Spline Estimator in Multi-Response Nonparametric Regression Model

Budi Lestari [1)], I Nyoman Budiantara [2)], Sony Sunaryo [2)], Muhammad Mashuri [2)]

[1)]*Department of Mathematics, Faculty of Mathematics and Natural Sciences, University of Jember*
[2)]*Department of Statistics, Faculty of Mathematics and Natural Sciences, Sepuluh Nopember Institute of Technology,*

ABSTRACT

In many applications two or more dependent variables are observed at several values of the independent variables, such as at time points. The statistical problems are to estimate functions that model their dependences on the independent variables, and to investigate relationships between these functions. Nonparametric regression model, especially smoothing splines provides powerful tools to model the functions which draw association of these variables. Penalized weighted least-squares is used to jointly estimate nonparametric functions from contemporaneously correlated data. In this paper we formulate the multi-response nonparametric regression model and give a theoretical method for both obtaining distribution of the response and estimating the nonparametric function in the model. We also estimate the smoothing parameters, the weighting parameters and the correlation parameter simultaneously by applying three methods: generalized maximum likelihood (GML), generalized cross validation (GCV) and leaving-out-one-pair cross validation (CV).

Keywords : Multi-response Nonparametric Regression Model, Penalized Weighted Least-Squares, Generalized Maximum Likelihood, Generalized Cross Validation, leaving-out-one-pair cross validation

INTRODUCTION

There are many writers who have studied spline estimators for estimating regression curve of nonparametric regression models. Kimeldorf & Wahba (1971), Craven & Wahba (1979) and Wahba (1990) proposed original spline estimator to estimates regression curve of smooth data. Cox (1983) and Cox & O'Sullivan (1996) used M-type spline to overcome outliers in nonparametric regression. Wahba (1983) proposed polynomial spline to obtain confidence interval based on posterior covariance function.

Oehlert (1992) and Koenker & Portnoy (1994) introduced relaxed spline and quantile spline, respectively. Budiantara *et al*. (1997) studied weighted spline estimator in nonparametric regression model with different variance. Wahba (2000) introduced some techniques for spline statistical model building by using reproducing kernel Hilbert spaces. Aydin (2007) showed goodness of spline estimator rather than kernel estimator in estimating nonparametric regression model for gross national product data. All these writers studied spline estimators in case of single response nonparametric models only.

In the real cases, we are frequently faced to the problem in which two or more dependent variables are observed at several values of the independent variables, such as at time points. Multi-response nonparametric regression model provide powerful tools to model the functions which draw association of these variables.

Many authors have considered nonparametric models for multiresponse data. Wegman (1981), Miller & Wegman (1987) and Flessler (1991) proposed algorithms for spline smoothing. Wahba (1992) developed the theory of general smoothing splines using reproducing kernel Hilbert spaces. Gooijer *et al.* (1991) and Fernez & Opsomer (2005) proposed methods of estimating nonparametric regression models with serially and spatially correlated errors, respectively. Wang *et al*. (2000) proposed spline smoothing for estimating nonparametric functions from bivariate data. Lestari (2007) studied spline smoothing for estimating three responses nonparametric regression models with the same variances of errors for the same response. Lestari (2008a) developed spline estimator in biresponse nonparametric regression model with unequal variances of errors and Lestari (2008b) developed penalized weighted least-squares estimator for bivariate nonparametric regression model with correlated errors. All, except Wang *et al*. (2000) and Lestari (2007, 2008a, & 2008b), assumed that the covariance matrix is known, which is usually not the case in practice. When the covariance matrix is unknown, it has to be estimated from the data and this can affect the estimates of the smoothing parameters (Wang 1998).

In this paper, we study mathematical statistics methods for obtaining distribution of responses, and estimating the nonparametric functions and the parameters in the multi-response nonparametric regression model. Here, we assume that the covariance parameters are unknown, and errors of the same responses have the same variances. Based on the multi-response nonparametric regression model given, we estimate multi-response nonparametric regression function by using penalized weighted least-squares. Next, we describe three methods: generalized maximum likelihood (GML), generalized cross validation (GCV) and leaving-out-one-pair cross validation (CV) to estimate the smoothing parameters, the weighting parameters and the correlation parameter simultaneously.

RESULTS AND DISCUSSION

Multi-response nonparametric regression models

Assume that data $\{y_{ki}, t_{ki}\}$ follows multi-response nonparametric regression model:

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

where $k = 1, 2, ..., p$; $i = 1, 2, ..., n_k$. It means that the i^{th} response of the k^{th} variable y_{ki} is generated by the k^{th} function f_k evaluated at the design point t_{ki} plus a random error ε_{ki}.

Assume $\varepsilon_{ki} \overset{i.i.d.}{\sim} N\left(0, \sigma_k^2\right)$ for fixed $k = 1, 2, ..., p$; and $Corr(\varepsilon_{ki}, \varepsilon_{li}) = \rho$ for $k \neq l$ and zero otherwise. It is a special case of our other paper, i.e., $Corr(\varepsilon_{ki}, \varepsilon_{li}) = \rho_i$, which has been submitted for an international journal

Here, for simplicity of notation, we assume that the domain of the functions are [0,1] and f_k is element of Sobolev space W_2, i.e., $f_k \in W_2 = \{ f : f, f'$ absolutely continuous, $\int_0^1 (f''(t))^2\, dt < \infty \}$. Our methods can be easily extended to the general smoothing spline models where the p domains are arbitrary (thus could be different) and the observations are linear functionals instead of evaluations (Wahba 1990, 1992).

Distribution of the responses

Suppose that we denote $\underline{t}_k = (t_{k1}, ..., t_{kn_k})^T$; $\underline{y}_k = (y_{k1}, ..., y_{kn_k})^T$; $\underline{\varepsilon}_k = (\varepsilon_{k1}, ..., \varepsilon_{kn_k})^T$; $\underline{f}_k = (f_k(t_{k1}), f_k(t_{k2}), ..., f_k(t_{kn_k}))^T$; $\underline{f} = (\underline{f}_1^T, ..., \underline{f}_p^T)^T$ and $\underline{y} = (\underline{y}_1^T, ..., \underline{y}_p^T)^T$, where the superscript T refers to transpose. For for $k = 1, 2, ..., p$, let $r_k = \dfrac{\sigma_k}{m}$, $m = \prod_{\substack{j=1 \\ j\neq k}}^{p} \sigma_j$; $\theta = \prod_{k=1}^{p} \sigma_k$; $\gamma_{ij} = \dfrac{\rho}{\sigma_k}\left(i, j = 1, 2, ..., p; k \neq i \neq j\right)$; and J_{qs} be a $n_q \times n_s$ matrix with $(i, j)^{th}$ element equal to 1 if the i^{th} element of $\underline{y}_q$ and the j^{th} element of $\underline{y}_s$ is a pair, and zero otherwise. Note that $J = I$, the identity matrix, when all observations come in pairs. By taking $E\left(\underline{y}\right)$ and $Var\left(\underline{y}\right)$, we obtain the distribution of responses, i.e., $\underline{y} \sim N(\underline{f}, \theta W^{-1})$, where

$$W^{-1} = \begin{bmatrix} r_1 I_{n_1} & \gamma_{12} J_{12} & \cdots & \gamma_{1p} J_{1p} \\ \gamma_{12} J_{12}^T & r_2 I_{n_2} & \cdots & \gamma_{2p} J_{2p} \\ . & . & \cdots & . \\ . & . & \cdots & . \\ . & . & \cdots & . \\ \gamma_{1p} J_{1p}^T & \gamma_{2p} J_{2p}^T & \cdots & r_p I_{n_p} \end{bmatrix} \qquad (2)$$

Spline estimator of multi-response nonparametric model

The nonparametric functions f_k are estimated by carrying out the following penalized weighted least-squares :

$$\underset{f_1, f_2, ..., f_p \in W_2}{Min} \{(\underline{y}-\underline{f})^T W(\underline{y}-\underline{f}) + \lambda_1 \int_0^1 (f_1''(t))^2 dt + \lambda_2 \int_0^1 (f_2''(t))^2 dt + \ldots + \lambda_p \int_0^1 (f_p''(t))^2 dt \} \quad (3)$$

The parameters λ_k ($k = 1, 2, ..., p$) control the trade-off between goodness-of-fit and the smoothness of the estimates and are referred to as smoothing parameters.

We extend method as in Wang (1998) (i.e., in case of single-response nonparametric regression model) to multi-response nonparametric regression model. Let $\phi_\nu(t) = t^{\nu-1}/(\nu-1)!$, $\nu = 1, 2, ..., p$; $R^1(s,t) = k_2(s)k_2(t) - k_4(s-t)$, where $k_\nu(.) = B_\nu(.)/\nu!$ and $B_\nu(.)$ is the ν^{th} Bernoulli polynomial. Let $T_k = \{\phi_\nu(t_{ki})\}_{i=1,\nu=1}^{n_k}$; $T = diag(T_1, ..., T_p)$; $\Sigma_k = \{R^1(t_{ki}, t_{kj})\}_{i=1,j=1}^{n_k, n_k}$ and $\Sigma = diag(\Sigma_1, \Sigma_2, ..., \Sigma_p)$. By extending method as in both Wang (1998) and Wahba (1990) to multi-response case, we can show that for fixed λ_k, γ_k for $k = 1, 2, ..., p$; and ρ, the solution to (3) is

$$\hat{f}_k(t) = \sum_{\nu=1}^{p} d_{k\nu}\phi_\nu(t) + \sum_{i=1}^{n_k} c_{ki} R^1(t, t_{ki}) \quad (4)$$

where $k = 1, 2, ..., p$; and

$\underline{c} = (c_{11}, ..., c_{1n_1}, c_{21}, ..., c_{2n_2}, ..., c_{p1}, ..., c_{pn_p})^T$;

$\underline{d} = (d_{11}, ..., d_{1p}, d_{21}, ..., d_{2p}, ..., d_{p1}, ..., d_{pp})^T$ are solutions to

$$\begin{pmatrix} T^T W T & T^T W \Sigma \\ \Sigma W T & \Sigma W \Sigma + diag(\lambda_1\Sigma_1, ..., \lambda_p\Sigma_p) \end{pmatrix} \begin{pmatrix} \underline{d} \\ \underline{c} \end{pmatrix} = \begin{pmatrix} T^T W \underline{y} \\ \Sigma W \underline{y} \end{pmatrix} \quad (5)$$

Note that $\underline{\hat{f}} = (\hat{f}_1(t_{11}), ..., \hat{f}_1(t_{1n_1}) \hat{f}_2(t_{21}), \ldots, \hat{f}_2(t_{2n_2}), ..., \hat{f}_p(t_{p1}), ..., \hat{f}_p(t_{pn_p}))^T$ is always unique when T is of full column rank, which are assumed to be true in this paper. It can be verified that a solution to

$$\left.\begin{aligned} \left(\Sigma + W^{-1} diag\left(\lambda_1 I_{n_1}, ..., \lambda_p I_{n_p}\right)\right)\underline{c} + T\underline{d} &= \underline{y} \\ T^T \underline{c} &= 0 \end{aligned}\right\} (6)$$

is also a solution to (5). Thus we need to solve simultaneous equation (6) for $\underline{c}$ and $\underline{d}$. In fact, $W^{-1}diag\left(\lambda_1 I_{n_1}, ..., \lambda_p I_{n_p}\right)$ is asymmetric if $\lambda_1 \neq \lambda_2 \neq ... \neq \lambda_p$ and $\rho \neq 0$. To calculate the coefficients $\underline{c}$ and $\underline{d}$, we use the following transformations:

$\tilde{\Sigma} = \Sigma diag\left(I_{n_1}/\lambda_1, I_{n_2}/\lambda_2, ..., I_{n_p}/\lambda_p\right)$ and $\tilde{c} = diag\left(\lambda_1 I_{n_1}, ..., \lambda_p I_{n_p}\right)\underline{c}$. Then (6) is equivalent to

$$\left.\begin{aligned} \left(\tilde{\Sigma} + W^{-1}\right)\underline{\tilde{c}} + T\underline{d} &= \underline{y} \\ T^T \underline{\tilde{c}} &= 0 \end{aligned}\right\} \quad (7)$$

Let $T_k = \begin{pmatrix} Q_{k1} & Q_{k2} \end{pmatrix}\begin{pmatrix} R_k \\ 0 \end{pmatrix}$, $k = 1, 2, ..., p$ be the *QR* decompositions. Let

$Q_1 = diag(Q_{11}, Q_{21}, Q_{31}, ..., Q_{p1})$;

$Q_2 = diag(Q_{12}, Q_{22}, Q_{32}, ..., Q_{p2})$;

$R = diag(R_1, R_2, ..., R_p)$; and $B = \tilde{\Sigma} + W^{-1}$. It can be shown that the solutions to (7) are

$$\underline{\tilde{c}} = Q_2(Q_2^T B Q_2)^{-1} Q_2^T \underline{y},$$

$$R\underline{d} = Q_1^T(\underline{y} - B\underline{\tilde{c}}) \quad (8)$$

Note that $\underline{\hat{f}} = A\underline{y}$ where

$$A = I - W^{-1} Q_2 (Q_2^T B Q_2)^{-1} Q_2^T \quad (9)$$

is the "hat matrix". Here, *A* is not symmetric, which is different from the usual independent case.

Estimations of parameters

We have assumed that the parameters λ_k, γ_{ij} (for $i, j, k = 1, 2, ..., p; k \neq i \neq j$) and ρ are fixed. In practice it is very important to estimate these parameters from the data. Since observations are correlated, popular methods such as the usual generalized maximum likelihood (GML) method and the generalized cross validation (GCV) method may underestimate the smoothing parameters

(Wang 1998). In this section we propose the following three methods to estimate the smoothing parameters λ_k, the weighting parameters r_k, and γ_{ij}; and the correlation parameter ρ simultaneously, i.e. an extension of the GML method based on a Bayesian model; an extension of the GCV method; and leaving-out-one-pair cross validation.

Wang (1998) proposed the GML and GCV methods for correlated observations with one smoothing parameter. Wang *et al.* (2000) proposed the GML and GCV methods for correlated observations with two smoothing parameters. In multi-response (with *p* responses) nonparametric regression model, there are *p* smoothing parameters which need to be estimated simultaneously together with the covariance parameters. Following an extension of derivation, we extend the GML and GCV in both Wang (1998) and Wang *et al.* (2000) as follows.

The GML estimates of λ_k, γ_{ij}, r_k and ρ are minimizers of the following GML function:

$$M(\lambda_k, \gamma_{ij}, r_k, \rho) = \frac{\underline{y}^T W(I-A)\underline{y}}{\left[\det^+(W(I-A))\right]^{\frac{1}{n-4}}} = \frac{\underline{z}^T (Q_2^T B Q_2)^{-1}\underline{z}}{\left[\det(Q_2^T B Q_2)^{-1}\right]^{\frac{1}{n-4}}} \qquad (10)$$

where $n = n_1 + n_2 + ... + n_p$; $\det^+$ is the product of the nonzero eigen values and $\underline{z} = Q_2^T \underline{y}$. The minimizers of $M(\lambda_k, \gamma_{ij}, r_k, \rho)$ are called GML estimates.

The GCV estimates of λ_k, γ_{ij}, r_k and ρ are minimizers of the following GCV function :

$$V(\lambda_k, \gamma_{ij}, r_k, \rho) = \frac{\left\|W(I-A)\underline{y}\right\|^2}{\left[Tr(W(I-A))\right]^2} = \frac{\underline{z}^T (Q_2^T B Q_2)^{-2}\underline{z}}{\left[Tr(Q_2^T B Q_2)^{-1}\right]^2} \qquad (11)$$

In the following we propose a cross validation method based on leaving-out-one-pair procedure. Suppose there are a total of *N* ($N \geq \max\{n_1, n_2, ..., n_p\}$) distinct time points and thus *N* pairs of observations. Any one observation in a pair may be missing. These pairs are numbered from 1 to *N*. We use the following notation: superscripts (*i*) to denote the collection of elements corresponding to the i^{th} pair; superscripts [*i*] to denote the collection of elements after deleting the i^{th} pair; superscripts {*i*} to denote solution of f_k without the i^{th} pair. When one observation in a pair is missing, superscripts indicate a single observation instead of a pair. The solutions to :

$$\underset{f_1,...,f_p \in W_2}{Min} \{(\underline{y}^{[i]} - \underline{f}^{[i]})^T W^{[i]} (\underline{y}^{[i]} - \underline{f}^{[i]}) + \lambda_1 \int_0^1 (f_1''(t))^2 dt + \lambda_2 \int_0^1 (f_2''(t))^2 dt + ... + \lambda_p \int_0^1 (f_p''(t))^2 dt \} \qquad (12)$$

are $\hat{f}_1^{\{i\}}$, $\hat{f}_2^{\{i\}}$, …, $\hat{f}_p^{\{i\}}$. Assume that there are *p* elements in the i^{th} pair (it is simple if there is only one). Denote i_1, i_2, …, i_p as the row numbers of this pair in $\underline{y}_1$, $\underline{y}_2$, …, $\underline{y}_p$, respectively. Define:

$$y_{kj}^* = \begin{cases} y_{kj}, \ j \neq i_k \\ \hat{f}_k^{\{i\}}(t_{ki_k}), \ j = i_k, k = 1, 2, ..., p \end{cases}$$

Suppose that we denote $\underline{y}_k^* = (y_{k1}^*, ..., y_{kn_k}^*)^T$, $\underline{y}^* = (\underline{y}_1^{*T}, ..., \underline{y}_p^{*T})^T$, and $\hat{f}^{\{i\}}(\underline{t}) = (\hat{f}_1^{\{i\}}(t_{11}), ..., \hat{f}_1^{\{i\}}(t_{1n_1}), \hat{f}_2^{\{i\}}(t_{21}), ..., \hat{f}_2^{\{i\}}(t_{2n_2}), ..., \hat{f}_p^{\{i\}}(t_{p1}), ..., \hat{f}_p^{\{i\}}(t_{pn_p}))$. Then we have the following *leaving-out-one-pair* lemma.

Lemma. *For fixed* $\lambda_k, \gamma_{ij}, r_k, \rho$, and i, we have $\hat{f}^{\{i\}}(\underline{t}) = A\underline{y}^*$

Proof :

Let $f(\underline{t}) = \{f_1(t_{11}), ..., f_1(t_{1n_1}), f_2(t_{21}), ..., f_2(t_{2n_2}), ..., f_p(t_{p1}), ..., f_p(t_{pn_p})\}$ and $\hat{f}^{\{i\}}(\underline{t}) = (\hat{f}_1^{\{i\}}(t_{11}), ..., \hat{f}_1^{\{i\}}(t_{1n_1}), \hat{f}_2^{\{i\}}(t_{21}), ..., \hat{f}_2^{\{i\}}(t_{2n_2}), ..., \hat{f}_p^{\{i\}}(t_{p1}), ..., \hat{f}_p^{\{i\}}(t_{pn_p}))$. Similarly define $f(\underline{t}^{[i]})$ and $\hat{f}^{\{i\}}(\underline{t}^{[i]})$ as $f(\underline{t})$ and $\hat{f}^{\{i\}}(\underline{t})$ respectively without the elements corresponding to the i^{th} pair. For any function $f_1, f_2, ..., f_p$ in W_2, we have :

$$\left(\underline{y}*-f\left(\underline{t}\right)\right)^T W\left(\underline{y}*-f\left(\underline{t}\right)\right)+\lambda_1\int_0^1\left(f_1''(t)\right)^2 dt+$$
$$\lambda_2\int_0^1\left(f_2''(t)\right)^2 dt+\ldots+\lambda_p\int_0^1\left(f_p''(t)\right)^2 dt$$
$$\geq\left(\underline{y}^{[i]}-f\left(\underline{t}^{[i]}\right)\right)^T W^{[i]}\left(\underline{y}^{[i]}-f\left(\underline{t}^{[i]}\right)\right)+$$
$$\lambda_1\int_0^1\left(f_1''(t)\right)^2 dt+\ldots+\lambda_p\int_0^1\left(f_p''(t)\right)^2 dt$$
$$\geq\left(\underline{y}^{[i]}-\hat{f}^{\{i\}}\left(\underline{t}^{[i]}\right)\right)^T W^{[i]}\left(\underline{y}^{[i]}-\hat{f}^{\{i\}}\left(\underline{t}^{[i]}\right)\right)+$$
$$\lambda_1\int_0^1\left(\left(\hat{f}_1^{\{i\}}(t)\right)''\right)^2 dt+\ldots+\lambda_p\int_0^1\left(\left(\hat{f}_p^{\{i\}}(t)\right)''\right)^2 dt$$
$$=\left(\underline{y}*-\hat{f}^{\{i\}}\left(\underline{t}^{[i]}\right)\right)^T W^{[i]}\left(\underline{y}*-\hat{f}^{\{i\}}\left(\underline{t}^{[i]}\right)\right)+$$
$$\lambda_1\int_0^1\left(\left(\hat{f}_1^{\{i\}}(t)\right)''\right)^2 dt+\ldots+\lambda_p\int_0^1\left(\left(\hat{f}_p^{\{i\}}(t)\right)''\right)^2 dt \quad (13)$$

where the first inequality holds because after switching rows and columns, we have

$$\left(\underline{y}*-f\left(\underline{t}\right)\right)^T W\left(\underline{y}*-f\left(\underline{t}\right)\right)=$$
$$\begin{pmatrix}\underline{y}*^{[i]}-f\left(\underline{t}^{[i]}\right)\\ \underline{y}*^{(i)}-f\left(\underline{t}^{(i)}\right)\end{pmatrix}^T\begin{pmatrix}W^{[i]} & 0\\ 0 & W^{(i)}\end{pmatrix}\begin{pmatrix}\underline{y}*^{[i]}-f\left(\underline{t}^{[i]}\right)\\ \underline{y}*^{(i)}-f\left(\underline{t}^{(i)}\right)\end{pmatrix}$$
$$\geq\left(\underline{y}^{[i]}-f\left(\underline{t}^{[i]}\right)\right)^T W^{[i]}\left(\underline{y}^{[i]}-f\left(\underline{t}^{[i]}\right)\right).$$

The second inequality holds because $\hat{f}_1^{\{i\}},\ldots,\hat{f}_p^{\{i\}}$ are solutions to (12). The last equality holds because of the definition of $\underline{y}*$. The inequality at (13) indicates that $\hat{f}_1^{\{i\}},\ldots,\hat{f}_p^{\{i\}}$ are solutions to (3) with $\underline{y}$ replaced by $\underline{y}*$. Therefore $\hat{f}^{\{i\}}\left(\underline{t}\right)=A\underline{y}*$.

As a consequence of this lemma, we do not need to solve separate minimization problems (12) for each deleting-one-pair set. All we need to do is to solve the following equations

$$\begin{bmatrix} m_{11} & m_{12} & \ldots & m_{1p}\\ m_{21} & m_{22} & \ldots & m_{2p}\\ . & . & \ldots & .\\ . & . & \ldots & .\\ . & . & \ldots & .\\ m_{p1} & m_{p2} & \ldots & m_{pp}\end{bmatrix}\begin{bmatrix} s_1\\ s_2\\ .\\ .\\ .\\ s_p\end{bmatrix}=\begin{bmatrix} u_1\\ u_2\\ .\\ .\\ .\\ u_p\end{bmatrix}$$

(14)

for $\hat{f}_k^{\{i\}}\left(t_{ki_k}\right)-y_{ki_k}$, where $m_{11}=1-a\left(i_1,i_1\right)$; $m_{12}=-a\left(i_1,n_1+i_2\right);\ldots;m_{1p}=-a(i_1,n_1+i_p)$; $m_{21}=-a\left(n_1+i_2,i_1\right)$; $m_{22}=1-a\left(n_1+i_2,n_1+i_2\right);\ldots;$ $m_{2p}=-a\left(i_2,n_2+i_p\right);\ldots;m_{p1}=-a\left(n_1+i_p,i_1\right)$ $m_{p2}=-a\left(n_2+i_p,i_2\right);\ldots;$ $m_{pp}=1-a\left(n_{p-1}+i_p,n_{p-1}+i_p\right)$ $s_1=\hat{f}_1^{\{i\}}\left(t_{1i_1}\right)-y_{1i_1}$; $s_2=\hat{f}_2^{\{i\}}\left(t_{2i_2}\right)-y_{1i_2};\ldots;$ $s_p=\hat{f}_p^{\{i\}}\left(t_{pi_p}\right)-y_{pi_p}$; $u_1=f_1^{\{i\}}\left(t_{1i_1}\right)-y_{1i_1}$; $u_2=f_2^{\{i\}}\left(t_{2i_2}\right)-y_{1i_2}$; … ; $u_p=f_p^{\{i\}}\left(t_{pi_p}\right)-y_{pi_p}$; and $a\left(i,j\right)$ are elements of the matrix A. If there is only one observation in the i^{th} pair, for example y_{1i_1}, we then have the following equation

$$\left(1-a\left(i_1,i_1\right)\right)\left(\hat{f}_1^{\{i\}}\left(t_{1i_1}\right)-y_{1i_1}\right)=\hat{f}_1\left(t_{1i_1}\right)-y_{1i_1} \quad (15)$$

Note that (15) is exactly the same as the "leaving-out-one" lemma in the independent case.

Let $\underline{\hat{f}}_k^{\{-\}}=\left(\hat{f}_k^{\{i_{k1}\}}\left(t_{k1}\right),\ldots,\hat{f}_k^{\{i_{kn_k}\}}\left(t_{kn_k}\right)\right)^T$ and $\underline{\hat{f}}^{\{-\}}=\left(\left(\underline{\hat{f}}_1^{\{-\}}\right)^T,\ldots,\left(\underline{\hat{f}}_p^{\{-\}}\right)^T\right)$, where i_{kj} denotes the index of the pair for observation y_{kj}. Define the cross validation score as

$$C\left(\lambda_k,\gamma_{ij},r_k,\rho\right)=\frac{1}{n}\left\|W\left(\underline{y}-\underline{\hat{f}}^{\{-\}}\right)\right\|^2 \quad (16)$$

Here, C estimates the weighted mean-square errors (WMSE) (Wang 1998). The minimizers of $C\left(\lambda_k,\gamma_{ij},r_k,\rho\right)$ are called cross validation estimates of the parameters.

CONCLUSION

The distribution of vector responses $\underline{y}$ is Multivariate Normal with mean $\underline{f}$ and variance θW^{-1}. General smoothing spline models provide flexibility for estimating nonparametric functions and are widely used in many areas. With multiple correlated responses it is better to estimate these functions jointly using the penalized weighted least-squares.

DAFTAR PUSTAKA

Aydin D. 2007. A Comparison of the Nonparametric Regression Models Using Smoothing Spline and Kernel Regression. *Proceedings of World Academy of Science, Engineering and Technology*. ISSN 1307-6884. Volume **26**: 730-734.

Budiantara IN, Subanar & Soejoeti Z. 1997. Weighted Spline Estimator. *Proc. 51st Session of the International Statistical Institute*, Istanbul : 333-334.

Cox DD 1983. Assymptotic for M-type Smoothing Spline. *The Annals. of Statistics*. **11**: 530-551.

Cox DD & O'Sullivan F. 1996. Penalized Likelihood Type Estimators for Generalized Nonparametric Regression. *J. Mult. Anal*. **56**: 185-206.

Craven P & 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.

Fernandez MF & Opsomer JD. 2005. Smoothing Parameter Selection Methods for Nonparametric Regression With Spatially Correlated Errors. *The Canadian Journal of Statistics*. **33 (2)**: 279-295.

Flessler JA. 1991. Nonparametric Fixed-Interval Smoothing With Vector Splines. *IEEE Trans. Signal Processing*. **39**:852 – 859.

Gallant AR. 1975. Seemingly Unrelated Nonlinear Regression, *J. Econometrics*. **3**: 35-50.

Gooijer JGD, Gannoun A & Larramendy I. 1991. *Nonparametric Regression With Serially Correlated Errors*. http://www.tinbergen.nl/discussionpapers/99063.pdf [12 January 2009].

Kimeldorf G & Wahba G. 1971. Some Result on Tchebycheffian Spline Functions. *J. of Math. Anal. And Applications.* **33**: 82-95.

Koenker RNgP & Portnoy S. 1994. Quantile Smoothing Splines. *Biometrics*. **81:** 673-680.

Lestari B. 2007. *Estimation of Three Responses Nonparametric Regression Models.* Paper for 8th National Seminar in Statistics (SNS VIII), November 3rd, 2007, Department of Statistics, Sepuluh Nopember Institute of Technology, Surabaya.

Lestari B. 2008a. Spline Estimator of Biresponse Nonparametric Regression Model with Unequal Variances of Errors. *Jurnal Penelitian Matematika dan Sains, FMIPA Universitas Negeri Surabaya*. **15** (2): 85-93.

Lestari B. 2008b. Penalized Weighted Least-squares Estimator for Bivariate Nonparametric Regression Model with Correlated Errors. *Prosiding Seminar Nasional Matematika dan Statistika 2008, Fak-SainsTeks UNAIR*, ISBN : 978-979-19096-0-0: 83-95.

Miller JJ & Wegman EJ. 1987. Vector Function Estimation Using Splines. *Journal of the Royal Statistical Society*, Series B. **17**: 173-180.

Oehlert GW. 1992. Relaxed Boundary Smoothing Spline. *The Annals. of Statistics.* **20**: 146-160.

Wegman EJ. 1981. Vector Spline and Estimation of filter Function. *Technometrics*. **23**: 83-89.

Wahba G. 1983. Bayesian "Confidence Intervals" for the Cross-Validated Smoothing Spline, *Journal of the Royal Statistical Society*, Series B. **45**: 133-150.

Wahba G. 1990. *Spline Models for Observational Data*. CBMS-NSF Regional Conference Series in Applied Mathematics. Volume 59, SIAM, Philadelphia.