Spline Estimator in Multi-Response Nonparametric Regression Model with Unequal Correlation of Errors

[1]Budi Lestari, [2]I. Nyoman Budiantara, [2]Sony Sunaryo and [2]Muhammad Mashuri
[1]Department of Mathematics, Faculty of Mathematics and Natural Sciences,
University of Jember, Jalan Kalimantan III/25 Kampus Tegalboto, Jember 68121, Indonesia
[2]Department of Statistics, Faculty of Mathematics and Natural Sciences,
Sepuluh Nopember Institute of Technology, Surabaya, Indonesia

Abstract: Problem statement: 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 provide powerful tools to model the functions which draw association of these variables. **Approach:** Penalized weighted least-squares was used to jointly estimate nonparametric functions from contemporaneously correlated data. We apply Generalized Maximum Likelihood (GML), Generalized Cross Validation (GCV) and leaving-out-one-pair Cross Validation (CV) for estimating the smoothing parameters, the weighting parameters and the correlation parameter **Results:** In this study we formulated the multi-response nonparametric regression model with unequal correlation of errors 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 GML, GCV and CV. **Conclusion:** Distribution of responses is normal. With multiple correlated responses it is better to estimate these functions jointly using the penalized weighted least-squares.

Key words: Multi-response nonparametric regression model, penalized weighted least-squares, generalized maximum likelihood, generalized cross validation, leaving-out-one-pair cross validation

INTRODUCTION

The functions which draw association of two or more dependent variables are observed at several values of the independent variables, such as at multiple time points, can be modeled by using smoothing spline. There are many writers who have studied spline estimators for estimating regression curve of nonparametric regression models. Kimeldorf and Wahba (1971); Craven and Wahba (1979) and Wahba (1985) proposed original spline estimator to estimates regression curve of smooth data. Cox (1983) and Cox and 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. Wahba (2000) compared between GCV and GML for choosing the smoothing parameter in the generalized spline smoothing problem. Oehlert (1992) and Koenker *et al.*

(1994) introduced relaxed spline and quantile spline, respectively. Budiantara *et al.* (1997) studied weighted spline estimator in nonparametric regression model with different variance. Wahba (1992) introduced some techniques for spline statistical model building by using reproducing kernel Hilbert spaces. Cardot *et al.* (2007) studied asymptotic property of smoothing splines estimators in functional linear regression with errors-in-variables. Liu *et al.* (2007) proposed smoothing spline estimation of variance functions. 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 multiple time points. Multi-response

Corresponding Author: Budi Lestari, Department of Mathematics, Faculty of Mathematics and Natural Sciences,
University of Jember, Jalan Kalimantan III/25 Kampus Tegalboto, Jember 68121, Indonesia

nonparametric regression model provide powerful tools to model the functions which draw association of these variables.

Many authors have considered nonparametric models for multi-response data. Wegman (1981); Miller and Wegman (1987) and Flessler (1991) proposed algorithms for spline smoothing. Wahba (1990) developed the theory of general smoothing splines using reproducing kernel Hilbert spaces. Gooijer *et al.* (1991) and Francisco-Fernandez and 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 (2008a) developed spline estimator in biresponse nonparametric regression model with unequal variances of errors, Lestari (2008b) studied penalized weighted least-squares estimator for bivariate nonparametric regression model with correlated errors, Lestari *et al.* (2010a; 2010b) proposed spline approach for estimating regression function of multi-response nonparametric regression model in special cases, i.e., variances and correlations of errors are the same for every response.

All, except Wang *et al.* (2000); Lestari (2008a; 2008b) and Lestari *et al.* (2010a; 2010b), 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 study, 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. In addition, errors have different correlation. 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.

MATERIALS AND METHODS

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 ith response of the kth variable y_{ki} is generated by the kth function f_k evaluated at the design point t_{ki} plus a random error ε_{ki}. Assume $\varepsilon_{ki} \overset{i.i.d.}{\sim} N(0, \sigma_k^2)$ for fixed k = 1, 2,...,p and the correlation of errors are different for every i = 1, 2,...,n_k, i.e., Corr($\varepsilon_{ki}, \varepsilon_{ki}$) = ρ_i for k ≠ l and zero otherwise. This correlation assumption means that correlation of errors are the same for every response.

There are four cases for this correlation assumption, i.e., (i) case: $n_1 < n_2 < ... < n_p$, (ii) case: $n_1 > n_2 > ... > n_p$, (iii) case: $n_1 = n_2 = ... = n_p$ and (iv) case: not all $n_i < n_{i+1}$ (for i = 1, 2,...,p-1), not all $n_i > n_{i+1}$ (for i = 1, 2,...,p-1) and not all $n_i = n_{i+1}$ (for i = 1, 2,...,p-1). In this study, we describe for the first case only, i.e., $n_1 < n_2 < ... < n_p$. The three other cases can be verified similarly. In addition, 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 functional instead of evaluations (Wahba, 1985; 1990).

Distribution of the responses: Based on the model (1), 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. Moreover, we may write model (1) in the vector equation as follows:

$$\underline{y} = \underline{f} + \underline{\varepsilon} \qquad (2)$$

Next, for k = 1, 2,...,p.

Let $r_k = \dfrac{\sigma_k}{m}$, $m = \prod\limits_{\substack{j=1 \\ j \neq k}}^{p} \sigma_j$; $\theta = \prod\limits_{k=1}^{p} \sigma_k$;

$\gamma_{ij} = \dfrac{1}{\sigma_k}(i, j = 1, 2,...,p; k \neq i \neq j)$ and Ω_{qs} be a $n_q \times n_s$ matrix with $(i, j)^{th}$ element equal to $\rho_i (i = 1, 2,...,n_q)$ if the ith element of $\underline{y}_q$ and the jth element of $\underline{y}_s$ is a pair and

zero otherwise. Therefore, by taking $E(\underline{y})$ and $\text{Var}(\underline{y})$ in (2), we obtain the distribution of responses, i.e., Multivariate Normal with mean $\underline{f}$ and variance θW^{-1}. It can be written as $\underline{y} \sim N(\underline{f}, \theta W^{-1})$, where:

$$W^{-1} = \begin{bmatrix} r_1 I_{n_1} & \gamma_{12}\Omega_{12} & \cdots & \gamma_{1p}\Omega_{1p} \\ \gamma_{12}\Omega_{12}^T & r_2 I_{n_2} & \cdots & \gamma_{2p}\Omega_{2p} \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \gamma_{1p}\Omega_{1p}^T & \gamma_{2p}\Omega_{2p}^T & \cdots & r_p I_{n_p} \end{bmatrix} \quad (3)$$

RESULTS AND DISCUSSION

Estimation nonparametric function of multi-response nonparametric regression model: The nonparametric functions f_k are estimated by carrying out the following penalized weighted least-squares:

$$\begin{aligned} &\underset{f_1,f_2,\ldots,f_p \in W_2}{\text{Min}} \{(\underline{y}-\underline{f})^T W(\underline{y}-\underline{f}) + \\ &\lambda_1 \int_0^1 (f_1''(t))^2 dt + \lambda_2 \int_0^1 (f_2''(t))^2 dt + \ldots + \\ &\lambda_p \int_0^1 (f_p''(t))^2 dt\} \end{aligned} \quad (4)$$

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 with unequal correlation of errors. Let $\phi_v(t) = t^{v-1}/(v-1)!$ for v = 1, 2; $R^1(s,t) = k_2(s)k_2(t)-k_4(s-t)$ where $k_v(.) = B_v(.)/v$ and $B_v(.)$ is the vth Bernoulli polynomial. Let $G_k = \{\phi_v(t_{ki})\}_{i=1,v=1}^{n_k,2}$; $G = \text{diag}(G_1,\ldots,G_p)$;

$\Sigma_k = \{R^1(t_{ki},t_{kj})\}_{i=1,j=1}^{n_k,n_k}$ and $\Sigma = \text{diag}(\Sigma_1,\Sigma_2,\ldots,\Sigma_p)$. By extending method as in both Wang (1998) and Wahba (1985) to the case of multi-response with unequal correlation of errors, we can show that for fixed λ_k, γ_{ij} and ρ_i, the solution to (4) is:

$$\hat{f}_k(t) = \sum_{v=1}^2 d_{kv}\phi_v(t) + \sum_{i=1}^{n_k} c_{ki}R^1(t,t_{ki}) \quad (5)$$

where, k = 1, 2,…p and $\underline{c} = (c_{11},\ldots,c_{1n_1},c_{21},\ldots,c_{2n_2},\ldots,c_{p1},\ldots,c_{pn_p})^T$; $\underline{d} = (d_{11},\ldots,d_{1p}, d_{21},\ldots,d_{2p},\ldots,d_{p1},\ldots,d_{pp})^T$ are solutions to:

$$\begin{pmatrix} G^T W G & G^T W\Sigma \\ \Sigma W G & \Sigma W\Sigma + \text{diag}(\lambda_1\Sigma_1,\ldots,\lambda_p\Sigma_p) \end{pmatrix} \begin{pmatrix} \underline{d} \\ \underline{c} \end{pmatrix}$$
$$= \begin{pmatrix} G^T W\underline{y} \\ \Sigma W\underline{y} \end{pmatrix} \quad (6)$$

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

$$\left. \begin{aligned} \left(\Sigma + W^{-1}\text{diag}\left(\lambda_1 I_{n_1},\ldots,\lambda_p I_{n_p}\right)\right)\underline{c} + G\underline{d} = \underline{y} \\ G^T\underline{c} = 0 \end{aligned} \right\} \quad (7)$$

is also a solution to (6). Thus we need to solve simultaneous Eq. 7 for $\underline{c}$ and $\underline{d}$. In fact, $W^{-1}\text{diag}\left(\lambda_1 I_{n_1},\ldots,\lambda_p I_{n_p}\right)$ is asymmetric if $\lambda_1 \neq \lambda_2 \neq \lambda_p$ and $\rho_i \neq 0 (i = 1, 2,\ldots,n_k)$. To calculate the coefficients $\underline{c}$ and $\underline{d}$, we use the following transformations: $\tilde{\Sigma} = \Sigma\text{diag}\left(I_{n_1}/\lambda_1, I_{n_2}/\lambda_2,\ldots,I_{n_p}/\lambda_p\right)$ and $\tilde{c} = \text{diag}\left(\lambda_1 I_{n_1},\ldots,\lambda_p I_{n_p}\right)\underline{c}$. Then (7) is equivalent to:

$$\left. \begin{aligned} (\tilde{\Sigma} + W^{-1})\tilde{\underline{c}} + G\underline{d} = \underline{y} \\ G^T\tilde{\underline{c}} = 0 \end{aligned} \right\} \quad (8)$$

Let $G_k = (Q_{k1} \quad Q_{k2})\begin{pmatrix} R_k \\ 0 \end{pmatrix}$, k = 1, 2,…,p be the QR decompositions. Let $Q_1 = \text{diag}(Q_{11},Q_{21},Q_{31},\ldots,Q_{p1})$; $Q_2 = \text{diag}(Q_{12},Q_{22},Q_{32},\ldots,Q_{p2})$; $R = \text{diag}(R_1,R_2,\ldots,R_p)$ and $B = \tilde{\Sigma} + W^{-1}$. It can be shown that the solutions to (8) are:

$$\begin{aligned} \tilde{\underline{c}} = Q_2(Q_2^T B Q_2)^{-1} Q_2^T \underline{y} \\ R\underline{d} = Q_1^T(\underline{y} - B\tilde{\underline{c}}) \end{aligned} \quad (9)$$

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

$$A = I - W^{-1}Q_2(Q_2^T B Q_2)^{-1}Q_2^T \quad (10)$$

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} and ρ_i, 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 study 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_i 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 Wang (1998); Wang *et al.* (2000) and Lestari *et al.* (2010a; 2010b), to the case of multi-response with unequal correlation of errors as follows.

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

$$
\begin{aligned}
M(\lambda_k, \gamma_{ij}, r_k, \rho_i) &= \frac{\underline{y}^T W(I-A)\underline{y}}{\left[\det^+(W(I-A))\right]^{\frac{1}{n-4}}} \\
&= \frac{\underline{z}^T (Q_2^T B Q_2)^{-1} \underline{z}}{\left[\det(Q_2^T B Q_2)^{-1}\right]^{\frac{1}{n-4}}}
\end{aligned}
\tag{11}
$$

Where:
n $= n_1+n_2+\ldots+n_p$
$\det^+$ = The product of the nonzero eigen values
$\underline{z}$ $= Q_2^T \underline{y}$

The minimizers of M (λ_k, γ_{ij}, r_k, ρ_i) are called GML estimates.

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

$$
\begin{aligned}
V(\lambda_k, \gamma_{ij}, r_k, \rho_i) &= \frac{\left\| W(I-A)\underline{y} \right\|^2}{\left[\mathrm{tr}(W(I-A)) \right]^2} \\
&= \frac{\underline{z}^T (Q_2^T B Q_2)^{-2} \underline{z}}{\left[\mathrm{tr}(Q_2^T B Q_2)^{-1} \right]^2}
\end{aligned}
\tag{12}
$$

In the following we propose a cross validation method based on leaving-out-one-pair procedure. Suppose there are a total of N ($N \geq \max\{n_1, n_2,\ldots,n_p\}$) 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 ith pair; superscripts [i] to denote the collection of elements after deleting the ith 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:

$$
\begin{aligned}
&\underset{f_1,\ldots,f_p \in W_2}{\mathrm{Min}} \{ (\underline{y}^{[i]} - \underline{f}^{[i]})^T W^{[i]} (\underline{y}^{[i]} - \underline{f}^{[i]}) + \\
&\lambda_1 \int_0^1 (f_1''(t))^2 dt + \lambda_2 \int_0^1 (f_2''(t))^2 dt +\ldots+ \\
&\lambda_p \int_0^1 (f_p''(t))^2 dt \}
\end{aligned}
\tag{13}
$$

are $\hat{f}_1^{\{i\}}$, $\hat{f}_2^{\{i\}}$, ..., $\hat{f}_p^{\{i\}}$. Assume that there are p elements in the ith pair (it is simple if there is only one). Denotes 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:

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

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

Lemma: For fixed λ_k, γ_{ij}, r_k, ρi and i, we have $\hat{f}^{\{i\}}(\underline{t}) = A\underline{y}*$.

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

$$\left(\underline{y}*-f\left(\underline{t}\right)\right)^{\mathrm{T}} W\left(\underline{y}*-f\left(\underline{t}\right)\right)+\lambda_1\int_0^1\left(f_1''(t)\right)^2 dt +$$

$$\lambda_2\int_0^1\left(f_2''(t)\right)^2 dt +...+\lambda_p\int_0^1\left(f_p''(t)\right)^2 dt$$

$$\geq\left(\underline{y}^{[i]}-f\left(\underline{t}^{[i]}\right)\right)^{\mathrm{T}} W^{[i]}\left(\underline{y}^{[i]}-f\left(\underline{t}^{[i]}\right)\right)+$$

$$\lambda_1\int_0^1\left(f_1''(t)\right)^2 dt +...+\lambda_p\int_0^1\left(f_p''(t)\right)^2 dt$$

$$\geq\left(\underline{y}^{[i]}-\hat{f}^{\{i\}}\left(\underline{t}^{[i]}\right)\right)^{\mathrm{T}} W^{[i]}\left(\underline{y}^{[i]}-\hat{f}^{\{i\}}\left(\underline{t}^{[i]}\right)\right)+ \qquad (14)$$

$$\lambda_1\int_0^1\left(\left(\hat{f}_1^{\{i\}}(t)\right)''\right)^2 dt +...+\lambda_p\int_0^1\left(\left(\hat{f}_p^{\{i\}}(t)\right)''\right)^2 dt$$

$$=\left(\underline{y}*-\hat{f}^{\{i\}}\left(\underline{t}^{[i]}\right)\right)^{\mathrm{T}} W^{[i]}\left(\underline{y}*-\hat{f}^{\{i\}}\left(\underline{t}^{[i]}\right)\right)+$$

$$\lambda_1\int_0^1\left(\left(\hat{f}_1^{\{i\}}(t)\right)''\right)^2 dt +...+\lambda_p\int_0^1\left(\left(\hat{f}_p^{\{i\}}(t)\right)''\right)^2 dt$$

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

$$\left(\underline{y}*-f\left(\underline{t}\right)\right)^{\mathrm{T}} W\left(\underline{y}*-f\left(\underline{t}\right)\right)=$$

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

$$\geq\left(\underline{y}^{[i]}-f\left(\underline{t}^{[i]}\right)\right)^{\mathrm{T}} W^{[i]}\left(\underline{y}^{[i]}-f\left(\underline{t}^{[i]}\right)\right)$$

The second inequality holds because $\hat{f}_1^{\{i\}},...,\hat{f}_p^{\{i\}}$ are solutions to (13). The last equality holds because of the definition of $\underline{y}*$. The inequality at (14) indicates that $\hat{f}_1^{\{i\}},...,\hat{f}_p^{\{i\}}$ are solutions to (4) 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 (13) for each deleting-one-pair set. All we need to do is to solve the following equations:

$$\begin{bmatrix}m_{11} & m_{12} & ... & m_{1p}\\ m_{21} & m_{22} & ... & m_{2p}\\ . & . & ... & .\\ . & . & ... & .\\ . & . & ... & .\\ m_{p1} & m_{p2} & ... & m_{pp}\end{bmatrix}\begin{bmatrix}s_1\\ s_2\\ .\\ .\\ .\\ s_p\end{bmatrix}=\begin{bmatrix}u_1\\ u_2\\ .\\ .\\ .\\ u_p\end{bmatrix} \qquad (15)$$

for $\hat{f}_k^{\{i\}}\left(t_{ki_k}\right)-y_{ki_k}$, where $m_{11}=1-a\left(i_1,i_1\right)$; $m_{12}=-a\left(i_1,n_1+i_2\right);...;m_{1p}=-a(i_1,n_1+i_p)$; $m_{21}=$

$-a\left(n_1+i_2,i_1\right)$; $\qquad m_{22}=1-a\left(n_1+i_2,n_1+i_2\right);...$;

$m_{2p}=-a\left(i_2,n_2+i_p\right);...;$ $\qquad m_{p1}=-a\left(n_1+i_p,i_1\right)$

$m_{p2}=-a\left(n_2+i_p,i_2\right);...$; $\qquad m_{pp}=1-a\left(n_{p-1}+i_p,n_{p-1}+i_p\right)$

$s_1=\hat{f}_1^{\{i\}}\left(t_{1i_1}\right)-y_{1i_1}$; $\qquad s_2=\hat{f}_2^{\{i\}}\left(t_{2i_2}\right)-y_{1i_2};...$;

$s_p=\hat{f}_p^{\{i\}}\left(t_{pi_p}\right)-y_{pi_p}$; $u_1=f_1^{\{i\}}\left(t_{1i_1}\right)-y_{1i_1}$; $u_2=f_2^{\{i\}}\left(t_{2i_2}\right)-y_{1i_2}$;

... ; $u_p=f_p^{\{i\}}\left(t_{pi_p}\right)-y_{pi_p}$ and a(i, j) are elements of the matrix A. If there is only one observation in the ith pair, for example y_{1i_1}, we then have the following equation:

$$\left(1-a\left(i_1,i_1\right)\right)\left(\hat{f}_1^{\{i\}}\left(t_{1i_1}\right)-y_{1i_1}\right)=\hat{f}_1\left(t_{1i_1}\right)-y_{1i_1} \qquad (16)$$

Note that (16) is exactly the same as the "leaving-out-one" lemma in the independent case.

Let $\qquad \hat{f}_k^{\{-\}}=\left(\hat{f}_k^{\{i_{k1}\}}\left(t_{k1}\right),...,\hat{f}_k^{\{t_{kn_k}\}}\left(t_{kn_k}\right)\right)^{\mathrm{T}}$ $\qquad$ and

$\underline{\hat{f}}^{\{-\}}=\left(\left(\underline{\hat{f}}_1^{\{-\}}\right)^{\mathrm{T}},...,\left(\underline{\hat{f}}_p^{\{-\}}\right)^{\mathrm{T}}\right)$, where i_{kj} denotes the index of the pair for observation y_{kj}. Define the cross validation score as:

$$C\left(\lambda_k,\gamma_{ij},r_k,\rho_i\right)=\frac{1}{n}\left\|W\left(\underline{y}-\underline{\hat{f}}^{\{-\}}\right)\right\|^2 \qquad (17)$$

Here, C estimates the Weighted Mean-Square Errors (WMSE) (Wang, 1998; Wang *et al.*, 2000). The minimizers of $C((\lambda_k,~\gamma_{ij},~r_k,~\rho_i)$ are called cross validation estimates of the parameters.

CONCLUSION

The distribution of vector of 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.

REFERENCES

Aydin, D., 2007. A comparison of the nonparametric regression models using smoothing spline and kernel regression. Proc. World Acad. Sci. Eng. Technol., 36: 253-257. http://www.waset.org/journals/waset/v36/v36-46.pdf

Budiantara, I.N., Subanar and Z. Soejoeti, 1997. Weighted spline estimator. Bull. Int. Stat. Inst., Istanbul, 51: 333-334.

Cardot, H., C. Crambes, A. Kneip and P. Sarda, 2007. Smoothing splines estimators in functional linear regression with errors-in-variables. Comput. Stat. Data Anal., 51: 4832-4848. DOI: 10.1016/j.csda.2006.07.029

Cox, D.D. and F. O'Sullivan, 1996. Penalized likelihood type estimators for generalized nonparametric regression. J. Multivar. Anal., 56: 185-206. http://ideas.repec.org/a/eee/jmvana/v56y1996i2p185-206.html

Cox, D.D., 1983. Asymptotic for M-type smoothing spline. Ann. Stat., 11: 530-551. DOI: 10.1214/aos/1176346159

Craven, P. and G. Wahba, 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. DOI: 10.1007/BF01404567

Francisco-Fernandez, M. and J.D. Opsomer, 2005. Smoothing parameter selection methods for nonparametric regression with spatially correlated errors. Can. J. Stat., 33: 279-295. http://www.jstor.org/stable/25046177

Flessler, J.A., 1991. Nonparametric fixed-interval smoothing with vector splines. IEEE Trans. Sig. Process., 39: 852-859. DOI: 10.1109/78.80907

Gooijer, J.G.D., A. Gannoun and I. Larramendy, 1999. Nonparametric regression with serially correlated errors. Tinbergen. http://ftp.tinbergen.nl/discussionpapers/99063.pdf

Kimeldorf, G. and G. Wahba, 1971. Some result on tchebycheffian spline functions. J. Math. Anal. Appli., 33: 82-95. DOI: 10.1016/0022-247X(71)90184-3

Koenker, R., N.G. Pin and S. Portnoy, 1994. Quantile smoothing splines. Biometrics, 81: 673-680. DOI: 10.1093/biomet/81.4.673

Lestari, B., 2008a. Spline estimator of biresponse nonparametric regression model with unequal variances of errors. J. Penelitian Math., 15: 85-93.

Lestari, B., 2008b. Penalized weighted least-squares estimator for bivariate nonparametric regression model with correlated errors. Proceeding of the National Seminar on Mathematics and Statistics, (MS'08), Airlangga University, Surabaya, pp: 83-95.

Lestari, B., I.N. Budiantara, S. Sunaryo and M. Mashuri, 2010a. Spline estimator in homoscedastic multi-response nonparametric regression model. Proceeding of the Indo MS International Conference on Mathematics and Its Application, Oct. 12-13, Yogyakarta, Indonesia, pp: 845-854.

Lestari, B., I.N. Budiantara, S. Sunaryo and M. Mashuri, 2010b. Spline estimator in multi-response nonparametric regression model. J. Basic Sci., 11: 17-22.

Liu, A., T. Tong and Y. Wang, 2007. Smoothing spline estimation of variance functions. J. Comput. Graph. Stat., 16: 312-329. DOI: 10.1198/106186007X204528

Miller, J.J. and E.J. Wegman, 1987. Vector function estimation using splines. J. Stat. Plann. Inferen., 17: 173-180. DOI: 10.1016/0378-3758(87)90110-8

Oehlert, G.W., 1992. Relaxed boundary smoothing spline. Ann. Stat., 20: 146-160. DOI: 10.1214/aos/1176348516

Wahba, G., 1983. Bayesian "confidence intervals" for the cross-validated smoothing spline. J. R. Stat. Soc. Ser. B, 45: 133-150. http://www.jstor.org/stable/2345632

Wahba, G., 1985. A comparison of GCV and GML for choosing the smoothing parameter in the generalized spline smoothing problem. The Ann. Stat., 13: 1378-1402. http://www.jstor.org/stable/2241360

Wahba, G., 1990. Spline Models for Observational Data. 1st Edn., SIAM, Philadelphia, Pennsylvania, ISBN: 10: 0898712440. pp: 180.

Wahba, G., 1992. Multivariate Function and Operator Estimation, Based on Smoothing Spline and Reproducing Kernels. In: Nonlinear Modeling and Forecasting, SFI Studies in the Science of Complexity, Casdagli, M. and S. Eubank (Eds.). Addison-Wesley, USA., pp: 95-112.

Wahba, G., 2000. An introduction to model building with reproducing Kernel Hilbert spaces. University of Wisconsin. http://www.stat.wisc.edu/~wahba/ftp1/interf/rootpar1r.pdf

Wang, Y., 1998. Smoothing spline models with correlated random errors. J. Am. Stat. Assoc., 93: 341-348. http://www.jstor.org/stable/2669630

Wang, Y., W. Guo and M.B. Brown, 2000. Spline smoothing for bivariate data with applications to association between hormones. Statistica Sinica, 10: 377-397. http://www3.stat.sinica.edu.tw/statistica/oldpdf/A10n23.pdf

Wegman, E.J., 1981. Vector spline and estimation of filter function. Technometrics, 23: 83-89. http://www.jstor.org/stable/1267980