Songklanakarin J. Sci. Technol.
42 (3), 533-548, May - Jun. 2020

http://www.sjst.psu.ac.th

Original Article

Spline estimator and its asymptotic properties in multiresponse nonparametric regression model

Budi Lestari[1], Fatmawati[2*], and I Nyoman Budiantara[3]

[1] *Department of Mathematics, Faculty of Mathematics and Natural Sciences, The University of Jember, Tegal Boto Campus, Jember, 68121 Indonesia*

[2] *Department of Mathematics, Faculty of Sciences and Technology, Airlangga University, UNAIR Campus C, Surabaya, 60115 Indonesia*

[3] *Department of Statistics, Sepuluh Nopember Institute of Technology, Sukolilo, Surabaya, 60111 Indonesia*

Received: 16 September 2018; Revised: 28 December 2018; Accepted: 14 February 2019

Abstract

In applications, we often meet the problem where more than one response variable is observed at several values of predictor variables, and these responses are correlated with each other. The multiresponse nonparametric regression model approach is appropriate to model the functions which represent relationship between response and predictor variables. This relationship is drawn by the regression function. The principal problem of this model approach is estimating of the regression function of this model. The spline estimator is one of the most popular estimators used for estimating it. In this paper we discuss methods to obtain a smoothing spline estimator for estimating the regression function, to get a covariance matrix estimator, and to choose an optimum smoothing parameter. In addition, we investigate the asymptotic properties of the smoothing spline estimator.

Keywords: multiresponse nonparametric regression, covariance matrix, spline estimator, smoothing parameter, asymptotic properties

1. Introduction

Statistical analysis often involves building mathematical models which examine association between response and predictor variables. Spline smoothing is a general class of powerful and flexible modeling techniques. Research on smoothing spline models has attracted a great deal of attention in recent years, and the methodology has been widely used in many areas. Smoothing spline estimator with its powerful and flexible properties is one of the most popular estimators used for estimating regression function of the nonparametric regression model. Several types of spline estimator have been considered by researchers to estimate the regression function. Original spline was used to estimate the regression function for smooth data by Kimeldorf and Wahba (1971), Craven and Wahba (1979), and Wahba (1990). M-type spline was proposed by Cox (1983), and Cox and O'Sullivan (1996) to overcome outliers in nonparametric regression. Construction of confidence interval for original spline model has been provided by Wahba (1983). A comparison between generalized cross validation and generalized maximum likelihood for choosing a smoothing parameter in the generalized spline smoothing problem was presented by Wahba (1985). Relaxed spline and quantile spline were introduced by Oehlert (1992), and Koenker, Pin, and Portnoy (1994), respectively. Smoothing spline for the case of correlated errors was discussed by Wang (1998). Reproducing kernel Hilbert spaces (RKHS) concept has been used by

*Corresponding author
Email address: fatmawati@fst.unair.ac.id

Wahba (2000) to build spline statistical model. Lee (2004) combined smoothing spline estimates of different smoothness to form a final improved estimate. Cardot, Crambes, Kneip, and Sarda (2007) investigated the asymptotic property of smoothing splines in functional linear regression with errors-in-variables. Liu, Tong, and Wang (2007) have discussed smoothing spline estimator for variance functions. Aydin (2007) compared goodness of spline and kernel in estimating nonparametric regression model for gross national product data. Aydin, Memmedhi, and Omay (2013) have studied the determination of an optimum smoothing parameter for nonparametric regression using smoothing spline. But, researchers mentioned above just discussed spline estimators for estimating regression function of single response nonparametric regression models. It means that they have not discussed spline estimators in the multiresponse nonparametric regression model.

In many real cases, we often find cases where more than one response variable is observed at several values of predictor variables, and there are correlations between the response and each other. Multiresponse nonparametric regression model provides potential methods to model the functions that represent the relationship of these variables. Some researchers have discussed estimating methods in the multiresponse nonparametric models. Wegman (1981), Miller and Wegman (1987), and Flessler (1991) provided spline smoothing algorithms. Wahba (1992) used RKHS method to develop the theory of general smoothing splines. Gooijer, Gannoun, and Larramendy (1999), and Fernandez and Opsomer (2005) proposed methods to estimate nonparametric regression models with serially and spatially correlated errors, respectively. Wang, Guo, and Brown (2000) used spline smoothing for estimating biresponse nonparametric regression model with the same correlation of errors. Lestari, Budiantara, Sunaryo, and Mashuri (2009), and Lestari, Budiantara, Sunaryo, and Mashuri (2010) used spline to estimate the multiresponse nonparametric regression model in cases of equal correlation of errors and unequal correlation of errors, respectively. Chamidah, Budiantara, Sunaryo, and Zain (2012) applied the multiresponse nonparametric regression model to design child growth chart. Lestari, Budiantara, Sunaryo, and Mashuri (2012) have studied spline to estimate the heteroscedastic multiresponse nonparametric regression model. Chamidah and Lestari (2016) discussed estimation of the homoscedastic multiresponse nonparametric regression model when the numbers of observations were unbalanced. Lestari, Fatmawati, and Budiantara (2017) estimated smoothing spline in the multiresponse nonparametric regression model by using RKHS method. Lestari, Fatmawati, Budiantara, and Chamidah (2018), and Lestari, Fatmawati, Budiantara, and Chamidah (2019) estimated regression functions and smoothing parameters using spline and kernel estimators. Yet, all these researchers assumed that the covariance matrix was known. When it is unknown, it has to be estimated from the data and it can affect the estimates of the smoothing parameters (Wang, 1998). Also, these researchers have not discussed the estimation of optimum smoothing parameter in the multiresponse nonparametric regression model when the variances of errors are not the same. In addition, none of these researchers have discussed the asymptotic properties of the spline estimator.

In nonparametric regression, we often consider asymptotic properties of estimator based on certain goodness criteria. In regression nonparametric analysis, investigating of the asymptotic properties of an estimator is an important step for obtaining convergence rate of regression function estimator. There are some criteria which have often been used by researches to determine goodness of spline estimator approach. Eubank (1988) proposed the mean square error criterion. Speckman (1985), and Carter, Eagleson, and Silver man (1994) have studied minimax criterion, while Cox (1983) and Cox and O'Sullivan (1996) considered the mean square error criterion for M-type spline approach. Eggermont, Eubank, and LaRiccia (2010) have studied convergence rate of spline estimator in the varying coefficient model. Li and Zhang (2010) established strong consistency and asymptotic normality of penalized spline in the varying-coefficient single-index model. Wang (2012) constructed the M-type estimator of regression function and has studied its asymptotic normality property. Ping and Lin (2013) discussed the asymptotic properties of spline estimator in the partly linear model for longitudinal data. Chen and Christensen (2015) have studied both asymptotic properties and convergence rate of some estimators in the nonparametric model. Although, these researches have discussed some asymptotic properties of some estimators, they discussed the asymptotic properties in single response nonparametric regression models only. They have not discussed the asymptotic properties in multiresponse nonparametric regression models.

In this paper, we develop the biresponse nonparametric regression model proposed by Wang et al. (2000) to the more general model, i.e., the multiresponse nonparametric regression model. Note that we need the covariance matrix to determine a weight matrix that will be used in the optimization penalized weighted least-square (PWLS) to obtain a smoothing spline estimator which depends on the smoothing parameter to estimate regression function of the model. Therefore, in this paper we discuss methods to obtain the smoothing spline estimator, to get the covariance matrix estimator, and to choose the optimum smoothing parameter. We also investigate the asymptotic properties of smoothing spline estimators of the multiresponse nonparametric regression model based on integrated mean square error (IMSE) criterion.

2. Methods

Firstly, we consider data (y_{ki}, t_{ki}), $k = 1,2,...,p$; $i = 1,2,...,n_k$ that follow the multiresponse nonparametric regression model as follows:

$$y_{ki} = f_k(t_{ki}) + \varepsilon_{ki} \tag{1}$$

where $f_1, f_2, ..., f_p$ are unknown regression functions assumed to be smoothed and contained in Sobolev space $W_2^m[a_k, b_k]$. ε_{ki} are zero-mean independent random errors with variance σ_{ki}^2. Based on (1), we can determine the covariance matrix of errors. The estimated multiresponse nonparametric regression model based on smoothing spline estimator can be obtained by carrying out optimization PWLS:

$$\underset{f_1,f_2,...,f_p\in W_2}{Min}\{(\sum_{k=1}^{p}n_k)^{-1}(\underset{\sim}{y}_1-\underset{\sim}{f}_1)'W_1(\underset{\sim}{\sigma}_1^2)(\underset{\sim}{y}_1-\underset{\sim}{f}_1)+...+(\underset{\sim}{y}_p-\underset{\sim}{f}_p)'W_p(\underset{\sim}{\sigma}_p^2)(\underset{\sim}{y}_p-\underset{\sim}{f}_p)+$$
$$\lambda_1\int_{a_1}^{b_k}(f_1^{(2)}(t))^2dt+...+\lambda_p\int_{a_p}^{b_p}(f_p^{(2)}(t))^2dt\} \quad (2)$$

Then, to get the solution of optimization PWLS in (2), we use RKHS method. Secondly, we estimate the covariance matrix of errors using maximum likelihood method, and estimate the optimum smoothing parameter based on the minimum value of generalized cross-validation (GCV). Finally, we determine the asymptotic properties of the smoothing spline estimator of regression function based on IMSE criterion.

3. Results and Discussion

Suppose that $\underset{\sim}{y}=(\underset{\sim}{y}_1,\underset{\sim}{y}_2,...,\underset{\sim}{y}_p)'$, $\underset{\sim}{f}=(\underset{\sim}{f}_1,\underset{\sim}{f}_2,..,\underset{\sim}{f}_p)'$, $\underset{\sim}{\varepsilon}=(\underset{\sim}{\varepsilon}_1,\underset{\sim}{\varepsilon}_2,...,\underset{\sim}{\varepsilon}_p)'$, and $\underset{\sim}{t}=(\underset{\sim}{t}_1,\underset{\sim}{t}_2,...,\underset{\sim}{t}_p)'$ where $\underset{\sim}{y}_k=(y_{k1},y_{k2},...,y_{kn_k})'$, $\underset{\sim}{f}_k=(f_k(t_{k1}),f_k(t_{k2}),...,f_k(t_{kn_k}))'$, $\underset{\sim}{\varepsilon}_k=(\varepsilon_{k1},\varepsilon_{k2},...,\varepsilon_{kn_k})'$, $\underset{\sim}{t}_k=(t_{k1},t_{k2},...,t_{kn_k})'$, $k=1,2,...,p$; $i=1,2,...,n_k$. Therefore, the model (1) can be written in the following matrix notation:

$$\underset{\sim}{y}=\underset{\sim}{f}+\underset{\sim}{\varepsilon} \quad (3)$$

where $E(\underset{\sim}{\varepsilon})=\underset{\sim}{0}$, and namely $Cov(\underset{\sim}{\varepsilon})=[W(\underset{\sim}{\sigma}^2)]^{-1}$.

3.1 Covariance matrix of random errors

The following theorem gives us how we determine the covariance matrix of random errors, $Cov(\underset{\sim}{\varepsilon})=[W(\underset{\sim}{\sigma}^2)]^{-1}$, in the multiresponse nonparametric regression model.

Theorem 1. Suppose that the data set (y_{ki},t_{ki}), $k=1,2,...,p$; $i=1,2,...,n_k$ follows the multiresponse nonparametric regression model given in (1):

$$y_{ki}=f_k(t_{ki})+\varepsilon_{ki}.$$

If $[W(\underset{\sim}{\sigma}^2)]^{-1}$ denotes the covariance matrix of errors, then

$$[W(\underset{\sim}{\sigma}^2)]^{-1}=diag(W_1(\underset{\sim}{\sigma}_1^2),W_2(\underset{\sim}{\sigma}_2^2),...,W_p(\underset{\sim}{\sigma}_p^2))$$

where matrix $W_k(\underset{\sim}{\sigma}_k^2)$ is given by:

$$W_k(\underset{\sim}{\sigma}_k^2)=\begin{bmatrix}\sigma_{k1}^2 & \sigma_{k(1,2)} & \cdots & \sigma_{k(1,n_1)}\\ \sigma_{k(2,1)} & \sigma_{k2}^2 & \cdots & \sigma_{k(2,n_2)}\\ \vdots & \vdots & \ddots & \vdots\\ \sigma_{k(n_k,1)} & \sigma_{k(n_k,2)} & \cdots & \sigma_{kn_k}^2\end{bmatrix},\ k=1,2,...,p.$$

Proof. $Cov(\underset{\sim}{\varepsilon})=E(\underset{\sim}{\varepsilon}-E(\underset{\sim}{\varepsilon}))(\underset{\sim}{\varepsilon}-E(\underset{\sim}{\varepsilon}))'$

$$=E[(\varepsilon_{11},...,\varepsilon_{1n_1},\varepsilon_{21},...,\varepsilon_{2n_2},...,\varepsilon_{p1},...,\varepsilon_{pn_p})'(\varepsilon_{11},...,\varepsilon_{1n_1},\varepsilon_{21},...,\varepsilon_{2n_2},...,\varepsilon_{p1},...,\varepsilon_{pn_p})]$$

$$=\begin{bmatrix}W_{11} & W_{12} & \cdots & W_{1p}\\ W_{21} & W_{22} & \cdots & W_{2p}\\ \vdots & \vdots & \ddots & \vdots\\ W_{p1} & W_{p2} & \cdots & W_{pp}\end{bmatrix}=[W(\underset{\sim}{\sigma}^2)]^{-1} \quad (4)$$

where $W_{kl} = \begin{cases} diag(\sigma_{k1}^2, \sigma_{k2}^2, ..., \sigma_{kn_k}^2), & k = l(k = 1, 2, ..., p) \\ diag(\sigma_{(kl)1}, \sigma_{(kl)2}, ..., \sigma_{(kl)n_k}), & k \neq l \end{cases}$.

Since $\sigma_{kl} = \rho_{kl}\sigma_k\sigma_l$ and $\rho_{kl} = \begin{cases} \rho_k, & k = l \\ 0, & k \neq l \end{cases}$, then we have: $W_{kl} = \begin{cases} diag(\sigma_{k1}^2, \sigma_{k2}^2, ..., \sigma_{kn_k}^2), & k = l(k = 1, 2, ..., p) \\ 0, & k \neq l \end{cases}$.

Therefore, we can write equation (4) as follows:

$$[W(\underset{\sim}{\sigma}^2)]^{-1} = diag(W_1(\underset{\sim}{\sigma}_1^2), W_2(\underset{\sim}{\sigma}_2^2), ..., W_p(\underset{\sim}{\sigma}_p^2))$$

where matrix $W_k(\underset{\sim}{\sigma}_k^2)$ is given by:

$$W_k(\underset{\sim}{\sigma}_k^2) = \begin{bmatrix} \sigma_{k1}^2 & \sigma_{k(1,2)} & \cdots & \sigma_{k(1,n_1)} \\ \sigma_{k(2,1)} & \sigma_{k2}^2 & \cdots & \sigma_{k(2,n_2)} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{k(n_k,1)} & \sigma_{k(n_k,2)} & \cdots & \sigma_{kn_k}^2 \end{bmatrix}, \; k = 1,2,..., p.$$

3.2 Estimation of regression function

Suppose that data set (y_{ki}, t_{ki}), $k = 1,2,..., p$, $i = 1, 2,..., n_k$, follows the model given in (1). We can determine the smoothing spline estimator for regression function of the model given in (1) by solving the following PWLS:

$$\begin{aligned} &\underset{f_1, f_2, ..., f_p \in W_2^m}{Min}\{(\sum_{k=1}^{p} n_k)^{-1}(\underset{\sim}{y}_1 - \underset{\sim}{f}_1)'W_1(\underset{\sim}{\sigma}_1^2)(\underset{\sim}{y}_1 - \underset{\sim}{f}_1) + ... + (\underset{\sim}{y}_p - \underset{\sim}{f}_p)'W_p(\underset{\sim}{\sigma}_p^2)(\underset{\sim}{y}_p - \underset{\sim}{f}_p) + \\ &\lambda_1\int_{a_1}^{b_1}(f_1^{(2)}(t))^2 dt + ... + \lambda_p\int_{a_p}^{b_p}(f_p^{(2)}(t))^2 dt\} \end{aligned} \quad (5)$$

where $\underset{\sim}{y}_1 = (y_{11}, y_{12}, ..., y_{1n_1})', \underset{\sim}{y}_2 = (y_{21}, y_{22}, ..., y_{2n_2})', ..., \underset{\sim}{y}_p = (y_{p1}, y_{p2}, ..., y_{pn_p})'$; $\underset{\sim}{f}_1 = (f_{11}, f_{12}, ..., f_{1n_1})', \underset{\sim}{f}_2 = (f_{21}, f_{22}, ..., f_{2n_2})', ..., \underset{\sim}{f}_p = (f_{p1}, f_{p2}, ..., f_{pn_p})'$, and W_k, represents the k^{th}-weight, $k = 1,2,..., p$ which is obtained from Theorem 1, and the smoothing parameter λ_k ($k = 1,2,..., p$) controls the trade-off between the goodness of fit and the smoothness of the estimate. The model given in (1) can be expressed into a general smoothing spline regression model as follows:

$$y_{ki} = L_{t_k} f_k + \varepsilon_{ki}, \quad k = 1,2,..., p; \; i = 1, 2, ..., n_k$$

where function $f_k \in$ F$_k$ (F$_k$ represents Hilbert space) is unknown and assumed to be smooth, and $L_{t_k} \in$ F$_k$ is a bounded linear functional.

Next, suppose that Hilbert space **F**$_k$ can be decomposed into direct sum of two subspaces G$_k$ and H$_k$ as follows:

F$_k$ = G$_k$ $\oplus$ H$_k$

where G$_k$ $\perp$ H$_k$. Also, suppose that $\{\varsigma_{k1}, \varsigma_{k2}, ..., \varsigma_{km_k}\}$ and $\{\vartheta_{k1}, \vartheta_{k2}, ..., \vartheta_{kn_k}\}$ are bases of spaces of G$_k$ and H$_k$, respectively. Then, we can express every function $f_k \in$ **F**$_k$ ($k = 1,2,..., p$) into the following expression:

$$f_k = g_k + h_k$$

where $g_k \in$ G$_k$ and $h_k \in$ H$_k$. Since $\{\varsigma_{k1}, \varsigma_{k2}, ..., \varsigma_{km_k}\}$ is basis of space G$_k$, and $\{\vartheta_{k1}, \vartheta_{k2}, ..., \vartheta_{kn_k}\}$ is basis of space H$_k$, then for every function $f_k \in$ **F**$_k$ ($k = 1,2,..., p$) follows:

$$f_k = \sum_{v=1}^{m_k} c_{kv}\varsigma_{kv} + \sum_{i=1}^{n_k} d_{ki}\vartheta_{ki} = \underset{\sim}{\varsigma}_k' b_k + \underset{\sim}{\vartheta}_k' \underset{\sim}{c}_k\,;\; k = 1,2,...,p\,;\; c_{kv}, d_{ki} \in \mathrm{R} \tag{6}$$

where $\underset{\sim}{\varsigma}_k = (\varsigma_{k1}, \varsigma_{k2}, ..., \varsigma_{km_k})'$, $\underset{\sim}{c}_k = (c_{k1}, c_{k2}, ..., c_{km_k})'$, $\underset{\sim}{\vartheta}_k = (\vartheta_{k1}, \vartheta_{k2}, ..., \vartheta_{kn_k})'$, and $\underset{\sim}{d}_k = (d_{k1}, d_{k2}, ..., d_{kn_k})'$. Finally, since $L_{t_{ki}}$ is the bounded linear function in $\mathbf{F}_{\boldsymbol{k}}$, and $f_k \in \mathbf{F}_{\boldsymbol{k}}$, then we have:

$$L_{t_{ki}} f_k = L_{t_{ki}}(g_k + h_k) = L_{t_{ki}}(g_k) + L_{t_{ki}}(h_k) = g_k(t_{ki}) + h_k(t_{ki}) = f_k(t_{ki}).$$

In the following theorem, we give a method to obtain the estimated regression function by using RKHS method, i.e., by carrying out the PWLS given in (5).

Theorem 2. If the data set (y_{ki}, t_{ki}), $k = 1,2,..., p$; $i = 1,2,..., n_k$ follows the multiresponse nonparametric regression model given in (1):

$$y_{ki} = f_k(t_{ki}) + \varepsilon_{ki}$$

then the estimated regression function of model (1) based on smoothing spline estimator is given by:

$$\underset{\sim}{\hat{f}}_\lambda = A\underset{\sim}{\hat{c}} + B\underset{\sim}{\hat{d}} = H(\underset{\sim}{\lambda})\underset{\sim}{y}$$

where $H(\underset{\sim}{\lambda}) = A[A'D^{-1}W(\underset{\sim}{\sigma}^2)A]^{-1}A'D^{-1}W(\underset{\sim}{\sigma}^2) + BD^{-1}W(\underset{\sim}{\sigma}^2) \times [I - A(A'D^{-1}W(\underset{\sim}{\sigma}^2)A)^{-1}A'D^{-1}W(\underset{\sim}{\sigma}^2)]$.

Proof. By considering model given in (1) and applying Riesz representation theorem (Wang, 2011), and because of $L_{t_{ki}}$ is bounded linear functional in $\mathbf{F}_{\boldsymbol{k}}$, then there is a representer $\delta_{ki} \in \mathbf{F}_{\boldsymbol{k}}$ of $L_{t_{ki}}$ which follows (Wang, 2011):

$$L_{t_{ki}} f_k = \langle \delta_{ki}, f_k \rangle = f_k(t_{ki}), \quad f_k \in \mathbf{F}_{\boldsymbol{k}}$$

where $\langle .,. \rangle$ represents an inner product. Based on (6) and by considering the properties of inner product, we have:

$$f_k(t_{ki}) = \langle \delta_{ki}, \underset{\sim}{\varsigma}_k' c_k + \underset{\sim}{\vartheta}_k' \underset{\sim}{d}_k \rangle = \langle \delta_{ki}, \underset{\sim}{\varsigma}_k' \underset{\sim}{c}_k \rangle + \langle \delta_{ki}, \underset{\sim}{\vartheta}_k' \underset{\sim}{d}_k \rangle. \tag{7}$$

Next, based on (7), for $k = 1$ we have:

$$f_1(t_{1i}) = \langle \delta_{1i}, \underset{\sim}{\varsigma}_1' \underset{\sim}{c}_1 \rangle + \langle \delta_{1i}, \underset{\sim}{\vartheta}_1' \underset{\sim}{d}_1 \rangle, \; i = 1,2,..., n_1;$$

So that, for $i = 1,2,3,..., n_1$ we have:

$$\underset{\sim}{f}_1(t_1) = \left(f_1(t_{11}), f_1(t_{12}), ..., f_1(t_{1n_1}) \right)' = A_1\underset{\sim}{c}_1 + B_1\underset{\sim}{d}_1$$

where: $\underset{\sim}{c}_1 = (c_{11}, c_{12}, ..., c_{1m_1})'$, $\underset{\sim}{d}_1 = (d_{11}, d_{12}, ..., d_{1n_1})'$,

$$A_1 = \begin{bmatrix} \langle \delta_{11}, \varsigma_{11} \rangle & \langle \delta_{11}, \varsigma_{12} \rangle & \cdots & \langle \delta_{11}, \varsigma_{1m_1} \rangle \\ \langle \delta_{12}, \varsigma_{11} \rangle & \langle \delta_{12}, \varsigma_{12} \rangle & \cdots & \langle \delta_{12}, \varsigma_{1m_1} \rangle \\ \vdots & \vdots & \vdots & \vdots \\ \langle \delta_{1n_1}, \varsigma_{11} \rangle & \langle \delta_{1n_1}, \varsigma_{12} \rangle & \cdots & \langle \delta_{1n_1}, \varsigma_{1m_1} \rangle \end{bmatrix}, \quad B_1 = \begin{bmatrix} \langle \delta_{11}, \vartheta_{11} \rangle & \langle \delta_{11}, \vartheta_{12} \rangle & \cdots & \langle \delta_{11}, \vartheta_{1n_1} \rangle \\ \langle \delta_{12}, \vartheta_{11} \rangle & \langle \delta_{12}, \vartheta_{12} \rangle & \cdots & \langle \delta_{12}, \vartheta_{1n_1} \rangle \\ \vdots & \vdots & \vdots & \vdots \\ \langle \delta_{1n_1}, \vartheta_{11} \rangle & \langle \delta_{1n_1}, \vartheta_{12} \rangle & \cdots & \langle \delta_{1n_1}, \vartheta_{1n_1} \rangle \end{bmatrix}.$$

In the similar process, we obtain: $\underset{\sim}{f}_2(t_2) = A_2\underset{\sim}{c}_2 + B_2\underset{\sim}{d}_2, \ldots, \underset{\sim}{f}_p(t_p) = A_p\underset{\sim}{c}_p + B_p\underset{\sim}{d}_p$. Therefore, the regression function $\underset{\sim}{f}(t)$ can be expressed as:

$$\underset{\sim}{f}(t) = \left(\underset{\sim}{f}_1(t_1), \underset{\sim}{f}_2(t_2), ..., \underset{\sim}{f}_p(t_p) \right)' = \left(A_1\underset{\sim}{c}_1, A_2\underset{\sim}{c}_2, ..., A_p\underset{\sim}{c}_p \right)' + \left(B_1\underset{\sim}{d}_1, B_2\underset{\sim}{d}_2, ..., B_p\underset{\sim}{d}_p \right)'$$
$$= diag(A_1, A_2, ..., A_p)(\underset{\sim}{c}_1, \underset{\sim}{c}_2, ..., \underset{\sim}{c}_p)' + diag(B_1, B_2, ..., B_p)(\underset{\sim}{d}_1, \underset{\sim}{d}_2, ..., \underset{\sim}{d}_p)' = A\underset{\sim}{c} + B\underset{\sim}{d} \,. \tag{8}$$

In equation (8), A is a (N $\times$M)-matrix, where $N = \sum_{k=1}^{p} n_k$, $M = \sum_{k=1}^{p} m_k$, and $\underset{\sim}{c}$ is a $(M \times 1)$-vector of parameters that are expressed as $A = diag\left(A_1, A_2, ..., A_p\right)$, and $\underset{\sim}{c} = \left(\underset{\sim}{c}_1', \underset{\sim}{c}_2', ..., \underset{\sim}{c}_p'\right)'$, respectively. Also, B is ($N\times N$)-matrix, and $\underset{\sim}{d}$ is a ($N\times 1$)-vector of parameters which are expressed as $B = diag\left(B_1, B_2, ..., B_p\right)'$, and $\underset{\sim}{d} = \left(\underset{\sim}{d}_1', \underset{\sim}{d}_2', ..., \underset{\sim}{d}_p'\right)'$, respectively. Therefore, we can write the multiresponse nonparametric regression model given in (1) or (3) as $\underset{\sim}{y} = A\underset{\sim}{c} + B\underset{\sim}{d} + \underset{\sim}{\varepsilon}$. To obtain the estimation of regression function, $\underset{\sim}{f}$, we use RKHS method by solving the following optimization:

$$\underset{\substack{f_k \in \mathfrak{I}_k \\ k=1,2,...,p}}{Min} \left\{ \left\| W^{\frac{1}{2}}(\underset{\sim}{\sigma}^2)\underset{\sim}{\varepsilon} \right\|^2 \right\} = \underset{\substack{f_k \in \mathfrak{I}_k \\ k=1,2,...,p}}{Min} \left\{ \left\| W^{\frac{1}{2}}(\underset{\sim}{\sigma}^2)(\underset{\sim}{y} - \underset{\sim}{f}) \right\|^2 \right\}, \tag{9}$$

with constraint $\int_{a_k}^{b_k} [f_k^{(m)}(t_k)]^2 \, dt_k < \gamma_k$, $\gamma_k \geq 0$. The solution to (9) is equivalent to solution to the penalized weighted least-square (PWLS):

$$\underset{\substack{f_k \in W_2^m[a_k, b_k] \\ k=1,2,...,p}}{Min} \left\{ N^{-1}(\underset{\sim}{y} - \underset{\sim}{f})'W(\underset{\sim}{\sigma}^2)(\underset{\sim}{y} - \underset{\sim}{f}) + \sum_{k=1}^{p} \lambda_k \int_{a_k}^{b_k} [f_k^{(m)}(t_k)]^2 \, dt_k \right\}, \tag{10}$$

where λ_k, $k = 1,2,...,p$ are smoothing parameters that control between goodness of fit represented by $N^{-1}(\underset{\sim}{y} - \underset{\sim}{f})'W(\underset{\sim}{\sigma}^2)(\underset{\sim}{y} - \underset{\sim}{f})$ and smoothness of the estimate measured by $\lambda_1 \int_{a_1}^{b_1} [f_1^{(m)}(t_1)]^2 \, dt_1 + ... + \lambda_p \int_{a_p}^{b_p} [f_p^{(m)}(t_p)]^2 \, dt_p$. To determine the solution to (10), we first decompose the penalty as follows:

$$\int_{a_1}^{b_1} [f_1^{(m)}(t_1)]^2 \, dt_1 = \left\| Pf_1 \right\|^2 = \langle Pf_1, Pf_1 \rangle = \langle \underset{\sim}{\vartheta}_1'\underset{\sim}{d}_1, \underset{\sim}{\vartheta}_1'\underset{\sim}{d}_1 \rangle = \underset{\sim}{d}_1'\langle \underset{\sim}{\vartheta}_1, \underset{\sim}{\vartheta}_1' \rangle \underset{\sim}{d}_1 = \underset{\sim}{d}_1' B_1 \underset{\sim}{d}_1 .$$

Therefore, the penalty in (10) is $\sum_{k=1}^{p} \lambda_k \int_{a_k}^{b_k} [f_k^{(m)}(t_k)]^2 \, dt_k \} = \underset{\sim}{d}'\lambda B\underset{\sim}{d}$ where $\lambda = diag(\lambda_1 I_{n_1}, ..., \lambda_p I_{n_p})$. So, the goodness of fit in (10) can be written as:

$$N^{-1}(\underset{\sim}{y} - \underset{\sim}{f})'W(\underset{\sim}{\sigma}^2)(\underset{\sim}{y} - \underset{\sim}{f}) = N^{-1}(\underset{\sim}{y} - A\underset{\sim}{c} - B\underset{\sim}{d})'W(\underset{\sim}{\sigma}^2)(\underset{\sim}{y} - A\underset{\sim}{c} - B\underset{\sim}{d}) .$$

By combining the goodness of fit and penalty, we obtain optimization PWLS:

$$\underset{\substack{\underset{\sim}{c} \in R^{pn} \\ \underset{\sim}{d} \in R^{pm}}}{Min} \left\{ Q(\underset{\sim}{c}, \underset{\sim}{d}) \right\} = \underset{\substack{f_k \in W_2^m[a_k, b_k] \\ k=1,2,...,p}}{Min} \{ N^{-1}(\underset{\sim}{y} - \underset{\sim}{f})'W(\underset{\sim}{\sigma}^2)(\underset{\sim}{y} - \underset{\sim}{f}) + \sum_{k=1}^{p} \lambda_k \int_{a_k}^{b_k} [f_k^{(m)}(t_k)]^2 \, dt_k \} . \tag{11}$$

Optimization PWLS in (11) gives $\hat{\underset{\sim}{c}} = [A'D^{-1}W(\underset{\sim}{\sigma}^2)A]^{-1} A'D^{-1}W(\underset{\sim}{\sigma}^2)\underset{\sim}{y}$ and $\hat{\underset{\sim}{d}} = D^{-1}W(\underset{\sim}{\sigma}^2)[I - A(A'D^{-1}W(\underset{\sim}{\sigma}^2)A)^{-1}$ $A'D^{-1}W(\underset{\sim}{\sigma}^2)]\underset{\sim}{y}$, where $D = W(\underset{\sim}{\sigma}^2)B + N\lambda I$. Finally, we get the estimated regression function of model given in (1):

$$\hat{\underset{\sim}{f}}_{\lambda} = \left(\hat{\underset{\sim}{f}}_{1,\lambda_1}, \hat{\underset{\sim}{f}}_{2,\lambda_2}, ..., \hat{\underset{\sim}{f}}_{p,\lambda_p} \right)' = A\hat{\underset{\sim}{c}} + B\hat{\underset{\sim}{d}} = H(\underset{\sim}{\lambda})\underset{\sim}{y} \tag{12}$$

where $H(\underset{\sim}{\lambda}) = A[A'D^{-1}W(\underset{\sim}{\sigma}^2)A]^{-1} A'D^{-1}W(\underset{\sim}{\sigma}^2) + BD^{-1}W(\underset{\sim}{\sigma}^2) \times [I - A(A'D^{-1}W(\underset{\sim}{\sigma}^2)A)^{-1} A'D^{-1}W(\underset{\sim}{\sigma}^2)]$.

3.3 Estimation of covariance matrix of random errors

Suppose that $[W(\underset{\sim}{\sigma}^2)]^{-1}$ represents a estimated covariance matrix of random errors. To obtain the estimated covariance matrix, we consider the data set (t_{ki}, y_{ki}), $k = 1,2,...,p$; $i = 1,2,...,n_k$ follows the model in (3). Assume that $\underset{\sim}{y} = (\underset{\sim}{y}_1, \underset{\sim}{y}_2, ..., \underset{\sim}{y}_p)'$ is random sample obtained from N-variates normally distributed population ($N = \sum_{k=1}^{p} n_k$) with mean $\underset{\sim}{f}$, and covariance $[W(\underset{\sim}{\sigma}^2)]^{-1}$. Based on this distribution, we get likelihood function as follows:

$$L(\underset{\sim}{f}, W(\underset{\sim}{\sigma}^2) \mid \underset{\sim}{y}) = \prod_{j=1}^{n} \left\{ \frac{1}{(2\pi)^{\frac{N}{2}} \left| [W(\underset{\sim}{\sigma}^2)]^{-1} \right|^{\frac{1}{2}}} \exp\left(-\frac{1}{2} (\underset{\sim}{y}_j - \underset{\sim}{f}_j)' [W(\underset{\sim}{\sigma}^2)] (\underset{\sim}{y}_j - \underset{\sim}{f}_j) \right) \right\}$$

Since $N = \sum_{k=1}^{p} n_k$ and $W(\underset{\sim}{\sigma}^2) = diag(W_1(\underset{\sim}{\sigma}_1^2), W_2(\underset{\sim}{\sigma}_2^2), ..., W_p(\underset{\sim}{\sigma}_p^2))$ then we have:

$$L(\underset{\sim}{f}, W(\underset{\sim}{\sigma}^2) \mid \underset{\sim}{y}) = \left(\frac{1}{(2\pi)^{\frac{nn_1}{2}} \left| [W_1(\underset{\sim}{\sigma}_1^2)]^{-1} \right|^{\frac{n}{2}}} \exp\left\{ -\frac{1}{2} \sum_{j=1}^{n} (\underset{\sim}{y}_{1j} - \underset{\sim}{f}_{1j})' W_1(\underset{\sim}{\sigma}_1^2) (\underset{\sim}{y}_{1j} - \underset{\sim}{f}_{1j}) \right\} \right) \times$$

$$\left(\frac{1}{(2\pi)^{\frac{nn_2}{2}} \left| [W_2(\underset{\sim}{\sigma}_2^2)]^{-1} \right|^{\frac{n}{2}}} \exp\left\{ -\frac{1}{2} \sum_{j=1}^{n} (\underset{\sim}{y}_{2j} - \underset{\sim}{f}_{2j})' W_2(\underset{\sim}{\sigma}_2^2) (\underset{\sim}{y}_{2j} - \underset{\sim}{f}_{2j}) \right\} \right) \times ... \times$$

$$\left(\frac{1}{(2\pi)^{\frac{nn_p}{2}} \left| [W_p(\underset{\sim}{\sigma}_p^2)]^{-1} \right|^{\frac{n}{2}}} \exp\left\{ -\frac{1}{2} \sum_{j=1}^{n} (\underset{\sim}{y}_{pj} - \underset{\sim}{f}_{pj})' W_p(\underset{\sim}{\sigma}_p^2) (\underset{\sim}{y}_{pj} - \underset{\sim}{f}_{pj}) \right\} \right).$$

Next, estimated covariance matrix of random error can be obtained by carrying out the following optimization:

$$L(\underset{\sim}{f}, W(\underset{\sim}{\sigma}^2) \mid \underset{\sim}{y}) = \left[\underset{W_1(\underset{\sim}{\sigma}_1^2)}{Max} \left(\frac{1}{(2\pi)^{\frac{nn_1}{2}} \left| [W_1(\underset{\sim}{\sigma}_1^2)]^{-1} \right|^{\frac{n}{2}}} \exp\left\{ -\frac{1}{2} \sum_{j=1}^{n} (\underset{\sim}{y}_{1j} - \underset{\sim}{f}_{1j})' W_1(\underset{\sim}{\sigma}_1^2) (\underset{\sim}{y}_{1j} - \underset{\sim}{f}_{1j}) \right\} \right) \right] \times$$

$$\left[\underset{W_2(\underset{\sim}{\sigma}_2^2)}{Max} \left(\frac{1}{(2\pi)^{\frac{nn_2}{2}} \left| [W_2(\underset{\sim}{\sigma}_2^2)]^{-1} \right|^{\frac{n}{2}}} \exp\left\{ -\frac{1}{2} \sum_{j=1}^{n} (\underset{\sim}{y}_{2j} - \underset{\sim}{f}_{2j})' W_2(\underset{\sim}{\sigma}_2^2) (\underset{\sim}{y}_{2j} - \underset{\sim}{f}_{2j}) \right\} \right) \right] \times ... \times$$

$$\left[\underset{W_p(\underset{\sim}{\sigma}_p^2)}{Max} \left(\frac{1}{(2\pi)^{\frac{nn_p}{2}} \left| [W_p(\underset{\sim}{\sigma}_p^2)]^{-1} \right|^{\frac{n}{2}}} \exp\left\{ -\frac{1}{2} \sum_{j=1}^{n} (\underset{\sim}{y}_{pj} - \underset{\sim}{f}_{pj})' W_p(\underset{\sim}{\sigma}_p^2) (\underset{\sim}{y}_{pj} - \underset{\sim}{f}_{pj}) \right\} \right) \right]. \quad (13)$$

Determining of value $L(\underset{\sim}{f}, W(\underset{\sim}{\sigma}^2) \mid \underset{\sim}{y})$ in (13) is equivalent to determining of maximum value of each component in (13) (Johnson & Wichern, 1982). It means that the maximum value of each component in (13) can be determined by giving the following equations:

$$[W_1(\hat{\underset{\sim}{\sigma}}_1^2)] = \frac{\hat{\underset{\sim}{\varepsilon}}_1 \hat{\underset{\sim}{\varepsilon}}_1'}{n} = \frac{(\underset{\sim}{y}_1 - \hat{\underset{\sim}{f}}_{1,\lambda_1})(\underset{\sim}{y}_1 - \hat{\underset{\sim}{f}}_{1,\lambda_1})'}{n}, \; ... \; ,$$

$$[W_p(\hat{\underset{\sim}{\sigma}}_p^2)] = \frac{\hat{\underset{\sim}{\varepsilon}}_p \hat{\underset{\sim}{\varepsilon}}_p'}{n} = \frac{(\underset{\sim}{y}_p - \hat{\underset{\sim}{f}}_{p,\lambda_p})(\underset{\sim}{y}_p - \hat{\underset{\sim}{f}}_{p,\lambda_p})'}{n}$$

where $\hat{\underset{\sim}{f}}_{k,\lambda}$ $(k = 1,2,..., p)$ is regression function estimator given in (12). Therefore, the maximum likelihood estimator for $[W(\underset{\sim}{\sigma}^2)]$ is given by:

$$[W(\hat{\underset{\sim}{\sigma}}^2)] = diag([W_1(\hat{\underset{\sim}{\sigma}}_1^2)],[W_2(\hat{\underset{\sim}{\sigma}}_2^2)],...,[W_p(\hat{\underset{\sim}{\sigma}}_p^2)])$$

$$= diag\left(\frac{(\underset{\sim}{y}_1 - \hat{\underset{\sim}{f}}_{1,\lambda_1})(\underset{\sim}{y}_1 - \hat{\underset{\sim}{f}}_{1,\lambda_1})'}{n}, \frac{(\underset{\sim}{y}_2 - \hat{\underset{\sim}{f}}_{2,\lambda_2})(\underset{\sim}{y}_2 - \hat{\underset{\sim}{f}}_{2,\lambda_2})'}{n},..., \frac{(\underset{\sim}{y}_p - \hat{\underset{\sim}{f}}_{p,\lambda_p})(\underset{\sim}{y}_p - \hat{\underset{\sim}{f}}_{p,\lambda_p})'}{n} \right)$$

Since $\hat{\underset{\sim}{f}}_{k,\lambda_k} = H\left(\lambda_k; \hat{\underset{\sim}{\sigma}}_k^2\right)\underset{\sim}{y}_k$, $k = 1,2,..., p$ then for $k = 1, 2,..., p$ we have:

$$\frac{(\underset{\sim}{y}_1 - \hat{\underset{\sim}{f}}_{1,\lambda_1})(\underset{\sim}{y}_1 - \hat{\underset{\sim}{f}}_{1,\lambda_1})'}{n} = \frac{(I_{n_1} - H(\lambda_1; \hat{\underset{\sim}{\sigma}}_1^2))\underset{\sim}{y}_1 \underset{\sim}{y}_1'(I_{n_1} - H(\lambda_1; \hat{\underset{\sim}{\sigma}}_1^2))'}{n} \text{ , ... ,}$$

$$\frac{(\underset{\sim}{y}_p - \hat{\underset{\sim}{f}}_{p,\lambda_p})(\underset{\sim}{y}_p - \hat{\underset{\sim}{f}}_{p,\lambda_p})'}{n} = \frac{(I_{n_p} - H(\lambda_p; \hat{\underset{\sim}{\sigma}}_p^2))\underset{\sim}{y}_p \underset{\sim}{y}_p'(I_{n_p} - H(\lambda_p; \hat{\underset{\sim}{\sigma}}_p^2))'}{n}.$$

Therefore, the maximum likelihood estimator for covariance matrix is given by:

$$[W(\hat{\underset{\sim}{\sigma}}^2)]^{-1} = diag\left(\left[\frac{(\underset{\sim}{y}_1 - \hat{\underset{\sim}{f}}_{1,\lambda_1})(\underset{\sim}{y}_1 - \hat{\underset{\sim}{f}}_{1,\lambda_1})'}{n} \right]^{-1},..., \left[\frac{(\underset{\sim}{y}_p - \hat{\underset{\sim}{f}}_{p,\lambda_p})(\underset{\sim}{y}_p - \hat{\underset{\sim}{f}}_{p,\lambda_p})'}{n} \right]^{-1} \right)$$

$$= diag([W_1(\hat{\underset{\sim}{\sigma}}_1^2)]^{-1},[W_2(\hat{\underset{\sim}{\sigma}}_2^2)]^{-1},...,[W_p(\hat{\underset{\sim}{\sigma}}_p^2)]^{-1}).$$

where:

$$[W_1(\hat{\underset{\sim}{\sigma}}_1^2)] = \frac{[I_{n_1} - H(\lambda_1; \hat{\underset{\sim}{\sigma}}_1^2)]\underset{\sim}{y}_1 \underset{\sim}{y}_1'[I_n - H(\lambda_1; \hat{\underset{\sim}{\sigma}}_1^2)]'}{n}, \text{ ... , } [W_p(\hat{\underset{\sim}{\sigma}}_p^2)] = \frac{[I_{n_p} - H(\lambda_p; \hat{\underset{\sim}{\sigma}}_p^2)]\underset{\sim}{y}_p \underset{\sim}{y}_p'[I_{n_p} - H(\lambda_p; \hat{\underset{\sim}{\sigma}}_p^2)]'}{n}.$$

3.4 Estimation of optimum smoothing parameter

Selecting an optimum (suitable) smoothing parameter value λ is crucial to good regression function fitting. There are a number of ways to choose λ, including minimizing Mallows's C_p , cross-validation (CV) score, generalized cross-validation (GCV) score, and Akaike's information criterion (AIC) (Li & Zhang, 2010). Ruppert and Carrol (1997) pointed out that Mallows's C_p and GCV were satisfactory for good regression function fitting based on spline estimator. Moreover, GCV approximating CV is a computationally expedient criterion, so it is popular in spline literature.

In this section we establish the estimation of optimum smoothing parameter to good function regression fitting. For this goal, we may express the estimated regression function given in (12) as follows:

$$\hat{\underset{\sim}{f}}_\lambda(t) = H(\lambda_1, \lambda_2,..., \lambda_p; \underset{\sim}{\sigma}^2)\underset{\sim}{y} \tag{14}$$

where $\underset{\sim}{\sigma}^2 = (\sigma_1^2, \sigma_2^2,..., \sigma_p^2)'$. The mean square error (MSE) of the smoothing spline estimator given in (14) is:

$$MSE(\lambda_1, \lambda_2,..., \lambda_p; \underset{\sim}{\sigma}^2) = \frac{(\underset{\sim}{y} - \hat{\underset{\sim}{f}}_\lambda(t))'W(\underset{\sim}{\sigma}^2)(\underset{\sim}{y} - \hat{\underset{\sim}{f}}_\lambda(t))}{\sum_{k=1}^{p} n_k}$$

$$= \frac{[(I - H(\lambda_1, \lambda_2, ..., \lambda_p; \underset{\sim}{\sigma}^2))\underset{\sim}{y}]' W(\underset{\sim}{\sigma}^2)[(I - H(\lambda_1, \lambda_2, ..., \lambda_p; \underset{\sim}{\sigma}^2))\underset{\sim}{y}]}{\sum_{k=1}^{p} n_k}$$

$$= \frac{\left\| [W(\underset{\sim}{\sigma}^2)]^{\frac{1}{2}}[(I - H(\lambda_1, \lambda_2, ..., \lambda_p; \underset{\sim}{\sigma}^2)]\underset{\sim}{y} \right\|^2}{\sum_{k=1}^{p} n_k}.$$

Next, we define the following quantity:

$$G(\lambda_1, \lambda_2, ..., \lambda_p; \underset{\sim}{\sigma}^2) = \frac{\left(\sum_{k=1}^{p} n_k\right)^{-1} \left\| [W(\underset{\sim}{\sigma}^2)]^{\frac{1}{2}}[(I - H(\lambda_1, \lambda_2, ..., \lambda_p; \underset{\sim}{\sigma}^2)]\underset{\sim}{y} \right\|^2}{\left[\left(\frac{1}{\sum_{k=1}^{p} n_k}\right) trace\left(I_{\sum_{k=1}^{p} n_k} - H(\lambda_1, \lambda_2, ..., \lambda_p; \underset{\sim}{\sigma}^2)\right)\right]^2} \tag{15}$$

Therefore, based on (15), the optimum smoothing parameter, $\underset{\sim}{\lambda}_{opt} = (\lambda_{1(opt)}, \lambda_{2(opt)}, ..., \lambda_{p(opt)})'$, is obtained by solving the following optimization:

$$G_{opt}(\lambda_{1(opt)}, \lambda_{2(opt)}, ..., \lambda_{p(opt)}; \underset{\sim}{\sigma}^2) = \underset{\lambda_1 \in R^+, \lambda_2 \in R^+, ..., \lambda_p \in R^+}{Min} \left\{ G(\lambda_1, \lambda_2, ..., \lambda_p; \underset{\sim}{\sigma}^2) \right\}$$

$$= \underset{\lambda_1 \in R^+, \lambda_2 \in R^+, ..., \lambda_p \in R^+}{Min} \left\{ \frac{\left(\sum_{k=1}^{p} n_k\right)^{-1} \left\| [W(\underset{\sim}{\sigma}^2)]^{\frac{1}{2}}[(I - H(\lambda_1, \lambda_2, ..., \lambda_p; \underset{\sim}{\sigma}^2)]\underset{\sim}{y} \right\|^2}{\left[\left(\frac{1}{\sum_{k=1}^{p} n_k}\right) trace\left(I_{\sum_{k=1}^{p} n_k} - H(\lambda_1, \lambda_2, ..., \lambda_p; \underset{\sim}{\sigma}^2)\right)\right]^2} \right\},$$

where norm $\|\underset{\sim}{v}\| = \sqrt{v_1^2 + v_2^2 + ... + v_p^2}$ for a p-dimension vector $\underset{\sim}{v} = (v_1, v_2, ..., v_p)'$.

3.5 Asymptotic properties of spline estimator

In this section, we investigate the assymptotic properties of spline estimator $\hat{\underset{\sim}{f}}_\lambda$ of the multiresponse nonparametric regression model. Our goal is to investigate asymptotic properties of spline estimator $\hat{\underset{\sim}{f}}_\lambda$ and we will judge the quantity of spline estimator $\hat{\underset{\sim}{f}}_\lambda$ by weighted integrated mean square error (IMSE).

Firstly, we decompose $IMSE(\underset{\sim}{\lambda})$ into two components as follows:

$$IMSE(\underset{\sim}{\lambda}) = E\int_a^b [(\underset{\sim}{f}(\underset{\sim}{t}) - \hat{\underset{\sim}{f}}_\lambda(\underset{\sim}{t}))' W(\underset{\sim}{\sigma}^2)(\underset{\sim}{f}(\underset{\sim}{t}) - \hat{\underset{\sim}{f}}_\lambda(\underset{\sim}{t}))] d\underset{\sim}{t} = bias^2(\underset{\sim}{\lambda}) + Var(\underset{\sim}{\lambda})$$

where $bias^2 = \int_a^b E[(\underset{\sim}{f}(t) - E(\hat{\underset{\sim}{f}}_\lambda(t))' W(\underset{\sim}{\sigma}^2)(\underset{\sim}{f}(t) - E(\hat{\underset{\sim}{f}}_\lambda(t))] d\underset{\sim}{t}$, and $Var(\underset{\sim}{\lambda}) = \int_a^b E[(E\hat{\underset{\sim}{f}}_\lambda(\underset{\sim}{t}) - \hat{\underset{\sim}{f}}_\lambda(\underset{\sim}{t}))' W(\underset{\sim}{\sigma}^2)(E\hat{\underset{\sim}{f}}_\lambda(\underset{\sim}{t}) - \hat{\underset{\sim}{f}}_\lambda(\underset{\sim}{t}))] d\underset{\sim}{t}$. Next, to investigate the asymptotic property of $bias^2(\underset{\sim}{\lambda})$, we first establish the solution of PWLS optimization in the folowing Theorem.

Theorem 3. If $\hat{\underset{\sim}{f}}_{\lambda}(\underset{\sim}{t})$ is solution which minimize the following PWLS:

$$(\sum_{k=1}^{p} n_k)^{-1}(\underset{\sim}{y}-\underset{\sim}{g}(\underset{\sim}{t}))'W(\underset{\sim}{\sigma}^2)(\underset{\sim}{y}-\underset{\sim}{g}(\underset{\sim}{t}))+\sum_{k=1}^{p}\lambda_k\int_a^b\left(g_k^{(m)}(t_k)\right)^2 dt_k$$

then the solution which minimize the following PWLS:

$$(\sum_{k=1}^{p} n_k)^{-1}(\underset{\sim}{f}(\underset{\sim}{t})-\underset{\sim}{g}(\underset{\sim}{t}))'W(\underset{\sim}{\sigma}^2)(\underset{\sim}{f}(\underset{\sim}{t})-\underset{\sim}{g}(\underset{\sim}{t}))+\sum_{k=1}^{p}\lambda_k\int_a^b\left(g_k^{(m)}(t_k)\right)^2 dt_k \text{ is } \hat{\underset{\sim}{g}}_{\lambda}^{*}(\underset{\sim}{t})=E\hat{\underset{\sim}{f}}_{\lambda}(\underset{\sim}{t}).$$

Proof. Theorem 2 has given the solution which minimize the following PWLS:

$$(\sum_{k=1}^{p} n_k)^{-1}(\underset{\sim}{y}-\underset{\sim}{g}(\underset{\sim}{t}))'W(\underset{\sim}{\sigma}^2)(\underset{\sim}{y}-\underset{\sim}{g}(\underset{\sim}{t}))+\sum_{k=1}^{p}\lambda_k\int_a^b\left(g_k^{(m)}(t_k)\right)^2 dt_k .$$

The solution is:

$$\hat{\underset{\sim}{f}}_{\lambda}(\underset{\sim}{t})=\{A[A'D^{-1}W(\underset{\sim}{\sigma}^2)A]^{-1}A'D^{-1}W(\underset{\sim}{\sigma}^2)+BD^{-1}W(\underset{\sim}{\sigma}^2)[I-A(A'D^{-1}W(\underset{\sim}{\sigma}^2)A)^{-1}A'D^{-1}W(\underset{\sim}{\sigma}^2)]\}\underset{\sim}{y}$$

Next, by taking $\underset{\sim}{f}(\underset{\sim}{t})=\underset{\sim}{y}$, then the value that minimize:

$$(\sum_{k=1}^{p} n_k)^{-1}(\underset{\sim}{f}(\underset{\sim}{t})-\underset{\sim}{g}(\underset{\sim}{t}))'W(\underset{\sim}{\sigma}^2)(\underset{\sim}{f}(\underset{\sim}{t})-\underset{\sim}{g}(\underset{\sim}{t}))+\sum_{k=1}^{p}\lambda_k\int_a^b\left(g_k^{(m)}(t_k)\right)^2 dt_k \text{ can be writen as:}$$

$$\hat{\underset{\sim}{g}}_{\lambda}^{*}(\underset{\sim}{t})=\{A[A'D^{-1}W(\underset{\sim}{\sigma}^2)A]^{-1}A'D^{-1}W(\underset{\sim}{\sigma}^2)+BD^{-1}W(\underset{\sim}{\sigma}^2)[I-A(A'D^{-1}W(\underset{\sim}{\sigma}^2)A)^{-1}A'D^{-1}W]\}\underset{\sim}{f}(\underset{\sim}{t})=E\hat{\underset{\sim}{f}}_{\lambda}(\underset{\sim}{t}).$$

Furthermore, we investigate the asymptotic property of $bias^2(\underset{\sim}{\lambda})$. For this goal, we first give the following assumption.

Assumption. (A0). For every k $(k=1,2,...,p)$, $t_{ki}=\dfrac{2i-1}{2n}$, $i=1,2,...,n.$

The asymptotic property of $bias^2(\underset{\sim}{\lambda})$ can be established in Theorem 4 under assumption (A0).

Theorem 4. If assumption (A0) hold, then $bias^2(\underset{\sim}{\lambda})\le O(\underset{\sim}{\lambda})$ as $n\to\infty$
where $O(\underset{\sim}{\lambda})$ represents "*big oh*" (Sen & Singer, 1993, and Wand & Jones, 1995).

Proof. Suppose $\hat{\underset{\sim}{g}}_{\lambda}(\underset{\sim}{t})$ is value which minimize:

$$\int_a^b(\underset{\sim}{f}(\underset{\sim}{t})-\underset{\sim}{g}(\underset{\sim}{t}))'W(\underset{\sim}{\sigma}^2)(\underset{\sim}{f}(\underset{\sim}{t})-\underset{\sim}{g}(\underset{\sim}{t}))d\underset{\sim}{t}+\sum_{k=1}^{p}\lambda_k\int_a^b[g_k^{(m)}(t_k)]^2 dt_k .$$

Since assumption (A0) then, as $n\to\infty$,

$$(\sum_{k=1}^{p} n_k)^{-1}(\underset{\sim}{f}(\underset{\sim}{t})-\underset{\sim}{g}(\underset{\sim}{t}))'W(\underset{\sim}{\sigma}^2)(\underset{\sim}{f}(\underset{\sim}{t})-\underset{\sim}{g}(\underset{\sim}{t}))\approx\int_a^b(\underset{\sim}{f}(\underset{\sim}{t})-\underset{\sim}{g}(\underset{\sim}{t}))'W(\underset{\sim}{\sigma}^2)(\underset{\sim}{f}(\underset{\sim}{t})-\underset{\sim}{g}(\underset{\sim}{t}))d\underset{\sim}{t}$$

It implies $\hat{\underset{\sim}{g}}_{\lambda}^{*}(\underset{\sim}{t})=E\hat{\underset{\sim}{f}}_{\lambda}(\underset{\sim}{t})\approx\hat{\underset{\sim}{g}}_{\lambda}(\underset{\sim}{t})$ (16)

So, for every $\underset{\sim}{g}\in W_2^m[a,b]$, we have $bias^2(\underset{\sim}{\lambda})=\int_a^b E(\underset{\sim}{f}(\underset{\sim}{t})-E\hat{\underset{\sim}{f}}(\underset{\sim}{t}))'W(\underset{\sim}{\sigma}^2)(\underset{\sim}{f}(\underset{\sim}{t})-E\hat{\underset{\sim}{f}}(\underset{\sim}{t}))d\underset{\sim}{t}$

$$\leq \int_a^b E(\underset{\sim}{f}(\underset{\sim}{t}) - E\hat{\underset{\sim}{f}}_\lambda(\underset{\sim}{t}))' W(\underset{\sim}{\sigma}^2)(\underset{\sim}{f}(\underset{\sim}{t}) - E\hat{\underset{\sim}{f}}_\lambda(\underset{\sim}{t}))d\underset{\sim}{t} + \sum_{k=1}^{p} \lambda_k \int_a^b \hat{\underset{\sim}{g}}_{k\lambda}^{(m)}(t_k)dt_k$$

By considering $E\hat{\underset{\sim}{f}}_\lambda(\underset{\sim}{t}) \approx \hat{\underset{\sim}{g}}_\lambda(\underset{\sim}{t})$ as given in (16), we obtain:

$$bias^2(\underset{\sim}{\lambda}) \leq \int_a^b E(\underset{\sim}{f}(\underset{\sim}{t}) - \hat{\underset{\sim}{g}}_\lambda(\underset{\sim}{t}))' W(\underset{\sim}{\sigma}^2)(\underset{\sim}{f}(\underset{\sim}{t}) - \hat{\underset{\sim}{g}}_\lambda(\underset{\sim}{t}))d\underset{\sim}{t} + \sum_{k=1}^{p} \lambda_k \int_a^b \hat{\underset{\sim}{g}}_{k\lambda}^{(m)}(t_k)dt_k \cdot$$

Thus, for every $\underset{\sim}{g} \in W_2^m[a,b]$, we have the following relationship:

$$bias^2(\underset{\sim}{\lambda}) \leq \int_a^b E(\underset{\sim}{f}(\underset{\sim}{t}) - \underset{\sim}{g}(\underset{\sim}{t}))' W(\underset{\sim}{\sigma}^2)(\underset{\sim}{f}(\underset{\sim}{t}) - \underset{\sim}{g}(\underset{\sim}{t}))d\underset{\sim}{t} + \sum_{k=1}^{p} \lambda_k \int_a^b g_k^{(m)}(t_k)dt_k \cdot \tag{17}$$

Since (17) hold for every $\underset{\sim}{g} \in W_2^m[a,b]$, then by taking $\underset{\sim}{g}(\underset{\sim}{t}) = \underset{\sim}{f}(\underset{\sim}{t})$ we get:

$$bias^2(\underset{\sim}{\lambda}) \leq \sum_{k=1}^{p} \lambda_k \int_a^b [g_k^{(m)}(t_k)]^2 dt_k = O(\underset{\sim}{\lambda}), \quad \text{as } n \to \infty.$$

Next, to determine the asymptotic property of $Var(\underset{\sim}{\lambda})$ we first define:

$$\Phi(\hat{\underset{\sim}{f}}_\lambda, h, \gamma) = R(\hat{\underset{\sim}{f}}_\lambda + \gamma h) + \sum_{k=1}^{p} \lambda_k J_k(\hat{\underset{\sim}{f}}_\lambda + \gamma h), \quad \gamma \in R, \text{ and } h \in W_2^m[a,b]$$

where $R(\underset{\sim}{g}) = (\sum_{k=1}^{p} n_k)^{-1}(\underset{\sim}{y} - \underset{\sim}{g}(\underset{\sim}{t}))' W(\underset{\sim}{\sigma}^2)(\underset{\sim}{y} - \underset{\sim}{g}(\underset{\sim}{t}))$, and $J_k(\underset{\sim}{g}) = \int_a^b [g_k^{(m)}(t_k)]^2 dt_k$.

For any $\underset{\sim}{f}, \underset{\sim}{g} \in W_2^m[a,b]$, we have

$$\Phi(\underset{\sim}{f}, \underset{\sim}{g}, \gamma) = (\sum_{k=1}^{p} n_k)^{-1}(\underset{\sim}{y} - \underset{\sim}{f}(\underset{\sim}{t}) - \gamma \underset{\sim}{g}(\underset{\sim}{t}))' W(\underset{\sim}{\sigma}^2)(\underset{\sim}{y} - \underset{\sim}{f}(\underset{\sim}{t}) - \gamma \underset{\sim}{g}(\underset{\sim}{t})) + \sum_{k=1}^{p} \lambda_k \int_a^b (f_k^{(m)}(t_k) + \gamma g_k^{(m)}(t_k))^2 dt_k$$

By carrying out $d\Phi(\underset{\sim}{f}, \underset{\sim}{g}, \gamma)/d\gamma = 0$ and $\gamma = 0$, it will give:

$$(\sum_{k=1}^{p} n_k)^{-1}(\underset{\sim}{y} - \underset{\sim}{f}(\underset{\sim}{t}))' W(\underset{\sim}{\sigma}^2)\underset{\sim}{g}(\underset{\sim}{t}) = \sum_{k=1}^{p} \lambda_k \int_a^b f_k^{(m)}(t_k) g_k^{(m)}(t_k)dt_k \cdot$$

Suppose $\varphi_1, \varphi_2, ..., \varphi_n$ are basis of natural splines and $f(t) = \sum_{j=1}^{n} \beta_j \varphi_j(t)$, then according to Eubank (1988), it will give:

$$\sum_{i=0}^{n} W_i(\underset{\sim}{\sigma}^2) g(t_i)(y_i - \sum_{j=1}^{n} \beta_j \varphi_j(t_i)) = n\lambda(-1)^m (2m-1)! \sum_{i=1}^{n} g(t_i) \sum_{j=1}^{n} \beta_j d_{ij} \cdot \tag{18}$$

Since (18) is hold for every $g \in W_2^m[a,b]$, then equation (23) is equivalent to determine β_j which helds:

$$y_i = \sum_{j=1}^{n} (n\lambda(-1)^m (2m-1)! W_i^{-1}(\underset{\sim}{\sigma}^2) d_{ij} + \varphi_j(t_i))\beta_j, \quad i = 1, 2, ..., n. \tag{19}$$

We can express (19) in the following matrix notation:

$$\underset{\sim}{y} = (n\lambda(-1)^m(2m-1)!W^{-1}(\underset{\sim}{\sigma}^2)K+\varphi)\underset{\sim}{\beta}$$

where $K=\{d_{ij}\}$, $i,j=1,2,...,n$, and $\varphi=\{\varphi_j(t_i)\}$, $i,j=1,2,...,n$. So, we obtain:

$$\underset{\sim}{y} = (n\lambda W^{-1}(\underset{\sim}{\sigma}^2)FB^{-1}F'W^{-1}(\underset{\sim}{\sigma}^2)\varphi+\varphi)\underset{\sim}{\beta}\ \text{, where } B=F'VF. \tag{20}$$

If we multiply equation (20) from the leaft side with φ', then we will have:

$$\varphi'\underset{\sim}{y} = (n\lambda\varphi' W^{-1}(\underset{\sim}{\sigma}^2)FB^{-1}F'W^{-1}(\underset{\sim}{\sigma}^2)\varphi+\varphi'\varphi)\underset{\sim}{\beta}.$$

So, we get the estimator $\hat{\underset{\sim}{\beta}}_\lambda$ of $\underset{\sim}{\beta}$ as follows:

$$\hat{\underset{\sim}{\beta}}_\lambda = diag\left((1+n\lambda\theta_1)^{-1},...,(1+n\lambda\theta_n)^{-1}\right)\varphi'\underset{\sim}{y}.$$

Thus, we can express the estimator $\hat{\underset{\sim}{f}}_\lambda(\underset{\sim}{t})$ as follows:

$$\hat{\underset{\sim}{f}}_\lambda(\underset{\sim}{t}) = \varphi(\underset{\sim}{t})\hat{\underset{\sim}{\beta}}_\lambda = \sum_{j=1}^{n}\frac{1}{1+n\lambda\theta_j}\varphi_j'\underset{\sim}{y}\varphi_j(\underset{\sim}{t}) \tag{21}$$

The asymptotic properties of $Var(\underset{\sim}{\lambda})$ can be established in Theorem 5 under assumption (A0).

Theorem 5. If assumption (A0) hold, then for $k=1,2,...,p$, $Var(\lambda_k)\le O\left(\frac{1}{n\lambda_k^{\frac{1}{2m}}}\right)$, as $n\to\infty$.

Proof. For $k=1,2,...,p$, the equation (21) gives:

$$\hat{f}_{\lambda_k}(\underset{\sim}{t}) = \sum_{j=1}^{n}\frac{1}{1+n\lambda_k\theta_j}\varphi_j'\underset{\sim}{y}\varphi_j(\underset{\sim}{t})$$

It implies: $Var(\hat{f}_{\lambda_k}(\underset{\sim}{t}))\le\left(\underset{1\le i\le n}{Max}\{W_i^{-1}(\underset{\sim}{\sigma}^2)\}\right)\sigma^2\sum_{j=1}^{n}\frac{\varphi_j^2(\underset{\sim}{t})}{(1+n\lambda_k\theta_j)^2}$

So, we have: $Var(\lambda_k)\le\left(\underset{1\le i\le n}{Max}\{W_i^{-1}(\underset{\sim}{\sigma}^2)\}\right)\sigma^2\sum_{j=1}^{n}\frac{\varphi_j^2(\underset{\sim}{t})}{(1+n\lambda_k\theta_j)^2}\int_a^b\varphi_j^2(\underset{\sim}{t})W(\underset{\sim}{\sigma}^2)d\underset{\sim}{t}.$

Speckman (1985), and Eubank(1988) have given the following approximation:

$$n^{-1}=n^{-1}\sum_{r=1}^{p}W_r(\underset{\sim}{\sigma}^2)\varphi_j^2(t_r)\approx\int_a^b\varphi_k^2(\underset{\sim}{t})W(\underset{\sim}{\sigma}^2)d\underset{\sim}{t}\ \text{ and }\ Var(\lambda_k)\le\left(\underset{1\le i\le n}{Max}\{W_i^{-1}(\underset{\sim}{\sigma}^2)\}\right)\sigma^2 n^{-1}\sum_{j=1}^{n}\frac{1}{(1+\lambda_k\gamma_j)^2}\ \text{ as } n\to\infty.$$

Furthermore, Eubank (1988) has given:

$$Var(\lambda_k)\le\left(\underset{1\le i\le n}{Max}\{W_i^{-1}(\underset{\sim}{\sigma}^2)\}\right)\sigma^2 n^{-1}\sum_{j=1}^{n}\frac{1}{[(1+\lambda_k(\pi j)^{2m}]^2}$$

By integral approximation, we obtain:

$$Var(\lambda_k) \le \left(\underset{1\le i\le n}{Max}\{W_i^{-1}(\underset{\sim}{\sigma}^2)\} \right) \frac{\sigma^2}{\pi n \lambda_k^{\frac{1}{2m}}} \int_a^b \frac{dt}{(1+t^{2m})^2} = \frac{1}{n\lambda_k^{\frac{1}{2m}}} K(m,\sigma) = O\left(\frac{1}{n\lambda_k^{\frac{1}{2m}}} \right)$$

where $K(m,\sigma) = \frac{\sigma^2}{\pi} \left(\underset{1\le i\le n}{Max}\{W_i^{-1}(\underset{\sim}{\sigma}^2)\} \right) \int_a^b \frac{dt}{(1+t^{2m})^2}$.

Finally, based on Theorem 4 and Theorem 5, we obtain the asymptotic property of spline estimator based on IMSE criterion as follows:

$$IMSE(\underset{\sim}{\lambda}) = bias^2(\underset{\sim}{\lambda}) + Var(\underset{\sim}{\lambda}) \le O(\underset{\sim}{\lambda}) + O(\underset{\sim}{s}) \tag{22}$$

where $\underset{\sim}{s} = \left(\frac{1}{n\lambda_1^{\frac{1}{2m}}} \quad \frac{1}{n\lambda_2^{\frac{1}{2m}}} \quad \cdots \quad \frac{1}{n\lambda_p^{\frac{1}{2m}}} \right)'$.

3.6 Numerical example

In this section we give a numerical example of estimation of the multiresponse nonparametric regression model based on smoothing spline method where performance of this method depends on the selection of smoothing parameters. For this example, we generate data for $n=100$, correlations $\rho_{12}=0.6$, $\rho_{13}=0.7$, $\rho_{23}=0.8$, and variances $\sigma_1^2=2$, $\sigma_2^2=3$, $\sigma_3^2=4$. The underlying multiresponse nonparametric regression model is:

$$\left. \begin{aligned} y_{1i} &= 5+3\sin(2\pi t_{1i}^2)+\varepsilon_{1i}, \quad i=1,2,...n \\ y_{2i} &= 3+3\sin(2\pi t_{2i}^2)+\varepsilon_{2i}, \quad i=1,2,...n \\ y_{3i} &= 1+3\sin(2\pi t_{3i}^2)+\varepsilon_{3i}, \quad i=1,2,...n \end{aligned} \right\}. \tag{23}$$

In this example, we conduct simulation to compare three different smoothing parameters (λ), i.e., $\lambda = 1e-09$ (small lambda), $\lambda = 2.27e-07$ (optimum lambda), and $\lambda = 1e-05$ (great lambda). A plot of GCV versus lambda (λ) that gives minimum GCV value at lambda, $\lambda_{optimum} = 2.27e-07$ is given in Figure 1.

Next, a plot of the estimated regression function of the model in (23) for optimum lambda, $\lambda_{optimum} = 2.27e-07$, is given in Figure 2. A plot of the estimated regression function of the model in (23) for small lambda, $\lambda = 1e-09$, is given in Figure 3. Also, a plot of the estimated regression function of the model in (23) for great lambda, $\lambda = 1e-05$, is given in Figure 4. Figures 2, 3, and 4 show that selection smoothing parameters, i.e., optimum lambda ($\lambda_{optimum} = 2.27e-07$), small lambda ($\lambda = 1e-09$), and great lambda ($\lambda = 1e-05$) give a good regression function fitting, a too rough regression function fitting, and a too smooth regression function fitting, respectively.

In the following Table 1, we give the comparison of three different smoothing parameters (λ) in estimating the regression function of the multiresponse nonparametric regression model.

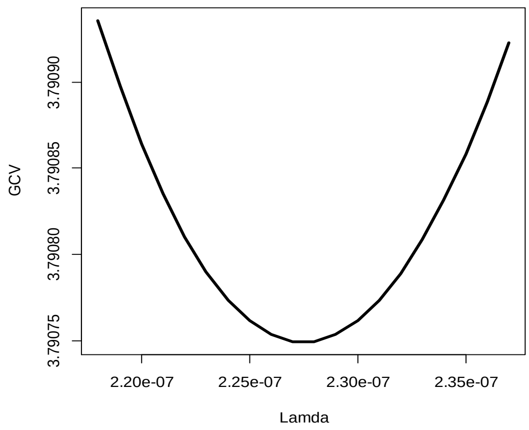

Figure 1. Plot of GCV versus lambda (λ) that gives minimum GCV value at lambda, $\lambda_{optimum} = 2.27e-07$.

Table 1 shows that the optimum smoothing parameter ($\lambda_{optimum} = 2.27e-07$) gives the most minimum GCV value compared with both small smoothing parameter ($\lambda = 1e-09$) and great smoothing parameter ($\lambda = 1e-05$). It means that the selection of optimum smoothing parameter gives the best regression function fitting of the multiresponse nonparametric regression model.

Table 1. Results of estimation for three different smoothing parameters.

Smoothing parameters	Minimum GCV values	Results of estimation
$\lambda = 1e-09$ (small lambda)	5.687333	A too rought regression function fitting.
$\lambda = 2.27e-07$ (optimum lambda)	3.79075	A good regression function fitting.
$\lambda = 1e-05$ (great lambda)	5.470145	A too smooth regression function fitting.

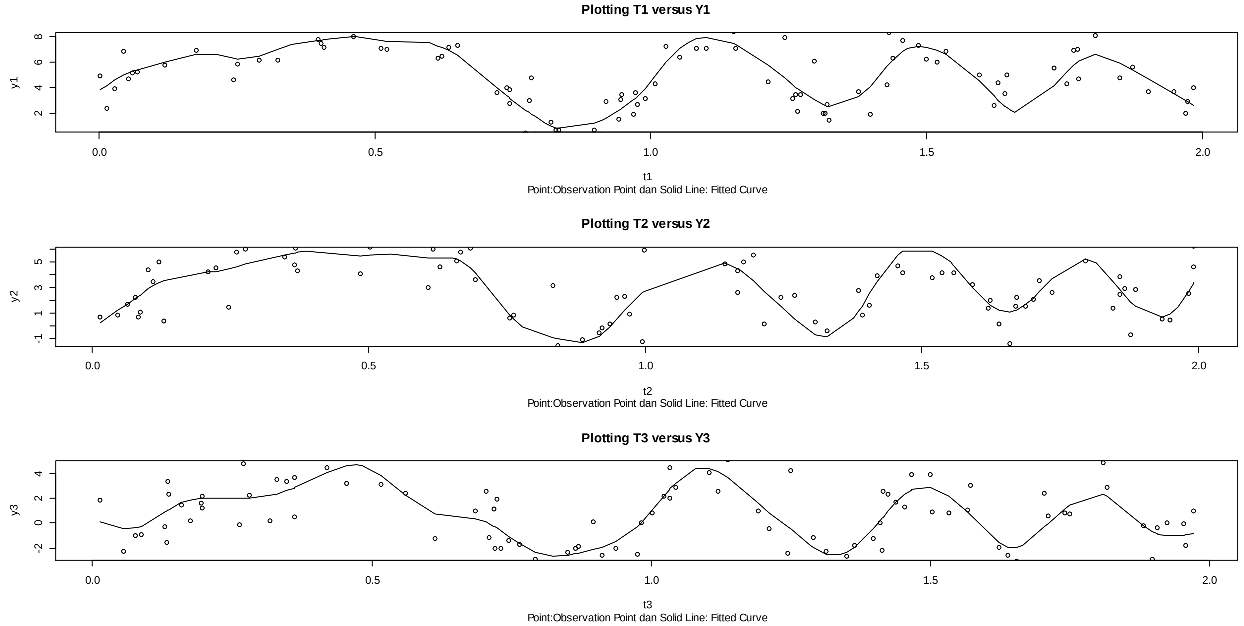

Figure 2. Plots of estimation of the first response (above), the second response (central), and the third response (below) of model (23) for optimum lambda, $\lambda_{optimum} = 2.27e-07$.

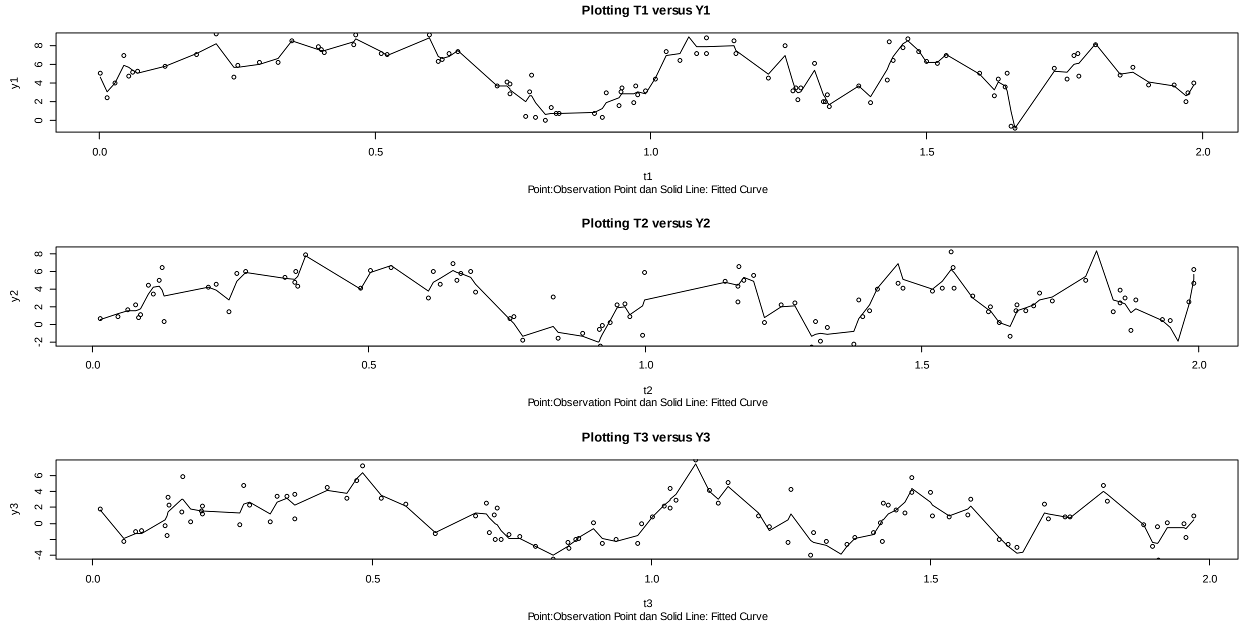

Figure 3. Plots of estimation of the first response (above), the second response (central), and the third response (below) of model (23) for small lambda, $\lambda = 1e-09$.

4. Conclusion

By assuming that the distribution of response variable is known, we can get the estimation of covariance matrix by using maximum likelihood method. The optimum smoothing parameter (λ) is obtained by minimizing the generalized cross validation (GCV). Optimum lambda (λ) gives a good regression function fitting. In estimating of the regression function of the multiresponse nonparametric regression model based on smoothing spline estimator, we use all observation points as knots. Smoothing spline estimate of the functions $\underset{\sim}{f}$ arises as a solution to the minimization problem, i.e., find $\underset{\sim}{\hat{f}}$ that minimizes the penalized weighted least-square (PWLS). The result shows that smoothing spline

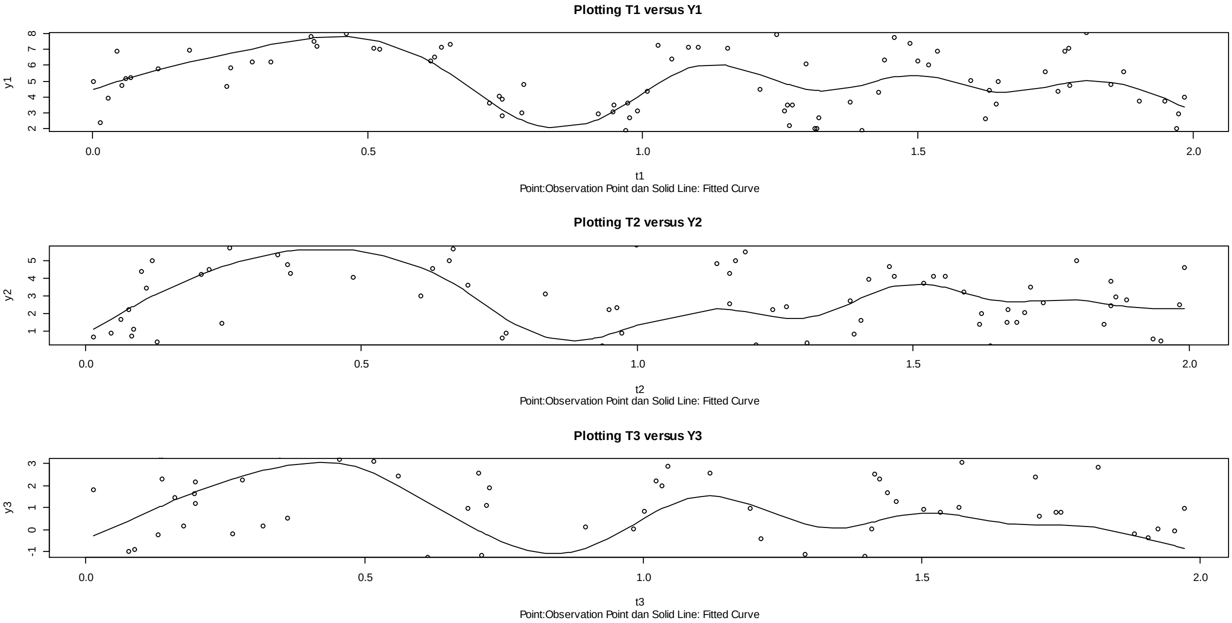

Figure 4. Plots of estimation of the first response (above), the second response (central), and the third response (below) of model (23) for great lambda, $\lambda = 1e-05$.

estimator, $\underset{\sim}{\hat{f}}_{\lambda}$, is an estimator which is linear in observations, and by taking expectation, $E(\underset{\sim}{\hat{f}}_{\lambda})$, it is a bias estimator for $\underset{\sim}{f}$. We can investigate the asymptotic properties of spline estimator ($\underset{\sim}{\hat{f}}_{\lambda}$) of regression functon ($\underset{\sim}{f}$) by decomposing $IMSE(\underset{\sim}{\lambda})$ into two components, i.e., $bias^2(\underset{\sim}{\lambda})$ and $Var(\underset{\sim}{\lambda})$, and the result has been given in (22).

References

Aydin, D. (2007). A comparison of the nonparametric regression models using smoothing spline and kernel regression. *Proceedings of World Academy of Science and Technology, 26*, 730-734.

Aydin, D., Memmedhi, M., & Omay, R. E. (2013). Smoothing parameter selection for nonparametric regression using smoothing spline. *European Journal of Pure and Applied Mathematics, 6*, 222-238.

Cardot, Crambes, H. C., Kneip, A., & Sarda, P. (2007). Smoothing splines estimators in functional linear regression with errors-in-variables. *Computational Statistics and Data Analysis, 51*, 4832-4848. doi:10.1016/j.csda.2006.07.029.

Carter, C. K., Eagleson, G. K., & Silverman, B. W. (1994). A comparison of reinch and speckman splines. *Biometrika, 79*, 81-91.

Chamidah, N., Budiantara, I. N., Sunaryo, S., & Zain, I. (2012). Designing of child growth chart based on multi-response local polynomial modeling. *Journal of Mathematics and Statistics, 8*(3), 342-347. doi:10.3844/jmssp.2012.342.347.

Chamidah, N., & Lestari, B. (2016). Spline estimator in homoscedastic multi-response nonparametric regression model in case of unbalanced number of observations. *Far East Journal of Mathematical Sciences (FJMS), 100*(9), 1433-1453. doi:10.17654/MS100091433.

Chen, X., & Christensen, T. M. (2015). Optimal uniform convergence rates and asymptotic normality for series estimators under weak dependence and weak conditions. *Journal of Econometric, 188*, 447-465.

Cox, D. D. (1983). Asymptotic for M-type smoothing spline. *Annals of Statistics, 11*, 530-551. doi:10.1214/aos/1176346159.

Cox, D. D., & O'Sullivan, F. (1996). Penalized likelihood type estimators for generalized nonparametric regression. *Journal of Multivariate Analysis, 56*, 185-206. doi:10.1006/jmva.1996.0010.

Craven, P., & Wahba, G. (1979). Smoothing noisy data with spline function: Estimating the correct degree of smoothing by the method of generalized cross validation. *Numerical Mathematics, 31*, 377-403.

Eggermont, P. P. B., Eubank, R. L., & LaRiccia, V. N. (2010). Convergence rates for smoothing spline estimators in varying coefficient models. *Journal of Statistical Planning and Inference, 140*, 369-381.

Eubank, R. L. (1988). *Spline smoothing and nonparametric regression*. New York, NY: Marcel Deker.

Fernandez, F. M., & Opsomer, J. D. (2005). Smoothing parameter selection methods for nonparametric regression with spatially correlated errors. *Canadian Journal of Statistics, 33*, 279-295. doi:10.1002/cjs.5550330208.

Flessler, J. A. (1991). Nonparametric fixed-interval smoothing with vector splines. *The IEEE Transactions on Signal Processing, 39*, 852–859. doi: 10.1109/78.80907.

Gooijer, J. G. D., Gannoun, A., & Larramendy, I. (1999). *Nonparametric regression with serially correlated errors* (1st ed). Rotterdam, The Netherlands: Tinbergen Institute.

Johnson, R. A., & Wichern, D. W. (1982). *Applied multivariate statistical analysis*. New York, NY: Prentice Hall.

Kimeldorf, G., & Wahba, G. (1971). Some result on tchebycheffian spline functions. *Journal of Mathematical Analysis and Application, 33*, 82-95.

Koenker, R., Pin, N. G., & Portnoy, S. (1994). Quantile smoothing splines. *Biometrics, 81*, 673-680. doi:10.1093/biomet/81.4.673.

Lee, T. C. M. (2004). Improved smoothing spline regression by combines estimates of different smoothness. *Statistics and Probability Letters, 67*, 133-140. doi:10.1016/j.spl.2004.01.003.

Lestari, B., Budiantara, I. N., Sunaryo, S., & Mashuri, M. (2009). Spline estimator in homoscedastic multiresponse nonparametric regression model. *Proceeding of IndoMS International Conference on Mathematics and Its Applications* (IICMA 2009), October 12th-13th 2009, 845-854, Yogyakarta-Indonesia,.

Lestari, B., Budiantara, I. N., Sunaryo, S., & Mashuri, M. (2010). Spline estimator in multiresponse nonparametric regression model with unequal correlation of errors. *Journal of Mathematics and Statistics, 6*(3), 327-332. doi:10.3844/jmssp.2010.327.332.

Lestari, B., Budiantara, I. N., Sunaryo, S., & Mashuri, M. (2012). Spline smoothing for multiresponse nonparametric regression model in case of heteroscedasticity of variance. *Journal of Mathematics and Statistics, 8*(3), 377-384. doi:10.3844/jmssp.2012.377.384.

Lestari, B., Fatmawati, & Budiantara, I. N. (2017). Estimasi fungsi regresi nonparametrik multirespon meng gunakan reproducing kernel Hilbert space berdasarkan estimator smoothing spline. *Proceeding of National Seminar on Mathematics and Its Applications* (SNMA2017), 243-250, Airlangga University, Surabaya-Indonesia.

Lestari, B., Fatmawati, Budiantara, I. N., & Chamidah, N. (2018). Estimation of regression function in multiresponse nonparametric regression model using smoothing spline and kernel estimators. *Journal of Physics: Conference Series, 1097*, 012091. doi:10.1088/1742-6596/1097/1/012091.

Lestari, B., Fatmawati, Budiantara, I. N., & Chamidah, N. (2019). Smoothing parameter selection method for multiresponse nonparametric regression model using smoothing spline and Kernel estimators approaches. *Journal of Physics: Conference Series, 1397*, 012064. doi:10.1088/1742-6596/1397/1/012064.

Li, J., & Zhang, R. (2010). Penalized spline varying-coefficient single-index model. *Communication in Statistics-Simulation and Computation, 39*, 221-239.

Liu, A., Tong, T., & Wang, Y. (2007). Smoothing spline estimation of variance functions. *Journal of Computational and Graphical Statistics, 16*, 312-329. doi:10.1198/106186007X204528.

Miller, J. J., & Wegman, E. J. (1987). Vector function estimation using splines. *Journal of Royal Statistical Society, Series B, 17*, 173-180. doi:10.1016/0378-3758(87)90110-8.

Oehlert, G. W. (1992). Relaxed boundary smoothing spline. *Annals of Statistics, 20*, 146-160. doi:10.1214/aos/1176348516.

Ping, T., & Lin, Y. (2013). Asymptotic properties of estimators in partially linear single-index model for longitudinal data. *Acta Mathematica Scientia, 30B* (3), 677-687.

Ruppert, D., & Carroll, R. (1997). Penalized regression splines. Working Paper, Cornell University, School of Operation Research and Industrial Engineering. Retrieved from www.orie.cornell.edu/davidr/papers.

Sen, P. K., & Singer, J. M. (1993). *Large sample methods in statistics: An introduction with applications*. London, England: Chapman and Hall.

Speckman, P. (1985). Spline smoothing and optimal rate of convergence in nonparametric regression models. *Annals of Statistics, 13*, 970-983.

Wahba, G. 1983. Bayesian confidence intervals for the cross-validated smoothing spline. *Journal of Royal Statistical Society, Series B., 45*, 133-150.

Wahba, G. (1985). A comparison of GCV and GML for choosing the smoothing parameter in the generalized spline smoothing problem. *Annals of Statistics, 13*, 1378-1402. doi:10.1214/aos/1176349743.

Wahba, G. (1992). Multivariate functional and operator estimation, based on smoothing spline and reproducing kernels. Retrieved from http://www.stat.wisc.edu/~wahba/ftp1/oldie/santafe1992.pdf.

Wahba, G. (2000). An introduction to model building with reproducing kernel hilbert spaces. Retrieved from http://www.stat.wisc.edu/~wahba/ftp1/interf/rootpar1r.pdf.

Wand, M. P. & Jones, M. C. (1995). Kernel smoothing. London, England: Chapman and Hall.

Wang, J-F. (2012). Asymptotic properties for an M-Estimator of the regression function with truncation and dependent data. *Journal of the Korean Statistical Society, 41*, 351-367.

Wang, Y. (1998). Smoothing spline models with correlated random errors. *Journal of American Statistical Association, 93*, 341-348. Retrieved from http://www.jstor.org/stable/2669630.

Wang, Y. (2011). *Smoothing splines: Methods and applications*. New York, NY: CRC Press.

Wang, Y., Guo, W., & Brown, W. B. (2000). Spline smoothing for bivariate data with applications to association between hormones. *Statistica Sinica, 10*, 377-397.

Wegman, E. J. (1981). Vector spline and estimation of filter function. *Technometrics, 23*, 83-89.