Journal of King Saud University – Science 35 (2023) 102664

Contents lists available at ScienceDirect

Journal of King Saud University – Science

journal homepage: www.sciencedirect.com

Original article

Determining confidence interval and asymptotic distribution for parameters of multiresponse semiparametric regression model using smoothing spline estimator

Budi Lestari [a,1], Nur Chamidah [b,1,*], I. Nyoman Budiantara [c,1], Dursun Aydin [d]

[a] Department of Mathematics, Faculty of Mathematics and Natural Sciences, The University of Jember, Jember 68121 Indonesia
[b] Department of Mathematics, Faculty of Science and Technology, Airlangga University, Surabaya 60115 Indonesia
[c] Department of Statistics, Faculty of Sciences and Data Analytics, Sepuluh Nopember Institute of Technology, Surabaya 60111 Indonesia
[d] Department of Statistics, Faculty of Science, Muğla Sıtkı Koçman University, Muğla 48000 Turkey

ARTICLE INFO

Article history:
Received 19 October 2021
Revised 24 March 2023
Accepted 29 March 2023
Available online 6 April 2023

Keywords:
Asymptotic distribution
Confidence interval
Nutritional status
Semiparametric regression
Smoothing spline

ABSTRACT

The multiresponse semiparametric regression (MSR) model is a regression model with more than two response variables that are mutually correlated, and its regression function is composed of parametric and nonparametric components. The study objectives are propose a new method for estimating the MSR model using smoothing spline. Also, find the confidence interval (CI) of parameters and the distribution asymptotically of the model parameters estimator. Methods used in this study are reproducing kernel Hilbert space (RKHS) method and a developed penalized weighted least squares (PWLS), and apply pivotal quantity, central limit theorem, and theorems of Cramer-Wold and Slutsky. The results are an 100 (1−α)% CI estimate and an asymptotic normal distribution for the parameters of the MSR model. In conclusion, the estimated MSR model is a combined components estimate of parametric and nonparametric which is linear to observation, and CIs of parameters depend on **t** distribution and estimator of parameters is asymptotically normally distributed. Future time, this study results can be used as theoretical bases to design standard growth charts of the toddlers which can then be used to assess the nutritional status of the toddlers.

1. Introduction

Regression models are widely applied to analyze functional association between response and predictor variables for prediction and interpretation purposes. Based on regression function shapes, the regression models consist of parametric regression (PR) and nonparametric regression (NR) models. The PR and NR models combination forms semiparametric regression (SR) models. The SR model will form MSR model when it has two or more variables of response that are mutually correlated.

In regression modeling, determining estimators of regression functions such as spline, kernel, PWLS, local linear, local polynomial, is main problem. Some estimators were used to estimate the regression functions, namely splines (Eubank, 1988; Wahba, 1990; Wang et al., 2000; Gu, 2002; Wang, 2011; Chamidah et al., 2019b; 2020a; Fatmawati et al., 2019; Khan & Shahna, 2019; Shahna & Khan, 2019; and Islamiyati et al., 2022;), kernel (Yilmaz et al., 2021), PWLS (Lestari et al., 2020; 2022), local linear (Chamidah et al., 2018; 2019c; 2020b), local polynomial (Chamidah et al., 2019a; Chamidah & Lestari, 2019). Next, both kernel and spline estimators in multiresponse NR (MNR) models and in NR model were discussed by Lestari et al. (2018; 2019) and Osmani et al. (2019), respectively. The estimators mentioned above except for the spline, are very dependent on the neighbors of the target point (bandwidth). Hence, if these estimators are applied to estimate fluctuated data model, we need small bandwidth and this will give the estimation curve too rough. These estimators only examine goodness of fit and not smoothness. Thus, these estimators are less reliable for estimating the fluctuated data models in the sub intervals, because these estimators will provide

* Corresponding author.
E-mail address: nur-c@fst.unair.ac.id (N. Chamidah).
[1] Research Group of Statistical Modeling in Life Science, Faculty of Science and Technology, Airlangga University, Surabaya 60115 Indonesia.

Peer review under responsibility of King Saud University.

Production and hosting by Elsevier

https://doi.org/10.1016/j.jksus.2023.102664
1018-3647/© 2023 The Author(s). Published by Elsevier B.V. on behalf of King Saud University.

estimation results with large mean square errors (MSE). This is different from the spline estimator which considers fit and smoothness factors. The ability of the spline estimators to estimate the MNR model for prediction purposes has been discussed by Fatmawati et al. (2019) and Lestari et al. (2020). Although there have been several previous studies discussing these estimators for estimating the regression function, these estimators were applied to NR and MNR models only. This means that previous researchers have not applied these estimators to estimate the uniresponse semiparametric regression (USR) model.

Furthermore, several estimators in USR models have been discussed by researchers namely splines (Gao & Shi, 1997; Wang & Ke, 2009; Diana et al., 2013; Mohaisen & Abdulhussein, 2015; Ramadan et al., 2019; Aydin et al., 2019; Chen & Ren, 2020; Fernandes et al., 2020; Chamidah et al., 2021), kernel (Yilmaz et al., 2021). While, Amini & Roozbeh (2015), Roozbeh (2018), and Roozbeh et al. (2020) estimated the restricted SR models using ridge, and selected optimal shrinkage parameter and kernel smoother bandwidth based on developed generalized cross validation (GCV) criterion. But, these previous researchers discussed estimators in USR models only. Although Wibowo et al. (2012) and Chamidah et al. (2022) estimated the MSR model using penalized spline and truncated spline, respectively, but these researchers have not yet applied smoothing spline to estimate MSR model regression function.

In this study we develop a estimation method for the MSR model, and determine the CI and asymptotic distribution of parameters estimator in the MSR model using smoothing spline. The smoothing spline can handle data with too smooth or too coarse character, and changes at certain sub-intervals. It considers both goodness of fit stated by WLS function and smoothness of model estimation stated by penalty function where balance between them are controlled by smoothing parameters. The smoothing spline becomes less practical when sample size n is large because it uses n knots. To overcome this practical problem, in this article we therefore provide asymptotic distribution determination of parameters estimator in MSR model.

2. Materials and Methods

Suppose a paired dataset $(y_{ki}, x_{ki1}, x_{ki2}, \cdots, x_{kiq_k}, t_{ki1}, t_{ki2}, \cdots, t_{kir_k})$, $k = 1, 2, \cdots, p$; $i = 1, 2, \cdots, n_k$; $q_k + r_k = n_k$ where relationship between $(x_{ki1}, x_{ki2}, \cdots, x_{kiq_k}, t_{ki1}, t_{ki2} \cdots, t_{kir_k})$ and y_{ki} meets the MSR model:

$$y_{ki} = f_k(x_{ki1}, x_{ki2}, \cdots, x_{kiq_k}) + g_k(t_{ki1}, t_{ki2} \cdots, t_{kir_k}) + \varepsilon_{ki} \quad (1)$$

where y_{ki} is value of i^{th} observation for k^{th} response, $f_k(x_{ki1}, x_{ki2}, \cdots, x_{kiq_k})$ is unknown function for k^{th} response, $g_k(t_{ki1}, t_{ki2} \cdots, t_{kir_k})$ is unknown smooth function for k^{th} response contained in Sobolev space $W_2^m[a_k, b_k]$, and ε_{ki} is random error with mean zero and variance σ_{ki}^2.

The MSR model regression function in (1) is composed of parametric function component namely $f_k(x_{ki1}, x_{ki2}, \cdots, x_{kiq_k})$, and nonparametric function components namely $g_k(t_{ki1}, t_{ki2} \cdots, t_{kir_k})$. So, we use WLS method to estimate $f_k(x_{ki1}, x_{ki2}, \cdots, x_{kiq_k})$, and use smoothing spline to estimate $g_k(t_{ki1}, t_{ki2} \cdots, t_{kir_k})$ by developing PWLS method proposed by Wang et al. (2000). Next, we apply pivotal quantity, central limit theorem, and theorems of Cramer-Wold and Slutsky to obtain CI and distribution asymptotically of the model parameters estimator of MSR model.

3. Results

Following are results of this study including regression function estimation, determination of CI parameter and asymptotic distribution for parameter estimator of MSR model.

3.1. Regression function estimation

We may present the MSR model (1) as follows:

$$y_{ki} - f_k(x_{ki1}, x_{ki2}, \cdots, x_{kiq_k}) = g_k(t_{ki1}, t_{ki2} \cdots, t_{kir_k}) + \varepsilon_{ki} \quad (2)$$

We can rewrite model (2) as follows:

$$y_{ki} - \mathbf{x}_{ki}^T \beta_k = g_k(\mathbf{t}_{ki}) + \varepsilon_{ki} \quad (3)$$

$$\mathbf{x}_{ki}^T \beta_k = f_k(x_{ki1}, x_{ki2}, \cdots, x_{kiq_k})$$

Suppose $\widehat{\boldsymbol{\beta}}_k$ is the true WLS estimate of $\boldsymbol{\beta}_k$. Hence, we can express model (3) as follows:

$$y_{ki}^* = g_k(\mathbf{t}_{ki}) + \varepsilon_{ki} \quad (4)$$

$$y_{ki}^* = y_{ki} - \mathbf{x}_{ki}^T \widehat{\boldsymbol{\beta}}_k \quad (5)$$

Next, let $\mathbf{y}^* = \left(\mathbf{y}_{1i}^*, \mathbf{y}_{2i}^*, \cdots, \mathbf{y}_{pi}^*\right)^T$; $\mathbf{g} = (\mathbf{g}_1, \mathbf{g}_2, \cdots, \mathbf{g}_p)^T$; $\varepsilon = (\varepsilon_{1i}, \varepsilon_{2i}, \cdots, \varepsilon_{pi})^T$; and $\mathbf{t} = (\mathbf{t}_{1i}, \mathbf{t}_{2i}, \cdots, \mathbf{t}_{pi})^T$ where $\mathbf{y}_{ki}^* = \left(y_{k1}^*, y_{k2}^*, \cdots, y_{kn_k}^*\right)^T$; $\varepsilon_{ki} = (\varepsilon_{k1}, \varepsilon_{k2}, \cdots, \varepsilon_{kn_k})^T$;
$\mathbf{t}_{ki} = (t_{k1}, t_{k2}, \cdots, t_{kn_k})^T$; and $k = 1, 2, \cdots, p$.

Hence, we can present the MSR model (4) in the following matrix equation:

$$\mathbf{y}^* = \mathbf{g} + \varepsilon \quad (6)$$

where $E(\varepsilon) = 0$, $Cov(\varepsilon) = \mathbf{W}^{-1}$ (namely).

The smoothing spline estimator of function $\mathbf{g}$ in model (6) can be determined by solving the PWLS:

$$\underset{\mathbf{g}_1,\cdots,\mathbf{g}_p \in W_2^m}{\text{Min}} \Big\{ N^{-1} \Big[(\mathbf{y}_1^* - \mathbf{g}_1)^T \mathbf{W}_1 (\mathbf{y}_1^* - \mathbf{g}_1) + (\mathbf{y}_2^* - \mathbf{g}_2)^T \mathbf{W}_2 (\mathbf{y}_2^* - \mathbf{g}_2) + \cdots + \left(\mathbf{y}_p^* - \mathbf{g}_p\right)^T \mathbf{W}_p \left(\mathbf{y}_p^* - \mathbf{g}_p\right) \Big] + \lambda_1 \int_{a_1}^{b_1} \left(\mathbf{g}_1^{(m)}(\mathbf{t}_1)\right)^2 d\mathbf{t}_1 + \lambda_2 \int_{a_2}^{b_2} \left(\mathbf{g}_2^{(m)}(\mathbf{t}_2)\right)^2 d\mathbf{t}_2 + \cdots + \lambda_p \int_{a_p}^{b_p} \left(\mathbf{g}_p^{(m)}(\mathbf{t}_p)\right)^2 d\mathbf{t}_p \Big\} \quad (7)$$

where $N = \sum_{k=1}^p n_k$; $\mathbf{W}_1, \mathbf{W}_2, \cdots, \mathbf{W}_p$ are weight matrices that are inverse of covariance matrix, $\mathbf{g} \in W_2^{\boldsymbol{m}}[a, b]$, and $\lambda_1, \lambda_2, \cdots, \lambda_p$ are smoothing parameters that set the balance between good fit and smoothness of estimation.

Based on Eq.(6), it is easy to show that the covariance matrix of random errors in MSR model (1) is:

$$\mathbf{W}^{-1} = diag\left(\mathbf{W}_1^{-1}, \mathbf{W}_2^{-1}, \cdots, \mathbf{W}_p^{-1}\right) \quad (8)$$

where $\mathbf{W}_k^{-1} = \begin{pmatrix} \sigma_{k1}^2 & \sigma_{k(1,2)} & \cdots & \sigma_{k(1,n_1)} \\ \sigma_{k(2,1)} & \sigma_{k2}^2 & \cdots & \sigma_{k(2,n_2)} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{k(n_k,1)} & \sigma_{k(n_k,2)} & \cdots & \sigma_{kn_k}^2 \end{pmatrix}$, and $k = 1, 2, \cdots, p$.

Solution to optimization PWLS in (7) is obtained by using RKHS method. We can read details of RKHS in Aronszajn (1950), Eubank (1988), Wahba (1990), Gu (2002), and Wang (2011). Firstly, we express the model (4) into general smoothing spline regression model (Wang, 2011):

$$y_{ki}^* = \mathscr{L}_{t_{ki}} g_k + \varepsilon_{ki} \quad (9)$$

where$i = 1, 2, \cdots, n_k$; $k = 1, 2, \cdots, p$; $g_k \in \mathscr{G}_k$ is a function which unknown and smooth contained in Hilbert space $\mathscr{G}_k$; and $\mathscr{L}_{t_{ki}} \in \mathscr{G}_k$ is a linear function and bounded.

Suppose we may decompose the Hilbert space $\mathscr{G}_k$ into direct sum of two subspaces $\mathscr{F}_k$ and $\mathscr{H}_k$ such that we have:

$$\mathscr{G}_k = \mathscr{F}_k \oplus \mathscr{H}_k \tag{10}$$

where $\mathscr{F}_k$ is orthogonal to $\mathscr{H}_k$. Hence, for every function $g_k \in \mathscr{G}_k$, $k = 1, 2, \cdots, p$ can be expressed as follows:

$g_k = f_k + h_k;\ f_k \in \mathscr{F}_k;\ h_k \in \mathscr{H}_k.$

Next, if $\{\delta_{k1}, \delta_{k2}, \cdots, \delta_{km_k}\}$ is basis of space $\mathscr{F}_k$ and $\{\xi_{k1}, \xi_{k2}, \cdots, \xi_{kn_k}\}$ is basis of space $\mathscr{H}_k$, then we can express every function $g_k \in \mathscr{G}_k$, $k = 1, 2, \cdots, p$ as follows:

$$g_k = \sum_{r=1}^{m_k} c_{kr}\delta_{kr} + \sum_{s=1}^{n_k} d_{ks}\xi_{ks} = \delta_k^T \mathbf{c}_k + \xi_k^T \mathbf{d}_k;\ c_{kr}, d_{ks} \in \mathbb{R} \tag{11}$$

where $\delta_k = (\delta_{k1}, \delta_{k2}, \cdots, \delta_{km_k})^T$; $\mathbf{c}_k = (c_{k1}, c_{k2}, \cdots, c_{km_k})^T$; $\xi_k = (\xi_{k1}, \xi_{k2}, \cdots, \xi_{km_k})^T$; and

$\mathbf{d}_k = (d_{k1}, d_{k2}, \cdots, d_{km_k})^T$

Hereinafter, since $\mathscr{L}_{t_{ki}} \in \mathscr{G}_k$ is bounded linear function and $g_k \in \mathscr{G}_k$, $k = 1, 2, \cdots, p$ then we have:

$$\mathscr{L}_{t_{ki}} g_k = \mathscr{L}_{t_{ki}}(f_k + h_k) = \mathscr{L}_{t_{ki}}(f_k) + \mathscr{L}_{t_{ki}}(h_k)$$

$$= f_k(t_{ki}) + h_k(t_{ki})$$

$$= g_k(t_{ki}) \tag{12}$$

Based on Eq. (12) and Riesz representation theorem (Wang, 2011), there is a representer $\omega_{ki} \in \mathscr{G}_k$ of $\mathscr{L}_{t_{ki}}$ such that:

$\mathscr{L}_{t_{ki}} g_k = \langle \omega_{ki}, g_k \rangle = g_k(t_{ki}); g_k \in \mathscr{G}_k$

where $\langle \cdot, \cdot \rangle$ notates a product of inner. By considering Eq.(11) and inner-product properties, the following equation is obtained:

$$g_k(t_{ki}) = \langle \omega_{ki}, \delta_k^T \mathbf{c}_k + \xi_k^T \mathbf{d}_k \rangle = \langle \omega_{ki}, \delta_k^T \mathbf{c}_k \rangle + \langle \omega_{ki}, \xi_k^T \mathbf{d}_k \rangle \tag{13}$$

Next, by using Eq. (13) for $k = 1$ we get:

$$g_1(t_{1i}) = \langle \omega_{1i}, \delta_1^T \mathbf{c}_1 \rangle + \langle \omega_{1i}, \xi_1^T \mathbf{d}_1 \rangle \tag{14}$$

Hence, based on Eq. (14) for $i = 1, 2, \cdots, n_1$ we have:

$$\mathbf{g}_1(\mathbf{t}_1) = (g_1(t_{11}), \cdots, g_1(t_{1n_1}))^T = \mathbf{A}_1\mathbf{c}_1 + \mathbf{B}_1\mathbf{d}_1$$

where $\mathbf{c}_1 = (c_{11}, c_{12}, \cdots, c_{1m_1})^T$; $\mathbf{d}_1 = (d_{11}, d_{12}, \cdots, d_{1n_1})^T$;

$$\mathbf{A}_1 = \begin{pmatrix} \langle \omega_{11}, \delta_{11} \rangle & \cdots & \langle \omega_{11}, \delta_{1m_1} \rangle \\ \vdots & \ddots & \vdots \\ \langle \omega_{1n_1}, \delta_{11} \rangle & \cdots & \langle \omega_{1n_1}, \delta_{1m_1} \rangle \end{pmatrix}; \quad \text{and}$$

$$\mathbf{B}_{(1)} = \begin{pmatrix} \langle \omega_{11}, \xi_{11} \rangle & \cdots & \langle \omega_{11}, \xi_{1n_1} \rangle \\ \vdots & \ddots & \vdots \\ \langle \omega_{1n_1}, \xi_{11} \rangle & \cdots & \langle \omega_{1n_1}, \xi_{1n_1} \rangle \end{pmatrix}.$$

Similarly, we get:

$\mathbf{g}_2(\mathbf{t}_2) = \mathbf{A}_2\mathbf{c}_2 + \mathbf{B}_2\mathbf{d}_2$, $\mathbf{g}_3(\mathbf{t}_3) = \mathbf{A}_3\mathbf{c}_3 + \mathbf{B}_3\mathbf{d}_3$ $\cdots$ $\mathbf{g}_p(\mathbf{t}_p) = \mathbf{A}_p\mathbf{c}_p + \mathbf{B}_p\mathbf{d}_p$.

Therefore, generally, the following expression of $\mathbf{g}(\mathbf{t})$ is obtained:

$$\mathbf{g}(\boldsymbol{t}) = (\mathbf{g}_1(\mathbf{t}_1), \mathbf{g}_2(\mathbf{t}_2), \cdots, \mathbf{g}_p(\mathbf{t}_p))^{\boldsymbol{T}}$$
$$= (\mathbf{A}_1\mathbf{c}_1, \mathbf{A}_2\mathbf{c}_2, \cdots, \mathbf{A}_p\mathbf{c}_p)^T + (\mathbf{B}_1\mathbf{d}_1, \mathbf{B}_2\mathbf{d}_2, \cdots, \mathbf{B}_p\mathbf{d}_p)^T$$

$$= diag(\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_p)\left(\mathbf{c}_1^T, \mathbf{c}_2^T, \cdots, \mathbf{c}_p^T\right)^T$$
$$+ diag(\mathbf{B}_1, \mathbf{B}_2, \cdots, \mathbf{B}_p)\left(\mathbf{d}_1^T, \mathbf{d}_2^T, \cdots, \mathbf{d}_p^T\right)^T$$

$$= \mathbf{Ac} + \mathbf{Bd} \tag{15}$$

where $\mathrm{N} = \sum_{k=1}^{p} n_k$; $\mathrm{M} = \sum_{k=1}^{p} m_k$; $\mathbf{A}$ is a matrix with dimension $\mathrm{N} \times \mathrm{M}$; $\mathbf{c}$ is a vector with dimension $\mathrm{M} \times 1$; $\mathbf{B}$ is a matrix with dimension $\mathrm{N} \times \mathrm{N}$; and $\mathbf{d}$ is a vector with dimension $\mathrm{N} \times 1$. Generally, based on Eq.(15), the MSR model (6) can be written as follows:

$$\mathbf{y}^* = \mathbf{Ac} + \mathbf{Bd} + \varepsilon \tag{16}$$

Hereafter, to obtain regression function estimation of MSR model (16), we determine the solution to PWLS (7) which can be presented as follows:

$$\underset{g_{k \in \mathscr{G}_k}}{\mathrm{Min}} \left\{ \| \mathrm{W}^{\frac{1}{2}} \varepsilon \|^2 \right\} = \underset{g_{k \in \mathscr{G}_k}}{\mathrm{Min}} \left\{ \| \mathrm{W}^{\frac{1}{2}} (\mathbf{y}^* - \mathbf{g}) \|^2 \right\}$$

with constraint $\int_{a_k}^{b_k} \left(g_k^{(m)}(t_k)\right)^2 dt_k < \gamma_k$, $\gamma_k \geq 0$. Solution to the PWLS optimization is same as the solution to the following PWLS optimization:

$$\underset{g_{k \in W_2^m[a_k, b_k]}}{\mathrm{Min}} \left\{ \mathrm{N}^{-1} (\mathbf{y}^* - \mathbf{g})^T \mathbf{W} (\mathbf{y}^* - \mathbf{g}) + \sum_{k=1}^{p} \lambda_k \int_{a_k}^{b_k} \left(g_k^{(m)}(t_k)\right)^2 dt_k \right\} \tag{17}$$

where $\lambda_1, \lambda_2, \cdots, \lambda_p$ are smoothing parameters. These smoothing parameters set the balance between $\mathrm{N}^{-1}(\mathbf{y}^* - \mathbf{g})^T \mathbf{W} (\mathbf{y}^* - \mathbf{g})$**,** as goodness of fit, and $\sum_{k=1}^{p} \lambda_k \int_{a_k}^{b_k} \left(g_k^{(m)}(t_k)\right)^2 dt_k$ as the smoothness. To solve PWLS optimization (17), we decompose the penalty in (17) such that we get:

$$\sum_{k=1}^{p} \lambda_k \int_{a_k}^{b_k} \left(g_k^{(m)}(t_k)\right)^2 dt_k = d^T \phi B d$$

where $\phi = diag(\lambda_1 I_{n_1}, \lambda_2 I_{n_2}, \cdots, \lambda_p I_{n_p})$. Also, we get the goodness of fit:

$$\mathrm{N}^{-1}(\mathbf{y}^* - \mathbf{g})^T \mathbf{W} (\mathbf{y}^* - \mathbf{g}) = \mathrm{N}^{-1}(\mathbf{y}^* - \mathbf{Ac} - \mathbf{Bd})^T \mathbf{W} (\mathbf{y}^* - \mathbf{Ac} - \mathbf{Bd})$$

Hence, by combining penalty and goodness of fit, we obtain PWLS optimization whose solutions are:

$$\widehat{\mathbf{c}} = \left(\mathbf{A}^T \mathbf{D}^{-1} \mathbf{W} \mathbf{A}\right)^{-1} \mathbf{A}^T \mathbf{D}^{-1} \mathbf{W} \mathbf{y}^*$$

and. $\widehat{\mathbf{d}} = \mathbf{D}^{-1} \mathbf{W} \left[\mathbf{I} - \mathbf{A} \left(\mathbf{A}^T \mathbf{D}^{-1} \mathbf{W} \mathbf{A}\right)^{-1} \mathbf{A}^T \mathbf{D}^{-1} \mathbf{W} \right] \mathbf{y}^*$

where $D = WB + N\phi I$. Therefore, the estimated regression function in nonparametric component of MSR model (1) or (6) is:

$$\widehat{\mathbf{g}} = \mathbf{A}\widehat{\mathbf{c}} + \mathbf{B}\widehat{\mathbf{d}} = \mathbf{H}_\lambda \mathbf{y}^* \tag{18}$$

$$\text{where } \mathbf{H}_\lambda = \mathbf{A}\left(\mathbf{A}^T \mathbf{D}^{-1} \mathbf{W} \mathbf{A}\right)^{-1} \mathbf{A}^T \mathbf{D}^{-1} \mathbf{W} + \mathbf{B} \mathbf{D}^{-1} \mathbf{W} \left[\mathbf{I} - \mathbf{A} \left(\mathbf{A}^T \mathbf{D}^{-1} \mathbf{W} \mathbf{A}\right)^{-1} \mathbf{A}^T \mathbf{D}^{-1} \mathbf{W} \right] \tag{19}$$

Based on Eq. (3), we can express Eq. (18) as:

$$\widehat{\mathbf{g}} = \mathbf{H}_\lambda \mathbf{y}^* = \mathbf{H}_\lambda (\mathbf{y} - \mathbf{X}\beta) \tag{20}$$

Hence, the sum of squared errors (SSE) is given by:

$$\mathrm{Q} = [\mathbf{y} - \mathbf{X}\beta - (\mathbf{H}_\lambda(\mathbf{y} - \mathbf{X}\beta))]^T [\mathbf{y} - \mathbf{X}\beta - (\mathbf{H}_\lambda(\mathbf{y} - \mathbf{X}\beta))] \tag{21}$$

Next, by minimizing the SSE, we obtain the estimation of parameter β namely $\widehat{\beta}$ as follows:

$$\widehat{\beta} = \left[\mathbf{X}^T (\mathbf{I} - \mathbf{H}_\lambda)^T (\mathbf{I} - \mathbf{H}_\lambda) \mathbf{X}\right]^{-1} \mathbf{X}^T (\mathbf{I} - \mathbf{H}_\lambda)^T (\mathbf{I} - \mathbf{H}_\lambda) \mathbf{y} \tag{22}$$

where $\widehat{\beta}$ is a WLS estimator for parameters in parametric component of MSR model (1). Furthermore, by substituting Eq. (22) into Eq. (18), we get estimator of $\mathbf{g}$ as follows:

$$\hat{\mathbf{g}} = \mathbf{H}_\lambda\left[\mathbf{I}-\mathbf{X}\left(\mathbf{X}^T(\mathbf{I}-\mathbf{H}_\lambda)^T(\mathbf{I}-\mathbf{H}_\lambda)\mathbf{X}\right)^{-1}\mathbf{X}^T(\mathbf{I}-\mathbf{H}_\lambda)^T(\mathbf{I}-\mathbf{H}_\lambda)\right]\mathbf{y} \quad (23)$$

where $\hat{\mathbf{g}}$ is smoothing spline estimator for regression function $\mathbf{g}$ in nonparametric component of MSR model (1).

Finally, by considering MSR model (1) and based on estimation results given by equations (22) and (23), we obtain MSR model estimation based on smoothing spline as follows:

$$\hat{\mathbf{y}} = (\mathbf{H}_{par}+\mathbf{H}_{nonpar})\mathbf{y} = \mathbf{H}\mathbf{y} \quad (24)$$

where $\mathbf{H} = \mathbf{H}_{par}+\mathbf{H}_{nonpar}$; $\mathbf{H}_{par} = \mathbf{X}\left[\mathbf{X}^T(\mathbf{I}-\mathbf{H}_\lambda)^T(\mathbf{I}-\mathbf{H}_\lambda)\mathbf{X}\right]^{-1}\mathbf{X}^T(\mathbf{I}-\mathbf{H}_\lambda)^T(\mathbf{I}-\mathbf{H}_\lambda)$; and $\mathbf{H}_{nonpar} = \mathbf{H}_\lambda\left[\mathbf{I}-\mathbf{X}\left(\mathbf{X}^T(\mathbf{I}-\mathbf{H}_\lambda)^T(\mathbf{I}-\mathbf{H}_\lambda)\mathbf{X}\right)^{-1}\mathbf{X}^T(\mathbf{I}-\mathbf{H}_\lambda)^T(\mathbf{I}-\mathbf{H}_\lambda)\right]$.

Based smoothing spline in MSR model, estimator of $\mathbf{g}$ given in (23) is called weighted partial smoothing spline estimator of regression function of MSR model (1).

3.2. Determining confidence interval of $\boldsymbol{\beta}$

To determine a CI, we use pivotal quantity (Sahoo, 2013). We assume that ε_{ki} in (1) follows Normal distribution that independent and identic with mean zero and variance σ^2_{ki} or we write ε_{ki} $i.i.dN(0,\sigma^2_{ki})$ where σ^2_{ki} is unknown. Next, the $100(1-\alpha)\%$ CI for β_{ki}, $k=1,2,\cdots,p$; $i=1,2,\cdots,n_k$ is designed such that we have a pivotal quantity of parameter β_{ki}:

$$T_{ki}(y_{ki},x_{ki1},\cdots,x_{kiq_k},t_{ki1},\cdots,t_{kir_k}) = \frac{\left(\hat{\beta}_{ki}-\beta_{ki}\right)}{\sqrt{Var\left(\hat{\beta}_{ki}\right)}}$$

$$= \frac{\left(\hat{\beta}_{ki}-\beta_{ki}\right)}{\sqrt{(MSE(\lambda))\left(\Omega^{\mathrm{T}}\Omega\right)^{-1}_{ii}}} \quad (25)$$

where $\Omega = (\mathbf{I}-\mathbf{H}_\lambda)^{\mathrm{T}}(\mathbf{I}-\mathbf{H}_\lambda)\mathbf{X}$, $\mathrm{MSE}(\lambda) = \frac{\left(\mathbf{y}-\Omega(\Omega^{\mathrm{T}}\Omega)^{-1}\Omega^{\mathrm{T}}\mathbf{y}\right)^{\mathrm{T}}\left(\mathbf{y}-\Omega(\Omega^{\mathrm{T}}\Omega)^{-1}\Omega^{\mathrm{T}}\mathbf{y}\right)}{\left(\sum_{k=1}^{p}n_k\right)-n_k(1+q_k+r_k)}$, β_{ki} is the i^{th} element for k^{th} response of parameters vector β, and $\left(\Omega^{\mathrm{T}}\Omega\right)^{-1}_{ii}$ is diagonal element of $\left(\Omega^{\mathrm{T}}\Omega\right)^{-1}$. We can use GCV or CV instead of MSE to overcome over fitting (Amini & Roozbeh, 2015; Roozbeh, 2018; and Roozbeh et al., 2020).

Hereinafter, if $\mathbf{Z} = \Omega\left(\Omega^{\mathrm{T}}\Omega\right)^{-1}\Omega^{\mathrm{T}}$, then $\mathrm{MSE}(\lambda)$ in (25) is given by:

$$\mathrm{MSE}(\lambda) = \frac{\mathbf{y}^T(\mathbf{I}-\mathbf{Z})\mathbf{y}}{\left(\sum_{k=1}^{p}n_k\right)-n_k(1+q_k+r_k)}$$

$$= \frac{\mathbf{y}^{\mathrm{T}}\mathbf{P}\mathbf{y}}{\left(\sum_{k=1}^{p}n_k\right)-n_k(1+q_k+r_k)} \quad (26)$$

where $\mathbf{P} = (\mathbf{I}-\mathbf{Z})$. Hence, the pivotal quantity (25) can be expressed as follows:

$$T_{ki}(y_{ki},x_{ki1},\cdots,x_{kiq_k},t_{ki1},\cdots,t_{kir_k}) = \frac{\left(\hat{\beta}_{ki}-\beta_{ki}\right)}{\sqrt{\frac{\mathbf{y}^{\mathrm{T}}\mathbf{P}\mathbf{y}}{\left(\sum_{k=1}^{p}n_k\right)-n_k(1+q_k+r_k)}\left(\Omega^{\mathrm{T}}\Omega\right)^{-1}_{ii}}} \quad (27)$$

The pivotal quantity (27) follows a t-student distribution with $\left(\sum_{k=1}^{p}n_k\right)-n_k(1+q_k+r_k)$ degree of freedom.

Furthermore, to determine the $100(1-\alpha)\%$ CI for β_{ki}, $k=1,2,\cdots,p$; $i=1,2,\cdots,n_k$, we must take the solution to probability equation:

$$P\left[L_{ki} \le T_{ki}(y_{ki},x_{ki1},\cdots,x_{kiq_k},t_{ki1},\cdots,t_{kir_k}) \le U_{ki}\right] = 1-\alpha \quad (28)$$

where L_{ki} is lower limit of CI and U_{ki} is upper limit of CI, and $(1-\alpha)$ is level of confidence. Next, we substitute Eq.(27) into Eq. (28) so that we get:

$$P\left[L_{ki} \le \frac{\left(\hat{\beta}_{ki}-\beta_{ki}\right)}{\sqrt{\frac{\mathbf{y}^{\mathrm{T}}\mathbf{P}\mathbf{y}}{\left(\sum_{k=1}^{p}n_k\right)-n_k(1+q_k+r_k)}\left(\Omega^{\mathrm{T}}\Omega\right)^{-1}_{ii}}} \le U_{ki}\right] = 1-\alpha \quad (29)$$

We can write Eq.(29) as:

$$P\left(\hat{\beta}_{ki}-U \le \beta_{ki} \le \hat{\beta}_{ki}-L\right) = 1-\alpha \quad (30)$$

where $=L_{ki}\sqrt{\left(\frac{\mathbf{y}^{\mathrm{T}}\mathbf{P}\mathbf{y}}{\left(\sum_{k=1}^{p}n_k\right)-n_k(1+q_k+r_k)}\mathbf{V}\right)}$; $U = U_{ki}\sqrt{\left(\frac{\mathbf{y}^{\mathrm{T}}\mathbf{P}\mathbf{y}}{\left(\sum_{k=1}^{p}n_k\right)-n_k(1+q_k+r_k)}\mathbf{V}\right)}$; $\mathbf{V} = \left(\Omega^{\mathrm{T}}\Omega\right)^{-1}_{ii}$ and $= \mathbf{I}-\Omega\left(\Omega^{\mathrm{T}}\Omega\right)^{-1}\Omega^{\mathrm{T}} = \mathbf{I}-\mathbf{Z}$.

If interval length of CI is shortest then the CI is good. Therefore, we find values of $L_{ki} \in \mathbb{R}$ and $U_{ki} \in \mathbb{R}$ that results length of CI in (30) is the shortest. If $length(L_{ki},U_{ki})$ is length of CI in (30), then we have:

$$length(L_{ki},U_{ki}) =$$

$$\left(\hat{\beta}_{ki}-L_{ki}\sqrt{\left(\frac{\mathbf{y}^{\mathrm{T}}\mathbf{P}\mathbf{y}}{\left(\sum_{k=1}^{p}n_k\right)-n_k(1+q_k+r_k)}\mathbf{V}\right)}\right) - \left(\hat{\beta}_{ki}-U_{ki}\sqrt{\left(\frac{\mathbf{y}^{\mathrm{T}}\mathbf{P}\mathbf{y}}{\left(\sum_{k=1}^{p}n_k\right)-n_k(1+q_k+r_k)}\mathbf{V}\right)}\right)$$

$$= (U_{ki}-L_{ki})\sqrt{\left(\frac{\mathbf{y}^{\mathrm{T}}\mathbf{P}\mathbf{y}}{\left(\sum_{k=1}^{p}n_k\right)-n_k(1+q_k+r_k)}\mathbf{V}\right)}$$

Hence, the shortest length of CI for β_{ki} is determined by taking the solution to optimization:

$$\min_{L_{ki},U_{ki}\in\mathbb{R}}\{length(L_{ki},U_{ki})\}$$

$$= \min_{L_{ki},U_{ki}\in\mathbb{R}}\left\{(U_{ki}-L_{ki})\sqrt{\left(\frac{\mathbf{y}^{\mathrm{T}}\mathbf{P}\mathbf{y}}{\left(\sum_{k=1}^{p}n_k\right)-n_k(1+q_k+r_k)}\mathbf{V}\right)}\right\} \quad (31)$$

that meets the condition:

$$\int_{L_{ki}}^{U_{ki}}T(s)ds = 1-\alpha \text{ or } K(U_{ki})-K(L_{ki})-(1-\alpha) = 0 \quad (32).$$

where $T(\cdot)$ represents distribution of probability of $t_{\left(\sum_{k=1}^{p}n_k\right)-n_k(1+q_k+r_k)}$ and $K(\cdot)$ represents distribution of cumulative probability of $t_{\left(\sum_{k=1}^{p}n_k\right)-n_k(1+q_k+r_k)}$. Next, by applying Lagrange method, it results equation as follows:

$$R(L_{ki}, U_{ki}, \gamma) = (U_{ki} - L_{ki})\sqrt{\left(\frac{\mathbf{y}^{\mathrm{T}}\mathbf{P}\mathbf{y}}{\left(\sum_{k=1}^{p} n_k\right) - n_k(1 + q_k + r_k)}\mathbf{V}\right)} +$$

$$\gamma(K(U_{ki}) - K(L_{ki}) - (1 - \alpha)) \tag{33}$$

where γ is constant of Lagrange. Hereafter, the following equations are obtained:

$$\frac{\partial R(L_{ki}, U_{ki}, \gamma)}{\partial L_{ki}} = 0 \Longleftrightarrow -\sqrt{\left(\frac{\mathbf{y}^{\mathrm{T}}\mathbf{P}\mathbf{y}}{\left(\sum_{k=1}^{p} n_k\right) - n_k(1 + q_k + r_k)}\mathbf{V}\right)} - \gamma K'(L_{ki}) = 0 \tag{34}$$

$$\frac{\partial R(L_{ki}, U_{ki}, \gamma)}{\partial U_{ki}} = 0 \Longleftrightarrow \sqrt{\left(\frac{\mathbf{y}^{\mathrm{T}}\mathbf{P}\mathbf{y}}{\left(\sum_{k=1}^{p} n_k\right) - n_k(1 + q_k + r_k)}\mathbf{V}\right)} + \gamma K'(U_{ki}) = 0 \tag{35}$$

$$\frac{\partial R(L_{ki}, U_{ki}, \gamma)}{\partial \gamma} = 0 \Longleftrightarrow K(U_{ki}) - K(L_{ki}) - (1 - \alpha) = 0 \tag{36}$$

From equations (34) and (35), we obtain the following relationship:

$$K'(L_{ki}) = K'(U_{ki}) \tag{37}$$

The Eq.(37) implies $L_{ki} = U_{ki}$ or $L_{ki} = -U_{ki}$. Since, $L_{ki} = U_{ki}$ is not satisfied, then the shortest CI can be determined from the L_{ki} and U_{ki} values which fulfill:

$$\int_{-\infty}^{L_{ki}} T(s)ds = \int_{U_{ki}}^{\infty} T(s)ds = \frac{\alpha}{2} \tag{38}$$

By using $(1 - \alpha)$ level of confidence , the L_{ki} and U_{ki} values which fulfill condition (38) can be obtained from the $t_{\left(\sum_{k=1}^{p} n_k\right) - n_k(1+q_k+r_k)}$ distribution table.

Consequently, the shortest smoothing spline CI for parameters of MSR model fulfills the following probability:

$$P\left[\widehat{\beta}_{ki} - U_{ki}\sqrt{\left(\frac{\mathbf{y}^{\mathrm{T}}\mathbf{P}\mathbf{y}}{\left(\sum_{k=1}^{p} n_k\right) - n_k(1 + q_k + r_k)}\mathbf{V}\right)} \le \beta_{ki} \le \widehat{\beta}_{ki} + U_{ki}\sqrt{\left(\frac{\mathbf{y}^{\mathrm{T}}\mathbf{P}\mathbf{y}}{\left(\sum_{k=1}^{p} n_k\right) - n_k(1 + q_k + r_k)}\mathbf{V}\right)}\right] = 1 - \alpha$$

where value of U_{ki} can be determined from Eq.(38) which is $\int_{U_{ki}}^{\infty} T(s)ds = \frac{\alpha}{2}$. Hence, we have:

$$P\left[\widehat{\beta}_{ki} - t_{\left(\frac{\alpha}{2};N - n_k(1+q_k+r_k)\right)}\sqrt{\left(\frac{\mathbf{y}^{\mathrm{T}}\mathbf{P}\mathbf{y}}{N - n_k(1 + q_k + r_k)}\mathbf{V}\right)} \le \beta_{ki} \le \widehat{\beta}_{ki} + t_{\left(\frac{\alpha}{2};N - n_k(1+q_k+r_k)\right)}\sqrt{\left(\frac{\mathbf{y}^{\mathrm{T}}\mathbf{P}\mathbf{y}}{N - n_k(1 + q_k + r_k)}\mathbf{V}\right)}\right] = 1 - \alpha$$

$$N = \sum_{k=1}^{p} n_k$$

Finally, by using distribution of t-student, the $100(1 - \alpha)\%$ CIs parameters β_{ki}, $k = 1, 2, \cdots, p$; $i = 1, 2, \cdots, n_k$ of MSR model (1) are:

$$\left(\widehat{\beta}_{ki} \pm t_{\left(\frac{\alpha}{2};N - n_k(1+q_k+r_k)\right)}\sqrt{\left(\frac{\mathbf{y}^{\mathrm{T}}\mathbf{P}\mathbf{y}}{N - n_k(1 + q_k + r_k)}\mathbf{V}\right)}\right) \tag{39}$$

where $N = \sum_{k=1}^{p} n_k$; $\mathbf{V} = \left(\Omega^{\mathrm{T}}\Omega\right)_{ii}^{-1}$; $\mathbf{P} = \mathbf{I} - \Omega\left(\Omega^{\mathrm{T}}\Omega\right)^{-1}\Omega^{\mathrm{T}}$; $\Omega = (\mathbf{I} - \mathbf{H}_{\lambda})^{\mathrm{T}}(\mathbf{I} - \mathbf{H}_{\lambda})\mathbf{X}$; and $\mathbf{H}_{\lambda}$ is given in (19). The asymptotic distribution of $\widehat{\beta}_{ki}$ is Normal as presented by Theorem 2 in section 3.3.

3.3. Determining asymptotic distribution

For investigating asymptotic distribution of $\widehat{\beta}$, we consider the following lemmas and theorem.

Lemma 1. *Suppose $\boldsymbol{H}_{\lambda}$ is matrix presented in (19) and $\mathbf{g} = \left(\mathbf{g}_1(\mathbf{t}), \cdots, \mathbf{g}_p(\mathbf{t})\right)^T$ then.*

$$N^{-1}\left\|\left(\mathbf{I} - \mathbf{H}_{\lambda}^{T}\right)\mathbf{W}\mathbf{g}\right\|^2 \le \lambda\int_a^b \left[\mathbf{g}^{(m)}(\mathbf{t})\right]^2 d\mathbf{t}$$

$$N = \sum_{k=1}^{p} n_k$$

Proof of Lemma 1. Suppose $\widehat{\mathbf{g}}$ in (23) is a estimator of smoothing spline function $\mathbf{g}$ which makes the PWLS (7) is minimum, then for $0 < w_i < \infty$ and $\mathbf{g} \in W_2^m[a, b]$ we have:

$$N^{-1}\left\|\left(\mathbf{I} - \mathbf{H}_{\lambda}^{T}\right)\mathbf{W}\mathbf{g}\right\|^2 \le N^{-1}\left\|\left(\mathbf{I} - \mathbf{H}_{\lambda}^{T}\right)\mathbf{W}\mathbf{g}\right\|^2 + \lambda\int_a^b \left[\widehat{\mathbf{g}}^{(m)}(\mathbf{t})\right]^2 d\mathbf{t}$$

$$= N^{-1}\sum_{i=1}^{n_k}(w_i\mathbf{g}(\mathbf{t}_i) - \widehat{\mathbf{g}}(\mathbf{t}_i))^2 + \lambda\int_a^b \left[\widehat{\mathbf{g}}^{(m)}\right]^2 d\mathbf{t}$$

$$\le N^{-1}\left\|\left(\mathbf{I} - \mathbf{H}_{\lambda}^{T}\right)\mathbf{W}\mathbf{g}\right\|^2 + \lambda\int_a^b \left[\mathbf{g}^{(m)}\right]^2 d\mathbf{t}$$

$$\le N^{-1}\sum_{i=1}^{n_k}(w_i\mathbf{g}(\mathbf{t}_i) - w_i\mathbf{g}(\mathbf{t}_i))^2 + \lambda\int_a^b \left[\mathbf{g}^{(m)}\right]^2 d\mathbf{t}$$

$= \lambda\int_a^b \left[\mathbf{g}^{(m)}\right]^2 d\mathbf{t}$. □.

Lemma 2. *If $\boldsymbol{H}_{\lambda}$ is matrix as given in (19) and $\lambda \to 0$ or $\mathbf{g}^{(m)}(\mathbf{t}) = 0$ then*

$$\lim_{N\to\infty}\frac{1}{\sqrt{N^3}}\sum_j\left|\left[\left(\mathbf{I} - \mathbf{H}_{\lambda}^{T}\right)\mathbf{W}\mathbf{g}\right]_j\right|^3 = 0$$

$$N = \sum_{k=1}^{p} n_k$$

Proof of Lemma 2. With a little algebraic explanation, we obtain:

$$\frac{1}{\sqrt{N^3}}\sum_j\left|\left[\left(\mathbf{I} - \mathbf{H}_{\lambda}^{T}\right)\mathbf{W}\mathbf{g}\right]_j\right|^3 = \frac{1}{\sqrt{N^3}} \times \sum_j\left|\left[\left(\mathbf{I} - \mathbf{H}_{\lambda}^{T}\right)\mathbf{W}\mathbf{g}\right]_j\right|\left|\left[\left(\mathbf{I} - \mathbf{H}_{\lambda}^{T}\right)\mathbf{W}\mathbf{g}\right]_j\right|^2$$

$$\leq \frac{1}{\sqrt{N^3}} \max_j \left| \left[\left(\mathbf{I} - \mathbf{H}_\lambda^T \right) \mathbf{Wg} \right]_j \right| \sum_j \left| \left[\left(\mathbf{I} - \mathbf{H}_\lambda^T \right) \mathbf{Wg} \right]_j \right|^2$$

Consequently, we have relationship:

$$\frac{1}{\sqrt{N^3}} \sum_j \left| \left[\left(\mathbf{I} - \mathbf{H}_\lambda^T \right) \mathbf{Wg} \right]_j \right|^3 \leq \frac{1}{\sqrt{N^3}} \sqrt{\sum_j \left[\left(\mathbf{I} - \mathbf{H}_\lambda^T \right) \mathbf{Wg} \right]_j^2} \times \sum_j \left| \left[\left(\mathbf{I} - \mathbf{H}_\lambda^T \right) \mathbf{Wg} \right]_j \right|^2$$

$$= \frac{1}{\sqrt{N^3}} \left(\mathbf{g}^{\mathrm{T}} \mathbf{W} (\mathbf{I} - \mathbf{H}_\lambda) \left(\mathbf{I} - \mathbf{H}_\lambda^{\mathrm{T}} \right) \mathbf{Wg} \right)^{3/2}$$

For $\lambda \to 0$ or $\mathbf{g}^{(m)}(\mathbf{t}) = 0$, Lemma 1 gives:

$\frac{1}{\sqrt{N^3}} \sum_j \left| \left[\left(\mathbf{I} - \mathbf{H}_\lambda^T \right) \mathbf{Wg} \right]_j \right|^3 = \boldsymbol{o}(1)$ or $\lim_{N \to \infty} \frac{1}{\sqrt{N^3}} \sum_j \left| \left[\left(\mathbf{I} - \mathbf{H}_\lambda^T \right) \mathbf{Wg} \right]_j \right|^3 = 0.$ □.

Theorem 1. *If $\boldsymbol{H}_\lambda$ is matrix presented in (19) and $\lambda \to 0$ or $\mathbf{g}^{(m)}(\mathbf{t}) = 0$ then*

$\frac{\mathbf{X}^{\mathrm{T}} \left[\left(\mathbf{I} - \mathbf{H}_\lambda^{\mathrm{T}} \right) \mathbf{Wg} + \mathbf{W}\varepsilon \right]}{\sqrt{N}} \xrightarrow{d} D^* \; N\left(0, \sigma^2 \Sigma \vartheta\right)$ as $N \to \infty$,

$$N = \sum_{k=1}^{p} n_k$$

Proof of Theorem 1. Here, we apply the Cramer-Wold theorem (Cramer & Wold, 1936; Sen & Singer, 1993). Firstly, a vector $\mathbf{a}$ is given such that:

$$\frac{\mathbf{a}^{\mathrm{T}} \mathbf{X}^{\mathrm{T}} \left[\left(\mathbf{I} - \mathbf{H}_\lambda^{\mathrm{T}} \right) \mathbf{Wg} + \mathbf{W}\varepsilon \right]}{\sqrt{N}} = \sum_j \mathbf{Z}_j$$

where $\boldsymbol{Z}_j = \frac{(\mathbf{Xa})_j \left[(\mathbf{W}\varepsilon)_j + \left(\left(\mathbf{I} - \mathbf{H}_\lambda^{\mathrm{T}} \right) \mathbf{Wg} \right)_j \right]}{\sqrt{N}}$ is zero mean independent random variable, namely $\boldsymbol{Z}_j$ has mean $\mathbf{0}$ and variance $\sum_j \mathrm{Var}(\mathbf{Z}_j) = \mathbf{a}^{\mathrm{T}} \sigma^2 \Sigma \left(\frac{1}{N} \sum_{i=1}^{n_k} w_i \right) \mathbf{a} + \left(\boldsymbol{a}^{\mathrm{T}} \Sigma \boldsymbol{a} \right) \frac{1}{N} \sum_{\boldsymbol{j}} \left(\left[\left(\mathbf{I} - \mathbf{H}_\lambda^{\mathrm{T}} \right) \mathbf{Wg} \right]_{\boldsymbol{j}} \right)^2$.

Next, the following assumptions (A1, A2, A3) are given:

(A1). $t_i = \frac{2i-1}{2n_k}$; $i = 1, 2, \cdots, n_k$; $k = 1, 2, \cdots, p$.

(A2). $\mathbf{X}_1, \mathbf{X}_2, \cdots, \mathbf{X}_{n_k}$ follow a distribution that independent and identic with mean zero and covariance Σ, and the third absolute moment is finite.

(A3). $\lim_{n_k \to \infty} \sum_{i=1}^{n_k} w_i = \vartheta < \infty$.

Taking into account the assumptions (A1, A2, A3) and Lemma 1, then for $\lambda \to 0$ or $\mathbf{g}^{(m)}(\mathbf{t}) = 0$, $\sum_j \mathrm{Var}(\mathbf{Z}_j)$ converges to $\mathbf{a}^{\mathrm{T}} \sigma^2 \Sigma \vartheta \mathbf{a}$. Hence, we have:

$$\sum_j E\left| \mathbf{Z}_j \right|^3 = \frac{1}{\sqrt{N^3}} \sum_j E\left(\left| (\mathbf{Xa})_j \right|^3 \left| (\mathbf{W}\varepsilon)_j + \left[\left(\mathbf{I} - \mathbf{H}_\lambda^T \right) \mathbf{Wg} \right]_j \right|^3 \right)$$

$$= \frac{1}{\sqrt{N^3}} E\left| (\mathbf{Xa})_1 \right|^3 \sum_j E\left(\left| (\mathbf{W}\varepsilon)_j + \left[\left(\mathbf{I} - \mathbf{H}_\lambda^T \right) \mathbf{Wg} \right]_j \right|^3 \right)$$

Hence, we have relationship::

$$\sum_j E\left| \mathbf{Z}_j \right|^3 \leq E\left| (\mathbf{Xa})_1 \right|^3 \left(\frac{1}{\sqrt{N}} \max_j \left(E\left| (\mathbf{W}\varepsilon)_j \right|^3 \right) + \frac{1}{\sqrt{N^3}} \sum_j \left| \left[\left(\mathbf{I} - \mathbf{H}_\lambda^T \right) \mathbf{Wg} \right]_j \right|^3 \right)$$

Since Lemma 2 and the third absolute moment of $(\mathbf{W}\varepsilon)_j$ is finite, the $\sum_j E\left| \mathbf{Z}_j \right|^3$ leads to zero. Hence, $\sum_j \mathbf{Z}_j$ converges to $N\left(0, \mathbf{a}^{\mathrm{T}} \sigma^2 \Sigma \vartheta \mathbf{a}\right)$ namely Normally distributed. □.

Based on these lemmas and theorem, estimator $\widehat{\beta}$ is asymptotically normally distributed. More details for this are given in the following theorem.

Theorem 2. *If $\widehat{\beta}$ is parameters estimator of smoothing spline in parametric component the MSR model (1), and $\lambda \to 0$ or $\mathbf{g}^{(m)}(\mathbf{t}) = 0$ then.*

$$\sqrt{N}\left(\widehat{\beta} - \beta \right) \xrightarrow{d} D \; N\left(0, \boldsymbol{\sigma}^2 \Sigma^{-1} \boldsymbol{\vartheta}^{-1}\right) \text{ as } N \to \infty$$

$$N = \sum_{k=1}^{p} n_k$$

Proof of Theorem 2. We can express $\sqrt{N}\left(\widehat{\beta} - \beta \right)$ as:

$$\sqrt{N}\left(\widehat{\beta} - \beta \right) = \left(\frac{\mathbf{X}^{\mathrm{T}} \left(\mathbf{I} - \mathbf{H}_\lambda^T \right) \mathbf{WX}}{N} \right)^{-1} \left\{ \frac{\mathbf{X}^{\mathrm{T}} \left[\left(\mathbf{I} - \mathbf{H}_\lambda^T \right) \mathbf{Wg} + \mathbf{W}\varepsilon \right]}{\sqrt{N}} - \frac{\mathbf{X}^{\mathrm{T}} \mathbf{H}_\lambda^T \mathbf{W}\varepsilon}{\sqrt{N}} \right\}$$

Hence, we obtain:

$\left(\frac{\mathbf{X}^{\mathrm{T}} \left(\mathbf{I} - \mathbf{H}_\lambda^T \right) \mathbf{WX}}{N} \right)^{-1} \xrightarrow{P} \Sigma^{-1} \boldsymbol{\vartheta}^{-1}$ for $N \to \infty$; and $\frac{\mathbf{X}^{\mathrm{T}} \mathbf{H}_\lambda^T \mathbf{W}\varepsilon}{\sqrt{N}} \xrightarrow{P} 0$, as $N \to \infty$.

From Theorem 1, we have:

$\frac{\mathbf{X}^{\mathrm{T}} \left[\left(\mathbf{I} - \mathbf{H}_\lambda^T \right) \mathbf{Wg} + \mathbf{W}\varepsilon \right]}{\sqrt{N}} \xrightarrow{d} D^* \; N\left(0\,,\ \sigma^2 \Sigma \vartheta\right)$ as $\to \infty$.

Next, by applying Slutsky theorem (Sen & Singer, 1993), we obtain:

$\sqrt{N}\left(\widehat{\beta} - \beta \right) \xrightarrow{d} D \; N\left(0, \boldsymbol{\sigma}^2 \Sigma^{-1} \boldsymbol{\vartheta}^{-1}\right)$ as $\to \infty$. □.

4. Discussion

The estimated regression function of MSR model is a combination between the estimated parametric component namely $\widehat{\beta}$, and the estimated nonparametric functions namely $\widehat{\mathbf{g}}$. In this case, $\widehat{\beta}$ is a WLS estimator for parameter β contained in component of parametric and $\widehat{\mathbf{g}}$ is smoothing spline regression function estimator of $\mathbf{g}$ contained in component of nonparametric of the MSR model. Hence, the smoothing spline MSR model estimation is to be linear to observations $\mathbf{y}$ where its hessian matrix $\mathbf{H}$ given by Eq.(24) is also a combination between hessian matrix of parametric component, $\mathbf{H}_{par}$, and hessian matrix of nonparametric component, $\mathbf{H}_{nonpar}$.

In interval estimation concept, a good CI is the one with the shortest interval length. Therefore, we determine lower limit value of CI (L_{ki}) and upper limit value of CI (U_{ki}) such that length of CI is the shortest. The shortest CIs for parameters of MSR model are given in Eq. (39) that depend on t-student distribution because variance of population is unknown. Hereafter, for more statistical inference purposes, the asymptotic distribution of MSR model parameters estimator was also undertaken, and finally we obtained that estimator $\widehat{\beta}$ in (22) is asymptotically normally distributed, namely $N\left(0, \boldsymbol{\sigma}^2 \Sigma^{-1} \boldsymbol{\vartheta}^{-1}\right)$ as given in proof of Theorem 2.

5. Conclusion

The estimated MSR model is a composed estimations between component of parametric and component of nonparametric, and its functional relationship is linear to observation. Also, the $100(1-\alpha)\%$ CIs for parameters β_{ki} ($k = 1, 2, \cdots, p$; $i = 1, 2, \cdots, n_k$) follow distribution of t-student namely $t_{\left(\frac{\alpha}{2}; N - n_k(1 + q_k + r_k)\right)}$, and the estimator $\widehat{\beta}$ is asymptotically normally distributed. Future time, this study results can be used as theoretical bases to design standard growth charts of the toddlers for assessing the nutritional status of the toddlers.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We thank the editors and reviewers who have given constructive corrections, criticisms, and suggestions which can be used to improve the quality of the manuscript.

Funding

We disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was funded by the DRPM, Ministry of Education and Culture, Republic of Indonesia through the PDUPT Grant No. 473/UN3.15/PT/2021.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jksus.2023.102664.

References

Amini, M., Roozbeh, M., 2015. Optimal partial ridge estimation in restricted semiparametric regression models. J. Multivariate Anal. 136, 26–40.
Aronszajn, N., 1950. Theory of reproducing kernels. Trans. Amer. Math. Soc. 68, 337–404.
Aydin, D., Ahmed, S.E., Yilmaz, E., 2019. Estimation of semiparametric regression model with right-censored high-dimensional data. J. Stat. Comp. Simul. 89 (6), 985–1004.
Chamidah, N., Lestari, B., Massaid, A., Saifudin, T., 2020a. Estimating mean arterial pressure affected by stress scores using spline nonparametric regression model approach. Commun. Math. Biol. Neurosci. 2020 (2020) 72, 1–12.
Chamidah, N., Lestari, B., 2019. Estimating of covariance matrix using multi-response local polynomial estimator for designing children growth charts: a theoretically discussion. J. Phys.: Conf. Ser. 1397, (1) 012072.
Chamidah, N., Tjahjono, E., Fadilah, A.R., Lestari, B., 2018. Standard growth charts for weight of children in East Java using local linear estimator. J. Phys.: Conf. Ser. 1097, (1) 012092.
Chamidah, N., Gusti, K.H., Tjahjono, E., Lestari, B., 2019a. Improving of classification accuracy of cyst and tumor using local polynomial estimator. TELKOMNIKA 17 (3), 1492–1500.
Chamidah, N., Lestari, B., Saifudin, T., 2019b. Modeling of blood pressures based on stress score using least square spline estimator in bi-response nonparametric regression. Int. J. Innov., Creat. Change 5 (3), 1200–1216.
Chamidah, N., Zaman, B., Muniroh, L., Lestari, B., 2019c. Standard growth charts of children in East Java province using a local linear estimator. Int. J. Innov., Creat. Change 13 (1), 45–67.
Chamidah, N., Yonani, Y.S., Ana, E., Lestari, B., 2020b. Identification the number of mycobacterium tuberculosis based on sputum image using local linear estimator. Bullet. Elect. Eng. Inform. 9 (5), 2109–2116.
Chamidah, N., Lestari, B., Wulandari, A.Y., Muniroh, L., 2021. Z-Score standard growth chart design of toddler weight using least square spline semiparametric regression. AIP Conf. Proc. 2329, 060031.
Chamidah, N., Lestari, B., Budiantara, I.N., Saifudin, T., Rulaningtyas, R., Aryati, A., Wardani, P., Aydin, D., 2022. Consistency and asymptotic normality of estimator for parameters in multiresponse multipredictor semiparametric regression model. Symmetry 14(2) 336, 1–18.
Chen, W., Ren, F.L., 2020. Polynomial-based smoothing estimation for a semiparametric accelerated failure time partial linear model. Open Access Library J. 7, 1–15.
Cramér, H., Wold, H., 1936. Some theorems on distribution functions. J. London Math. Soc. 11 (4), 290–295.
Diana, R., Budiantara, I.N., Purhadi, Darmesto, S., 2013. Smoothing spline in semiparametric additive regression model with Bayesian approach. J. Math. Stats. 9 (3), 161–168.
Eubank, R.L., 1988. Spline Smoothing and Nonparametric Regression. Marcel Dekker, New York.
Fatmawati, Budiantara, I.N., Lestari, B., 2019. Comparison of smoothing and truncated spline estimators in estimating blood pressures models. Int. J. Innov., Creat. Change 5 (3), 1177–1199.
Fernandes, A.A.R., Widiastuti, D.A., Nurjanah, 2020. Smoothing spline semiparametric regression model assumption using PWLS Approach. Int. J. Adv. Sci. Technol. 29 (4), 2059–2070.
Gao, J., Shi, P., 1997. M-Type smoothing splines in nonparametric and semiparametric regression models. Statistica Sinica 7 (4), 1155–1169.
Gu, C., 2002. Smoothing Spline ANOVA Models. Springer-Verlag, New York.
Islamiyati, A., Kalondeng, A., Sunusi, N., Zakir, M., Amir, A.K., 2022. Biresponse nonparametric regression model in principal component analysis with truncated spline estimator. J. King Saud Univ.-Sci. 34(3) 101892, 1–9.
Khan, A., Shahna, 2019. Non-polynomial quadratic spline method for solving fourth order singularly perturbed boundary value problems. J. King Saud Univ.–Sci. 31 (4), 479–484.
Lestari, B., Fatmawati, Budiantara, I.N., Chamidah, N., 2018. Estimation of regression function in multiresponse nonparametric regression model using smoothing spline and kernel estimators. J. Phys.: Conf. Ser. 1097, (1) 012091.
Lestari, B., Fatmawati, Budiantara, I.N., Chamidah, N., 2019. Smoothing parameter selection method for multiresponse nonparametric regression model using spline and kernel estimators approaches. J. Phys.: Conf. Ser. 1397, (1) 012064.
Lestari, B., Fatmawati, Budiantara, I.N., 2020. Spline estimator and its asymptotic properties in multiresponse nonparametric regression model. Songklanakarin J. Sci. Technol. 42 (3), 533–548.
Lestari, B., Chamidah, N., Aydin, D., Yilmaz, E., 2022. Reproducing kernel Hilbert space approach to multiresponse smoothing spline regression function. Symmetry 14(11) 2227, 1–22.
Mohaisen, A.J., Abdulhussein, A.M., 2015. Spline semiparametric regression models. J. Kufa Math. Comp. 2 (3), 1–10.
Osmani, F., Hajizadeh, E., Mansouri, P., 2019. Kernel and regression spline smoothing techniques to estimate coefficient in rates model and its application in psoriasis. Medic. J. Islamic Rep. Iran 33 (90), 1–5.
Ramadan, W., Chamidah, N., Zaman, B., Muniroh, L., Lestari, B., 2019. Standard growth chart of weight for height to determine wasting nutritional status in East Java based on semiparametric least square spline estimator. IOP Conf. Ser.: Mater. Sci. Eng. 546, 052063.
Roozbeh, M., 2018. Optimal QR-based estimation in partially linear regression models with correlated errors using GCV criterion. Comput. Stats. Data Anal. 117, 45–61.
Roozbeh, M., Arashi, M., Hamzah, N.A., 2020. Generalized cross-validation for simultaneous optimization of tuning parameters in ridge regression. Iranian J. Sci. Technol. Transactions A: Science 44, 473–485.
Sahoo, P., 2013. Probability and Mathematical Statistics. University of Louisville, Lousville.
Sen, P.K., Singer, J.M., 1993. Large Sample in Statistics: An Introduction with Applications. Chapman & Hall, London.
Shahna, Khan, A., 2019. Approximation for higher order boundary value problems using non-polynomial quadratic spline base don off-step points. J. King Saud Univ.–Sci. 31 (4), 737–745.
Wahba, G., 1990. Spline Models for Observational Data. SIAM, Philadelphia.
Wang, Y., 2011. Smoothing Splines: Methods and Applications. Chapman and Hall, London.
Wang, Y., Guo, W., Brown, M.B., 2000. Spline smoothing for bivariate data with applications to association between hormones. Statistica Sinica 10 (2), 377–397.
Wang, Y., Ke, C., 2009. Smoothing spline semiparametric nonlinear regression models. J. Comp. Graphical Stats. 18 (1), 165–183.
Wibowo, W., Haryatmi, S., Budiantara, I.N., 2012. On multiresponse semiparametric regression model. J. Math. Stats. 8 (4), 489–499.
Yilmaz, E., Yusbasi, B., Aydin, D., 2021. Choice of smoothing parameter for kernel type ridge estimators in semiparametric regression models. REVSTAT-Stat. J. 19 (1), 47–69.