Smoothing parameter selection method for multiresponse nonparametric regression model using smoothing spline and Kernel estimators approaches

by Fatmawati Fatmawati

Journal of Physics: Conference Series

PAPER • OPEN ACCESS

Smoothing parameter selection method for multiresponse nonparametric regression model using smoothing spline and Kernel estimators approaches

Smoothing parameter selection method for multiresponse nonparametric regression model using smoothing spline and Kernel estimators approaches

B Lestari[1], Fatmawati[2], I N Budiantara[3], and N Chamidah[2]

[1]Faculty of Mathematics and Natural Sciences, The University of Jember, 37 Kalimantan Street, Campus of Tegal Boto, Jember 68121, Indonesia.
[2]Mathematics Department, Faculty of Science and Technology, Airlangga University, Mulyorejo Street, Campus C UNAIR, Surabaya 60115, Indonesia.
[3]Statistics Department, Sepuluh Nopember Institute of Technology, Campus of Sukolilo, Surabaya 60111, Indonesia.

Corresponding author's e-mail: fatmawati@fst.unair.ac.id

Abstract. The principle problem in multiresponse nonparametric regression model is how we estimate the regression functions which draw association between some dependent (response) variables and some independent (predictor) variables where there are correlations between responses. There are many techniques used to estimate the regression function. Two of them are spline and kernel smoothing techniques. Speaking about smoothing techniques, not only in uniresponse spline and kernel nonparametric regression models but also in multiresponse spline and kernel nonparametric regression models, the estimations of regression functions depend on smoothing parameters. In the privious researches the covariance matrices were assumed to be known. Matrix of covariance is not assumed known in this research. The goals of this research are selecting of optimal smoothing parameters for the model we consider through spline and kernel smoothing techniques. Optimal smoothing parameters can be obtained by taking the solution to generalized cross validation (GCV) optimization problem. The obtained results of this research are the optimal smoothing parameter for smoothing spline estimator approach and the optimal smoothing parameter namely optimal bandwidth for kernel estimator approach.

1. Introduction

In statistical analysis we frequently involve a building mathematical model which examines the functional association between one or more than one dependent (response) and one or more than one independent (predictor) variables. Discussing about association between some dependent (response) variables and some independent (predictor) variables which is drawn by a function, we may not neglect an usual model namely a regression model. In analysis of regression which uses parametric and nonparametric regression models approaches we often meet to principle problem, i.e., estimate functions of regression. For approaching model of nonparametric regression where it has no specific pattern of the regression function, we can use some estimators to estimate function of regression in the model, for example spline, kernel, local linear, etc. In this approach, spline estimator is more flexible to estimate regression function than other estimators. Researchers who have used spline estimator in various cases of models of nonparametric regression with single response are [1-21]. Next, [22-28] used this estimator to estimate functions of regressions in models of nonparametric regression with two responses and more than two responses. On the other hand, [29-35] have proposed kernel estimator to estimate models of nonparametric regression with single response. Furthermore, [36-38] have studied these estimators, i.e., spline and kernel, for estimating functions of regression in models of nonparametric regression with both single response and multi-responses.

Estimate regression functions in models of nonparametric regression with single response that uses smoothing techniques, either spline estimator or kernel estimator, depends on smoothing parameters [2, 10, 20, and 39]. In single response spline nonparametric regression, if smoothing parameter value is very small then it will give a very rought estimator of nonparametric regression function [20]. In contrary, if smoothing parameter value is very large then it will give a very smooth estimator of nonparametric regression function [20]. Therefore, we need to select optimal smoothing parameter in order to obtain estimator that is suitable with data. For this need, some researchers have proposed some selection methods for selecting the optimal smoothing parameters, for instance [2] proposed cross validation (CV) method, [10] proposed unbiased risk (UBR) method, and [20 and 39-40] proposed generalized cross validation (GCV) methods.

All researchers who have been mentioned previously, discussed the estimating methods of regression functions in either uniresponse or multiresponse nonparametric regression models only [1-38]. These researchers have not discussed the selecting of smoothing parameters methods. Further, although [2, 10, 20, and 39-40] have discussed the selecting of smoothing parameters methods, but these researchers have not discussed the selecting of smoothing parameters methods for the multiresponse nonparametric regression model by using kernel estimator. Also, previous researchers have not estimated the covariance matrix. Therefore, in this research we discuss method to select optimal smoothing parameters for the considered model by using both smoothing spline and kernel techniques.

2. Results of research and discussion

Here, we provide a mathematical statistics method for selecting the smoothing parameters for the model that we consider through both smoothing spline and kernel techniques.

2.1. Smoothing parameters selection method by using smoothing spline estimator approach

Firstly, suppose (y_{ij}, x_{ij}) is data set which accords with the following multiresponse nonparametric regression model:

$$y_{ij} = g_i(x_{ij}) + \varepsilon_{ij}\ ,\quad a_i \le x_i \le b_i\ ,\quad i = 1,2,\ldots,p\ ,\quad j = 1,2,\ldots,n_i \tag{1}$$

where y_{ij} represents i^{th} response of j^{th} observation, unknown smooth regression functions $g_1, g_2, \ldots, g_p$ are included in Sobolev space $W_2^m[a_i, b_i]$, and independent random errors ε_{ij} have zero-mean and variance σ_{ij}^2 [28 and 37].

Secondly, based on (1) we have $\underset{\sim}{y} = (\underset{\sim}{y}_1, \underset{\sim}{y}_2, \ldots, \underset{\sim}{y}_p)'$, $\underset{\sim}{f} = (\underset{\sim}{f}_1, \underset{\sim}{f}_2, \ldots, \underset{\sim}{f}_p)'$, $\underset{\sim}{\varepsilon} = (\underset{\sim}{\varepsilon}_1, \underset{\sim}{\varepsilon}_2, \ldots, \underset{\sim}{\varepsilon}_p)'$, and $\underset{\sim}{x} = (\underset{\sim}{x}_1, \underset{\sim}{x}_2, \ldots, \underset{\sim}{x}_p)'$ where $\underset{\sim}{y}_i = (y_{i1}, y_{i2}, \ldots, y_{in_i})'$, $\underset{\sim}{g}_i = (g_i(x_{i1}), g_i(x_{i2}), \ldots, g_i(x_{in_i}))'$, $\underset{\sim}{\varepsilon}_i = (\varepsilon_{i1}, \varepsilon_{i2}, \ldots, \varepsilon_{in_i})'$, $\underset{\sim}{x}_i = (\underset{\sim}{x}_{i1}, \underset{\sim}{x}_{i2}, \ldots, \underset{\sim}{x}_{in_i})'$. So, for $j = 1,2,\ldots,n_i$ and $i = 1,2,\ldots,p$, the equation (1) can be expressed in the following vector notation:

$$\underset{\sim}{y} = \underset{\sim}{g} + \underset{\sim}{\varepsilon} \tag{2}$$

where $\underset{\sim}{\varepsilon}$ have zero-mean and covariance of $[W(\underset{\sim}{\sigma}^2)]^{-1} = diag(W_1(\underset{\sim}{\sigma}_1^2), W_2(\underset{\sim}{\sigma}_2^2), \ldots, W_p(\underset{\sim}{\sigma}_p^2))$ [28]. Estimation of regression functions $\underset{\sim}{g}$ in (2) through smoothing spline method occurs as a solution to minimization problem of penalized weighted least-square (PWLS). Thus, we must find $\hat{\underset{\sim}{g}}$ that can make the following PWLS minimum:

$$\underset{g_1,\ldots,g_p\in W_2^m}{Min}\{(\sum_{i=1}^{p} n_i)^{-1}(\underset{\sim}{y}_1-\underset{\sim}{g}_1)'W_1(\underset{\sim}{\sigma}_1^2)(\underset{\sim}{y}_1-\underset{\sim}{g}_1)+\ldots+(\underset{\sim}{y}_p-\underset{\sim}{g}_p)'W_p(\underset{\sim}{\sigma}_p^2)(\underset{\sim}{y}_p-\underset{\sim}{g}_p)+\lambda_1\int_{a_1}^{b_1}(g_1^{(2)}(\underset{\sim}{x}_1))^2 d\underset{\sim}{x}_1+\ldots+\lambda_p\int_{a_p}^{b_p}(g_p^{(2)}(\underset{\sim}{x}_p))^2 d\underset{\sim}{x}_p\} \quad (3)$$

for pre-specified value of $\underset{\sim}{\lambda}=(\lambda_1,\lambda_2,\ldots,\lambda_p)'$. Equation (3) shows that the first expression represents sum squares of errors (MSE) which penalizes lack of fit, and the second expression with its weight $\underset{\sim}{\lambda}$ represents penalty of roughness. Thus, the penalty of roughness penalizes the curvature of $\underset{\sim}{g}$. Further, we call λ_i $(i=1,2,\ldots,p)$ as parameter of smoothing. If $0<\lambda_i<+\infty$ then the solution of (3) will be different from interpolation to a linear model. Next, equation (3) will be dominated by the roughness penalty, and estimating of smoothing spline will lead to be a constant, if $\lambda_i\to+\infty$. On the other hand, the penalty of roughness will not occur in (3) if $\lambda_i\to 0$, and the data will be interpolated by estimating of smoothing spline. It means that smoothing parameter λ_i controls the trade-off between the goodness of fit:

$$(\sum_{i=1}^{p} n_i)^{-1}(\underset{\sim}{y}_1-\underset{\sim}{g}_1)'W_1(\underset{\sim}{\sigma}_1^2)(\underset{\sim}{y}_1-\underset{\sim}{g}_1)+\ldots+(\underset{\sim}{y}_p-\underset{\sim}{g}_p)'W_p(\underset{\sim}{\sigma}_p^2)(\underset{\sim}{y}_p-\underset{\sim}{g}_p)$$

and smoothness of the estimate:

$$\lambda_1\int_{a_1}^{b_1}(g_1^{(2)}(\underset{\sim}{x}_1))^2 d\underset{\sim}{x}_1+\ldots+\lambda_p\int_{a_p}^{b_p}(g_p^{(2)}(\underset{\sim}{x}_p))^2 d\underset{\sim}{x}_p \cdot$$

The solution to optimization in (3) is a smoothing spline estimator. Function basis of this estimator is a spline of natural cubic where its knots are $x_1,x_2,\ldots,x_{n_i}$ $(i=1,2,\ldots,p)$. It means that interpolation of special structured spline which depends on the smoothing parameter λ_i value selection gets turn into a representative approach of the regression functions g_i $(i=1,2,\ldots,p)$ in model (1). Next, by applying reproducing kernel Hilbert space (RKHS) method, we determine regression function estimate, and its result is:

$$\hat{\underset{\sim}{g}}_{\lambda}(\underset{\sim}{x})=\left(\hat{g}_{1,\lambda_1}(\underset{\sim}{x}_1),\hat{g}_{2,\lambda_2}(\underset{\sim}{x}_2),\ldots,\hat{g}_{p,\lambda_p}(\underset{\sim}{x}_p)\right)'=K\hat{\underset{\sim}{d}}+\Sigma\hat{\underset{\sim}{c}}=A(\underset{\sim}{\lambda})\underset{\sim}{y} \quad (4)$$

where

$$A(\underset{\sim}{\lambda})=K[K'M^{-1}W(\underset{\sim}{\sigma}^2)K]^{-1}K'M^{-1}W(\underset{\sim}{\sigma}^2)+\Sigma M^{-1}W(\underset{\sim}{\sigma}^2)[I-K(K'M^{-1}W(\underset{\sim}{\sigma}^2)K)^{-1}K'M^{-1}W(\underset{\sim}{\sigma}^2)]$$

is called as "hat matrix" and a non-singular matrix, so it is a invertible matrix [28 and 37], and $\hat{\underset{\sim}{g}}_{\lambda}(\underset{\sim}{x})$ is estimated regression function based on smoothing spline estimator where its function basis is natural cubic spline with knots $x_1,x_2,\ldots,x_{n_i}$ $(i=1,2,\ldots,p)$, for a fixed parameter of smoothing $\underset{\sim}{\lambda}>\underset{\sim}{0}$. So, $A(\underset{\sim}{\lambda})$ is a smoother matrix that is positive-definite (symmetrical), and it does not depend on $\underset{\sim}{y}$, but it depends on parameter of smoothing $\underset{\sim}{\lambda}$ and knots $x_1,x_2,\ldots,x_{n_i}$ $(i=1,2,\ldots,p)$ [28, 37 and 40].

Furthermore, we may express the estimated regression function of model in (4) as follows:

$$\hat{\underset{\sim}{g}}_{\lambda}(\underset{\sim}{x})=A(\lambda_1,\ldots,\lambda_p;\underset{\sim}{\sigma}^2)\underset{\sim}{y} \quad (5)$$

where $\underset{\sim}{\sigma}^2=(\sigma_1^2,\ldots,\sigma_p^2)'$. Mean square error (MSE) of (5) can be determined as follows:

$$MSE(\lambda_1,\ldots,\lambda_p;\underset{\sim}{\sigma}^2)=\frac{(\underset{\sim}{y}-\hat{\underset{\sim}{g}}_{\lambda}(\underset{\sim}{x}))'W(\underset{\sim}{\sigma}^2)(\underset{\sim}{y}-\hat{\underset{\sim}{g}}_{\lambda}(\underset{\sim}{x}))}{\sum_{i=1}^{p} n_i}$$

$$=\frac{\left(\underset{\sim}{y}-A(\lambda_1,...,\lambda_p;\underset{\sim}{\sigma}^2)\underset{\sim}{y}\right)'W(\underset{\sim}{\sigma}^2)\left(\underset{\sim}{y}-A(\lambda_1,...,\lambda_p;\underset{\sim}{\sigma}^2)\underset{\sim}{y}\right)}{N}$$

$$=\frac{\left[\left(W(\underset{\sim}{\sigma}^2)\right)^{\frac{1}{2}}\left(\left(I_N-A(\lambda_1,...,\lambda_p;\underset{\sim}{\sigma}^2)\right)\underset{\sim}{y}\right)\right]'\left[\left(W(\underset{\sim}{\sigma}^2)\right)^{\frac{1}{2}}\left(\left(I_N-A(\lambda_1,...,\lambda_p;\underset{\sim}{\sigma}^2)\right)\underset{\sim}{y}\right)\right]}{N}$$

$$=\frac{\left\|\left(W(\underset{\sim}{\sigma}^2)\right)^{\frac{1}{2}}\left(I_N-A(\lambda_1,...,\lambda_p;\underset{\sim}{\sigma}^2)\right)\underset{\sim}{y}\right\|^2}{N}$$

where norm $\|\underset{\sim}{a}\|$ is defined as $\|\underset{\sim}{a}\|=\sqrt{a_1^2+a_2^2+...+a_p^2}$ for $\underset{\sim}{a}=(a_1,a_2,...,a_p)'$, $N=\sum_{i=1}^{p}n_i$, I_N is an identity matrix, and $W(\underset{\sim}{\sigma}^2)$ is a diagonal matrix that has been given by [28, 37 and 40], so that $\left(W(\underset{\sim}{\sigma}^2)\right)^{\frac{1}{2}}$ is also a diagonal matrix and it has the square root .

Next, we define a quantity (further it is called as GCV function) as follows:

$$G(\lambda_1,...,\lambda_p;\underset{\sim}{\sigma}^2)=\frac{N^{-1}\left\|\left(W(\underset{\sim}{\sigma}^2)\right)^{\frac{1}{2}}\left(I_N-A(\lambda_1,...,\lambda_p;\underset{\sim}{\sigma}^2)\right)\underset{\sim}{y}\right\|^2}{\left[N^{-1}trace\left(I_N-A(\lambda_1,...,\lambda_p;\underset{\sim}{\sigma}^2)\right)\right]^2}$$

where $(I_N-A(\lambda_1,...,\lambda_p;\underset{\sim}{\sigma}^2)$ is a non-zero matrix, because $\lambda_i>0$, $i=1,2,...,p$.
Finally, we obtain optimal smoothing parameter, namely $\underset{\sim}{\lambda}_{optimal}=(\lambda_{1(optimal)},...,\lambda_{p(optimal)})'$ by taking the solution to the following optimization:

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

$$=\underset{\lambda_1\in R^+,...,\lambda_p\in R^+}{Min}\left\{\frac{N^{-1}\left\|\left(W(\underset{\sim}{\sigma}^2)\right)^{\frac{1}{2}}\left(I_N-A(\lambda_1,...,\lambda_p;\underset{\sim}{\sigma}^2)\right)\underset{\sim}{y}\right\|^2}{\left[N^{-1}trace\left(I_N-A(\lambda_1,...,\lambda_p;\underset{\sim}{\sigma}^2)\right)\right]^2}\right\}. \quad (6)$$

2.2. Smoothing parameters selection method by using Kernel estimator approach.
In analysis of nonparametric regression, to estimate the function g in model (1) through kernel smoothing technique, basically we use a decreasing distance function in x-space that is a weighted raw data average.

For model given in (1), according to [37] for predicting at x_{ij}, the weighted raw data average scheme by relating it with i^{th}-response observations, namely y_{ik}, is given by:

$$v_{(i)jk}=\frac{K_i(\frac{t_{ij}-t_{ik}}{h_i})}{\sum_{k=1}^{n_i}K_i(\frac{t_{ij}-t_{ik}}{h_i})}=\frac{K_i(u)}{\sum_{k=1}^{n_i}K_i(u)}\quad,\quad i=1,2,...,p \quad (7)$$

where $K(u)$ is called as kernel function which is a decreasing function of u, and $h_i > 0$ is a parameter of smoothing called as bandwidth. The kernel function $K(u)$ should be a symmetric function. It usually give a probability density function for example Gaussian function [34-35].

By considering equation (7), we get the estimated regression function of model given in (1) to estimate function of regression at point of fit x_{ij} which is expressed as:

$$\hat{\underset{\sim}{g}}_k(x_{ij}) = \hat{y}_{ij} = \sum_{k=1}^{n_i} v_{(i)jk} y_{ik} = \underset{\sim}{v}'_{ik} \underset{\sim}{y}_i \ , \quad i = 1,2,...,p \ , \quad j = 1,2,...,n_i . \tag{8}$$

where each point of n_i points is given by a unequal weight $v_{(i)jk}$, $k = 1,2,...,n_i$ for any point of fit x_{ij}. Thus, we can express the estimated regression function in (8) as:

$$\hat{\underset{\sim}{g}} = V\underset{\sim}{y} \tag{9}$$

where $V = diag\left(V_{i1}, V_{i2}, ..., V_{in_i}\right)$, and

$$V_{ij} = \begin{pmatrix} v_{(i)11} & v_{(i)12} & \cdots & v_{(i)1n_i} \\ v_{(i)21} & v_{(i)22} & \cdots & v_{(i)2n_i} \\ \vdots & \vdots & \ddots & \vdots \\ v_{(i)n_i1} & v_{(i)n_i2} & \cdots & v_{(i)n_in_i} \end{pmatrix}.$$

Therefore, we can determine estimator of regression function of model given in (1) based on kernel estimator, and its result is [37]:

$$\hat{\underset{\sim}{g}}(\underset{\sim}{x}) = \left(\hat{\underset{\sim}{g}}_1(\underset{\sim}{x}_1), \hat{\underset{\sim}{g}}_2(\underset{\sim}{x}_2), ..., \hat{\underset{\sim}{g}}_p(\underset{\sim}{x}_p)\right)' = diag\left(V_{i1}, V_{i2}, ..., V_{in_i}\right)\left(\underset{\sim}{y}_1, \underset{\sim}{y}_2, ..., \underset{\sim}{y}_p\right)'. \tag{10}$$

Since the estimated regression function given in (9) depends on smoothing parameter namely bandwidth $h_i > 0$, $i = 1,2,..., p$ then we also may write (9) in the following equation:

$$\hat{\underset{\sim}{g}}_h(\underset{\sim}{x}) = B(h_1, ..., h_p; \underset{\sim}{\sigma}^2)\underset{\sim}{y} \tag{11}$$

Furthermore, based on equations (9)-(11) we can determine the mean square error (MSE) of (11) as follows:

$$MSE(h_1, ..., h_p; \underset{\sim}{\sigma}^2) = \frac{(\underset{\sim}{y} - \hat{\underset{\sim}{g}}_h(\underset{\sim}{x}))' V(\underset{\sim}{\sigma}^2)(\underset{\sim}{y} - \hat{\underset{\sim}{g}}_h(\underset{\sim}{x}))}{\sum_{i=1}^{p} n_i}$$

$$= \frac{\left(\underset{\sim}{y} - B(h_1, ..., h_p; \underset{\sim}{\sigma}^2)\underset{\sim}{y}\right)' V(\underset{\sim}{\sigma}^2)\left(\underset{\sim}{y} - B(h_1, ..., h_p; \underset{\sim}{\sigma}^2)\underset{\sim}{y}\right)}{N}$$

$$= \frac{\left[\left(V(\underset{\sim}{\sigma}^2)\right)^{\frac{1}{2}}\left(\left(I_N - B(h_1, ..., h_p; \underset{\sim}{\sigma}^2)\right)\underset{\sim}{y}\right)\right]' \left[\left(V(\underset{\sim}{\sigma}^2)\right)^{\frac{1}{2}}\left(\left(I_N - B(h_1, ..., h_p; \underset{\sim}{\sigma}^2)\right)\underset{\sim}{y}\right)\right]}{N}$$

$$= N^{-1}\left\|\left(V(\underset{\sim}{\sigma}^2)\right)^{\frac{1}{2}}\left(I_N - B(h_1, ..., h_p; \underset{\sim}{\sigma}^2)\right)\underset{\sim}{y}\right\|^2$$

where norm $\|\underset{\sim}{a}\|$ is defined as $\|\underset{\sim}{a}\| = \sqrt{a_1^2 + a_2^2 + ... + a_p^2}$ for $\underset{\sim}{a} = (a_1, a_2, ..., a_p)'$, $N = \sum_{i=1}^{p} n_i$, I_N is an identity matrix , and $V(\underset{\sim}{\sigma}^2)$ is a diagonal matrix that has been given by [28 and 37], so that $\left(V(\underset{\sim}{\sigma}^2)\right)^{\frac{1}{2}}$ is also a diagonal matrix and it has the square root.

Next, we define a quantity as follows:

$$G(h_1,...,h_p;\underset{\sim}{\sigma}^2) = \frac{N^{-1}\left\|\left(V(\underset{\sim}{\sigma}^2)\right)^{\frac{1}{2}}\left(I_N - B(h_1,...,h_p;\underset{\sim}{\sigma}^2)\right)\underset{\sim}{y}\right\|^2}{\left[N^{-1}trace\left(I_N - B(h_1,...,h_p;\underset{\sim}{\sigma}^2)\right)\right]^2}$$

where $(I_N - B(h_1,...,h_p;\underset{\sim}{\sigma}^2)$ is a non-zero matrix because $h_i > 0$, $i = 1,2,...,p$.

Finally, we obtain the optimal smoothing parameter called as optimal bandwidth $\underset{\sim}{h}_{optimal} = (h_{1(optimal)},...,h_{p(optimal)})'$ by taking the solution to the following optimization:

$$G_{optimal}(h_{1(optimal)},...,h_{p(optimal)};\underset{\sim}{\sigma}^2) = \underset{h_1\in R^+,...,h_p\in R^+}{Min}\left\{G(h_1,...,h_p;\underset{\sim}{\sigma}^2)\right\}$$

$$= \underset{h_1\in R^+,...,h_p\in R^+}{Min}\left\{\frac{N^{-1}\left\|\left(V(\underset{\sim}{\sigma}^2)\right)^{\frac{1}{2}}\left(I_N - H(h_1,...,h_p;\underset{\sim}{\sigma}^2)\right)\underset{\sim}{y}\right\|^2}{\left[N^{-1}trace\left(I_N - H(h_1,...,h_p;\underset{\sim}{\sigma}^2)\right)\right]^2}\right\}. \qquad (12)$$

3. Conclusion

The estimated regression functions in model (1) by using both smoothing spline estimator and kernel estimator approaches depends on smoothing parameters $(\underset{\sim}{\lambda})$ and bandwidths $(\underset{\sim}{h})$, respectively. In addition, by minimazing GCV functions given in (6) and in (12) we obtained optimal smoothing parameter for smoothing spline estimator approach and optimal smoothing parameter called as optimal bandwidth for kernel estimator approach, respectively.

Acknowledgement

Many thanks to the Faculty of Science and Technology, Airlangga University, Surabaya, Indonesia which has funded authors to present this paper in the 6th ICRIEMS 2019.

References

[1] Kimeldorf G and Wahba G 1971 Some Result on Tchebycheffian Spline Functions *J. of Math. Anal. and Applications* **33**(1) 82-95.

[2] Craven P and Wahba G 1979 Smoothing Noisy Data With Spline Function: Estimating the Correct Degree of Smoothing by the Method of Generalized Cross Validation Numer. Math. 31 377-403.

[3] Wahba G 1985 A Comparison of GCV and GML for Choosing the Smoothing Parameter in the Generalized Spline Smoothing Problem The Annals of Statistics 13(4) 1378-1402.

[4] Cox D D 1983 Asymptotic for M-type Smoothing Spline The Annals. of Statistics 11(2) 530-551.

[5] Cox D D and O'Sullivan F 1996 Penalized Likelihood Type Estimators for Generalized Nonparametric Regression J. Mult. Anal. 56 185-206.

[6] Wahba G 1983 Bayesian Confidence Intervals for the Cross-Validated Smoothing Spline Journal of the Royal Statistical Society Series B 45 133-150.
[7] Oehlert G W 1992 Relaxed Boundary Smoothing Spline The Annals. of Statistics 20 146-160.
[8] Koenker R, Pin N G and Portnoy S 1994 Quantile Smoothing Splines Biometrics 81(4) 673-680.
[9] Budiantara I N, Subanar and Soejoeti Z 1997 Weighted Spline Estimator Proc. 51st Session of the International Statistical Institute Istanbul 51 333-334.
[10] Wang Y 1998 Smoothing Spline Models With Correlated Random Errors J. Amer. Statist. Assoc. 93(93) 341-348.
[11] Wahba G 2000 An Introduction to Model Building with Reproducing Kernel Hilbert Spaces Technical Report No. 1020, 18 April 2000, Department of Statistics, University of Wisconsin, Madison.
[12] Cardot, Crambes H C, Kneip A and Sarda P 2007 Smoothing Splines Estimators in Functional Linear Regression with Errors-in-Variables Computational Statistics & Data Analysis 51 4832-4848.
[13] Liu A, Tong T and Wang Y 2007 Smoothing Spline Estimation of Variance Functions Journal of Computational and Graphical Statistics 16(2) 312-329.
[14] Wegman E J 1981 Vector spline and estimation of filter function Technometrics 23 83-89.
[15] Miller J J and Wegman E J 1987 Vector function estimation using splines J. Royal Stat. Society, Series B 17 173-180.
[16] Flessler J A 1991 Nonparametric fixed-interval smoothing with vector splines IEEE Trans. Signal Proceeding 39 852–859.
[17] Wahba G 1992 Multivariate functional and operator estimation, based on smoothing spline and reproducing kernels.
[18] Gooijer J G D, Gannoun A and Larramendy I 1999 Nonparametric regression with serially correlated errors 1st Edition Tinbergen Institute Rotterdam pp:19.
[19] Wang Y 2011 Smoothing Splines: Methods and Applications CRC Press New York.
[20] Wahba G 1990 Spline Models for Observational Data SIAM Philadelphia Pennsylvania.
[21] Eubank R L 1999 Nonparametric Regression and SmoothingSpline Marcel Deker New York.
[22] Wang Y, Guo W and Brown W B 2000 Spline smoothing for bivariate data with applications to association between hormones Statistica Sinica 10 377-397.
[23] Chamidah N and 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 1433-1453.
[24] Lestari B, Anggraeni D, and Saifudin T 2017 Konstruksi dan Estimasi Matriks Kovariansi dalam Model Regresi Nonparametrik Multirespon Berdasarkan Estimator Smoothing Spline untuk Beberapa Kasus Ukuran Sampel Proceeding of National Seminar on Mathematics and Its Applications (SNMA) 2017 Airlangga University Surabaya pp. 238-242.
[25] Lestari B, Fatmawati and Budiantara I N 2017 Estimasi Fungsi Regresi Nonparametrik Multirespon Menggunakan Reproducing Kernel Hilbert Space Berdasarkan Estimator Smoothing Spline Proceeding of National Seminar on Mathematics and Its Applications (SNMA) 2017 Airlangga University Surabaya pp. 243-250.
[26] Lestari B, Anggraeni D and Saifudin T 2018 Estimation of Covariance Matrix based on Spline estimator in Homoscedastic Multi-Responses Nonparametric Regression Model in Case of Unbalance Number of Observations Far East J.Math.Sciences (FJMS) 108(2) 341-355.
[27] Islamiyati A, Fatmawati and Chamidah N 2018 Estimation of Covariance Matrix on Bi-Response Longitudinal Data Analysis with Penalized Spline Regression Journal of Physics: Conf. Series Vol. 979 pp. 012093.
[28] Lestari B, Fatmawati and Budiantara I N 2019 Spline Estimator and Its Asymptotic Properties in Multiresponse Nonparametric Regression Model Songklanakarin Journal of Science and Technology In Press.

[29] Green P J and Silverman B W 1994 Nonparametric Regression and Generalized Linear Models Chapman Hall New York.
[30] Hardle W 1991 Applied Nonparametric Regression Cambridge University Press Cambridge.
[31] Schimek M G 2000 Smoothing and Regression John Wiley & Sons New York.
[32] Watson G S 1964 Smooth Regression Analysis Sankhya Series A 26 359-372.
[33] Nadaraya E A and Watson 1964 On Estimating Regression Theory Prob. Applications 10 186-190.
[34] Yatchew A 2003 Semiparametric Regression for the Applied Econometrician Cambridge University Press Cambridge.
[35] Wand M P and Jones M C 1995 Kernel Smoothing Chapman Hall New York pp. 114-141.
[36] Aydin D 2007 A comparison of the nonparametric regression models using smoothing spline and kernel regression World Acad. Sci. Eng. Technol. 26 730-734.
[37] Lestari B, Fatmawati, Budiantara I N and Chamidah N 2018 Estimation of regression function in multi-response nonparametric regression model using smoothing spline and kernel estimators Journal of Physics: Conference Series 1097 012091.
[38] Lestari B, Chamidah N and Saifudin T 2019 Estimasi fungsi regresi dalam model regresi nonparametrik birespon menggunakan estimator smoothing spline dan estimator kernel Jurnal Matematika, Statistika and Komputasi (JMSK) 15(2) 20-24.
[39] Fernandez F M and Opsomer J D 2005 Smoothing parameter selection methods for nonparametric regression with spatially correlated errors Canadian J. Statistics 33 279-295.
[40] Lestari B, Fatmawati and Budiantara I N 2019 Smoothing spline estimator in multiresponse nonparametric regression for predicting blood pressures and pulse: a theoretically discussion Proceeding of Global Conference on Engineering and Applied Science (GCEAS) July 16-18, 2019 in Sapporo, Hokkaido, Japan 71618 pp. 81-93.

Smoothing parameter selection method for multiresponse nonparametric regression model using smoothing spline and Kernel estimators approaches

ORIGINALITY REPORT

10%	8%	8%	1%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	hdl.handle.net Internet Source	2%
2	journal.stats.id Internet Source	1%
3	tubdok.tub.tuhh.de Internet Source	1%
4	issuu.com Internet Source	1%
5	www.wseas.us Internet Source	1%
6	N. V. Taylor, C. B. Allen, A. Gaitonde, D. P. Jones. "A structure-coupled CFD method for time-marching flutter analysis", The Aeronautical Journal, 2016 Publication	1%
7	Rahmat Hidayat, I. Nyoman Budiantara, Bambang W. Otok, Vita Ratnasari. "The	1%

regression curve estimation by using mixed smoothing spline and kernel (MsS-K) model", Communications in Statistics - Theory and Methods, 2020
Publication

8	Oktavian, Rama, Vika Amidelsi, Agung Rasmito, and Gede Wibawa. "Vapor pressure measurements of ethanol–isooctane and 1-butanol–isooctane systems using a new ebulliometer", Fuel, 2013. Publication	<1%
9	Edy Setiti Wida Utami, Sucipto Hariyanto, Yosephine Sri Wulan Manuhara. "In vitro propagation of the endangered medicinal orchid, Dendrobium lasianthera J.J.Sm through mature seed culture", Asian Pacific Journal of Tropical Biomedicine, 2017 Publication	<1%
10	Bai, Huang, Sheng Li, and Xiongxiong He. "Sensing Matrix Optimization Based on Equiangular Tight Frames with Consideration of Sparse Representation Error", IEEE Transactions on Multimedia, 2016. Publication	<1%
11	www.stat.purdue.edu Internet Source	<1%
12	www.maths.manchester.ac.uk	

Internet Source

<1%

13 N. Ukhrowiyah, M. Yasin. "One step linear reconstruction method for continuous wave diffuse optical tomography", AIP Publishing, 2017
Publication
<1%

14 thescipub.com
Internet Source
<1%

15 "Mathematical and Statistical Methods in Insurance and Finance", Springer Science and Business Media LLC, 2008
Publication
<1%

16 dl4a.org
Internet Source
<1%

17 Congwen Luo. "On symmetric extended MS-algebras whose congruences are permutable", Acta Mathematica Scientia, 2011
Publication
<1%

18 R. Apsari, E. Nahdliyatun, D. Winarni. "The regeneration of thermal wound on mice skin (Mus Musculus) after Q-Switch Nd: YAG laser irradiation for cancer therapy candidate", AIP Publishing, 2017
Publication
<1%

Smoothing parameter selection method for multiresponse nonparametric regression model using smoothing spline and Kernel estimators approaches

GRADEMARK REPORT

FINAL GRADE

/0

GENERAL COMMENTS

Instructor

PAGE 1

PAGE 2

PAGE 3

PAGE 4

PAGE 5

PAGE 6

PAGE 7

PAGE 8

PAGE 9