smoothing Spline in Semiparametric Multivariable Regression Model with Bayesian Approach, *Model Assisted Statistics and Applications*, 9 (2014),159-166.

[10] E. T. Astuti, I. N. Budiantara, S. Sunaryo and M. Dokhi, Statistical Modeling for Mortality Data Using Local Generalized Poisson Regression Model, *International Journal of Applied Mathematics and Statistics*, 33 (2013), 92-101.

[11] I. N. Budiantara, Subanar and Z. Soejoeti, Weighted Spline Estimator, *Bulletin of the International Statistical Insitute*, 51 (1997), 333-334.

[12] A. A. R. Fernandes, I. N. Budiantara, B. W.Otokand Suhartono, Spline Estimator for Bi-responses Nonparametric Regression Model for Longitudinal Data, *Applied Mathematical Sciences*, 8 (2014), 5653-5665.

[13] B. Lestari, I. N. Budiantara, S. Sunaryo and M. Mashuri, Spline Smoothing forMultiresponse NonparametricRegression Model in Case of Heteroscedasticity of Variance, *Journal of Mathematics and Statististics*, 8 (2012), 377-384.

[14] B. Lestari, I. N. Budiantara, S. Sunaryo and M. Mashuri, Spline Estimator in Multiresponse Nonparametric Regression Model with Unequal Correlation of Errors, *Journal of Mathematics and Statististics*, 6 (2010), 327-332.

[15] W. Wibowo, S. Haryatmi and I. N. Budiantara, On MultiresponseSemiparametric Regression Model, *Journal of Mathematics*

and Statistics, 8 (2012), 489-499.

[16] I. N. Budiantara,*Spline Dalam Regresi Nonparametrik dan Semiparametrik: Sebuah Pemodelan Statistika Masa Kini dan Masa Mendatang*, Pidato Pengukuhan Untuk Jabatan Guru Besar (Profesor) pada Jurusan Statistika, Institut Teknologi Sepuluh Nopember (ITS) Surabaya, Indonesia, ITS press, 2009.

[17] M. Kayri and G. Zirhhoglu, Kernel Smoothing Function and Choosing Bandwitdh for Nonparametric Regression Methods, *Ozean Journal of Applied Sciences*, 2 (2009), 49-60.

[18] 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*, 140 (2010), 1389-1397.

[19] M. Bilodeau, Fourier Smoother and Additive Models, *The Canadian Journal of Statistics*, 3 (1992), 257-269.

[20] R. Pane, I. N. Budiantara,I.Zain and B. W. Otok,Parametric and Nonparametric Estimators in Fourier Series Semiparametric Regression and Their Characteristics, *Applied Mathematical Sciences*, 8 (2014), 5053-5064.

[21] I. N. Budiantara,V. Ratnasari, M. Ratna, and I. Zain, The Combination of Spline and Kernel Estimator for Nonparametric Regression and its Properties,*Applied Mathematical Sciences*, 9 (2015), 6083-6094.