Reproducing Kernel Hilbert Space and Penalized Weighted Least Square in Nonparametric Regression

Adji Achmad Rinaldo Fernandes
Department of Mathematics
Faculty of Mathematics and Natural Sciences
University of Brawijaya
Jalan Veteran Malang-Indonesia

I Nyoman Budiantara, Bambang Widjanarko Otok, Suhartono
Department of Statistics
Faculty of Mathematics and Natural Sciences
Sepuluh Nopember Institute of Technology
Jalan Arif Rahman Hakim Surabaya-Indonesia

Copyright © 2014 Adji Achmad Rinaldo Fernandes, I Nyoman Budiantara, Bambang Widjanarko Otok and Suhartono. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract
Reproducing Kernel Hilbert Space (RKHS) play a central role to solve the Penalized Weighted Least Square (PWLS) in Spline Estimator of nonparametric regression analysis. The purposes of this research is to obtain the RKHS approach in PWLS to solve the estimator of regression curve. Base of RKHS, the curve nonparametric regression form is \( f(x) = \mathbf{Td} + \mathbf{V}\zeta \). Solving the weighted of PLWS coming from variance-covariance \( \hat{W} \) is equals to solving the \( \hat{\Sigma}_{11,1}, \hat{\Sigma}_{11,2}, \ldots, \hat{\Sigma}_{22,1}, \ldots, \hat{\Sigma}_{12,rr} \). For the purposes of \( \hat{f} \) estimation, RKHS approach with completes the PWLS criterion is \( \hat{f}_\alpha = \mathbf{A}^* (\hat{\zeta}) y \), with
\[
\mathbf{A}(\hat{\zeta}) = \mathbf{T} (\mathbf{T}'\mathbf{M}^{-1}\hat{W})^{-1} \mathbf{T}'\mathbf{M}^{-1}\hat{W} + \mathbf{VM}^{-1}\hat{W}[\mathbf{I} - \mathbf{T}(\mathbf{T}'\mathbf{M}^{-1}\hat{W})^{-1}\mathbf{T}'\mathbf{M}^{-1}\hat{W}]^{-1}
\]
Keywords: Reproducing Kernel, Hilbert Space, PWLS

1 Introduction

One of the uses of regression analysis is in the analysis of longitudinal data, which is a combination of cross-section data and time-series, that is the observations which are made as many as \( r \) mutually independent subjects (cross-section) with each subject is repeatedly observed in \( n \) period of time (time-series) and between observations within the same subjects which are correlated [4]. In longitudinal data \((x_{it}, y_{it})\), the relationship between the predictor variable \( x_{it} \) with response variables \( y_{it} \) follows the regression model can be presented as follows:

\[
y_{it} = f_{i}(x_{it}) + \epsilon_{it}, i = 1,2,\ldots, r; t = 1,2,\ldots, n.
\]

For longitudinal data with predictor variable \( x \) as the observation time (design time points) [11], and \( f \) is the regression curve relationship between the predictor variables with the response variable \( y \) for to-\( i \) subject. The curve \( f \) can be approached in three ways: parametric, nonparametric, or semi-parametric. The parametric regression approach is used when it is assumed that the shape of the curve \( f \) is known, while the nonparametric regression approach is used when the shape of the curve \( f \) is unknown. On the other hand, semi-parametric regression approach is used when it is assumed that the shape of the curve is partially known, and some others are unknown [5].

Reproducing Kernel Hilbert Space (RKHS) play a central role to solve the Penalized Weighted Least Square (PWLS) in Spline Estimator of nonparametric regression analysis [6]. Based on the above background, the purposes of this research is to obtain the RKHS approach in PWLS to solve the estimator of regression curve on longitudinal data.

2 Materials and Methods

Nonparametric regression model for longitudinal data which involves \( r \) subject on \( n \) observation in each subject in equation (1). \( \epsilon_{it} \) variable is an error random variable assumed to be normally distributed to N-variat \((N = rn)\), with zero mean and variance-covariance matrix \( W^{-1} \) is follows:
Reproducing Kernel Hilbert space

The spline approach generally defines \( f_{ki} \) in equation (1) in form of an unknown regression curve, but \( f_{ki} \) is only assumed as smooth, in a sense of being contained in a specified function space, especially Sobolev space or written as

\[
W_{2^m}[a_{ki}, b_{ki}] = \left\{ f_{ki} : \int \left[ f_{ki}^{(m)}(x) \right]^2 dx < \infty \right\},
\]

for a positive integer \( m \). Optimization Penalized Weighted Least Square (PWLS) involves weighting in form of random error variance-covariance matrix as has been described in equation (1). To obtain the estimate of the regression curve \( f_{ki} \) using optimization PWLS that is the completion of optimization as follows [5]:

\[
\text{Min}_{\mathbf{f}_{ki} \in W^m_{2^m}[a_{ki}, b_{ki}]} \left\{ N^{-1}(\mathbf{y} - \mathbf{f})' \mathbf{W}(\mathbf{y} - \mathbf{f}) + \sum_{k=1}^r \sum_{l=1}^r \lambda_{kl} \int_{a_{ki}}^{b_{ki}} \left( f_{ki}^{(m)}(x_{ki}) \right)^2 dx_{ki} \right\}. \tag{4}
\]

The PWLS optimization in equation (4) in addition to considering the weight, also considers the use of 2r smoothing parameter \( \lambda_{ki} \) as a controller between the goodness of fit (the first segment) and the roughness penalty (second segment).

Suppose we want to find a function \( f_{ij} \) that interpolates between the points \( (x_{ij}, \eta_{ij}), \ell = 1, 2, \ldots, p; i = 1, 2, \ldots, r; t = 1, 2, \ldots, n \) and the boundary value of \( a = x_{i1} < x_{i2} < \ldots < x_{im} = b \). The functions \( f_{ij} \in W^2_m \) where

\[
W^2_m = \{ f : f \text{ is absolutely continuous on } [a, b], f' \in L^2[0,1] \},
\]

and we called in Sobolev Space [6].

If we restrict the boundary value in \( 0 = x_{i1} < x_{i2} < \ldots < x_{im} = 1 \) with transformation, then the value of \( f(0) = 0 \) is not really necessary, but simplifies the presentation of derivative value of \( f' \) or \( f'' \). Defining an inner product of \( W^2_m \) by:
\[
\langle f, g \rangle = \int_{0}^{1} f'(x)g'(x)dx,
\]
(6)

Implies a norm over the space \( W_m^2 \) that is small for “smooth” functions [6]. To
tackle the interpolation problem, is given:

\[
\langle f_i, R_{i\ell} \rangle = \int_{0}^{1} f'(x)R'_{i\ell}(x)dx = f_i(x_{i\ell}) - f_i(0) = f_i(x_{i\ell}).
\]
(7)

Thus interpolator \( f_i \) satisfies a system of equation (8), namely:

\[
f_i(x_{i\ell}) = \langle R_{i\ell}, f_i \rangle = \eta_{i\ell}, \ell = 1, 2, ..., p; i = 1, 2, ..., r; t = 1, 2, ..., n,
\]
(8)

and the smoothest function \( f \) satisfies an equation (9)

\[
f_i(x_{i\ell}) = \sum_{\ell=1}^{p} \sum_{i=1}^{r} \sum_{t=1}^{n} \xi_{i\ell t} R_{i\ell t}(x)
\]
(9)

The \( \xi_{i\ell t} \)'s are the solutions to the system of real linear equations obtained by
substituting of \( \hat{f}_i \) satisfies into (7),

\[
\sum_{\ell=1}^{p} \sum_{i=1}^{r} \sum_{t=1}^{n} \langle R_{i\ell t}, R_{i\ell t} \rangle \xi_{i\ell t} = \eta_{i\ell t}
\]
(10)

Note that

\[
\langle R_{i\ell t}, R_{i\ell t} \rangle = R_i(x_{i\ell}) = R_i(x_{i\ell}) = \min(x_{i\ell}, x_{i\ell})
\]
(11)

and define the function \( R_{x_i}(x_i) = \min(x_i, x_i) \) which turns out to be a reproduc-
ing kernel.

3 Result and Discussion

Assume that the data follow the nonparametric regression models for
longitudinal data:

\[
y_i = L_{x_i} f + \varepsilon_i, \quad (y_{11r}, y_{12r}, ..., y_{1tr}, y_{21r}, y_{22r}, ..., y_{2tr})' \quad \text{as the response variable},
\]
(12)

with \( y_i = (y_{11r}, y_{12r}, ..., y_{1tr}, y_{21r}, y_{22r}, ..., y_{2tr})' \) as the response variable, \( L_{x_i} \) limited
Reproducing Kernel Hilbert space

In order to obtain the function (12), RKHS is presented in the following. Hilbert spaces that display certain properties on certain linear operators are RKHS. The function \( f \) is the unknown function and is assumed smooth in a sense of contained in space \( H \).

Suppose the basis for the space \( H_0 \) is \( \{ \phi_{i_1}, \phi_{i_2}, \ldots, \phi_{i_m} \} \) and the basis for the space \( H_1 \) is \( \{ \xi_{i_1}, \xi_{i_2}, \ldots, \xi_{i_n} \} \), then for each function \( f_i \in H \) can be presented individually as:

\[
f_i = g_i + h_i, \quad (13)
\]

Furthermore, for every function \( f_i \in H \) can be presented individually as :

\[
f_i = g_i + h_i = \sum_{j=1}^{m} d_{ij} \phi_j + \sum_{t=1}^{n} c_{it} \xi_t = \phi_i^T d + \xi_i^T c\quad (14)
\]

\( L_x \) is a limited linear function in the space \( H \) and function \( f_i \in H \), obtained

\[
L_{x_i} f_i = L_{x_i} (g_i + h_i) = f_i(x_i) \quad (15)
\]

Based on the Riesz Representation Theorem and \( L_{x_i} \) is a limited linear function in the space \( H \), obtained a single value \( \eta_i \in H \) which is the representative of \( L_{x_i} \), and completes the equation :

\[
L_{x_i} f_i = \langle \eta_i, f_i \rangle = f_i(x_i), \quad f_i \in H \quad (16)
\]

Based on the equation (15), then \( f_i(x_i) \) in the equation (16) can be expressed as :

\[
f_i(x_i) = \langle \eta_i, \phi_i^T d \rangle + \langle \eta_i, \xi_i^T c \rangle \quad (17)
\]

The description of the equation (14) for \( k = 1,i = 1 \), obtained :

\[
f_i(x_i) = \langle \eta_i, \phi_i^T d \rangle + \langle \eta_i, \xi_i^T c \rangle, \quad t = 1,2,\ldots,n.
\]
\[
\begin{bmatrix}
    f_1(x_{11}) \\
    f_1(x_{12}) \\
    \vdots \\
    f_1(x_{1n}) \\
\end{bmatrix} = \begin{bmatrix}
    d_{11} \langle \eta_{11}, \phi_{11} \rangle + \ldots + d_{1n} \langle \eta_{11}, \phi_{1n} \rangle + c_{11} \langle \eta_{11}, \xi_{11} \rangle + \ldots + c_{1n} \langle \eta_{11}, \xi_{1n} \rangle \\
    d_{21} \langle \eta_{12}, \phi_{11} \rangle + \ldots + d_{2n} \langle \eta_{12}, \phi_{1n} \rangle + c_{21} \langle \eta_{12}, \xi_{11} \rangle + \ldots + c_{2n} \langle \eta_{12}, \xi_{1n} \rangle \\
    \vdots \\
    d_{n1} \langle \eta_{1n}, \phi_{11} \rangle + \ldots + d_{nn} \langle \eta_{1n}, \phi_{1n} \rangle + c_{n1} \langle \eta_{1n}, \xi_{11} \rangle + \ldots + c_{nn} \langle \eta_{1n}, \xi_{1n} \rangle \\
\end{bmatrix} \]

\[
= \begin{bmatrix}
    \langle \eta_{11}, \phi_{11} \rangle & \ldots & \langle \eta_{11}, \phi_{1n} \rangle \\
    \langle \eta_{12}, \phi_{11} \rangle & \ldots & \langle \eta_{12}, \phi_{1n} \rangle \\
    \vdots & \vdots & \vdots \\
    \langle \eta_{1n}, \phi_{11} \rangle & \ldots & \langle \eta_{1n}, \phi_{1n} \rangle \\
\end{bmatrix}
\begin{bmatrix}
    d_{11} \\
    d_{12} \\
    \vdots \\
    d_{1n} \\
\end{bmatrix} + \begin{bmatrix}
    \langle \eta_{11}, \xi_{11} \rangle & \ldots & \langle \eta_{11}, \xi_{1n} \rangle \\
    \langle \eta_{12}, \xi_{11} \rangle & \ldots & \langle \eta_{12}, \xi_{1n} \rangle \\
    \vdots & \vdots & \vdots \\
    \langle \eta_{1n}, \xi_{11} \rangle & \ldots & \langle \eta_{1n}, \xi_{1n} \rangle \\
\end{bmatrix}
\begin{bmatrix}
    c_{11} \\
    c_{12} \\
    \vdots \\
    c_{1n} \\
\end{bmatrix}
\]

\[
f_i(x_i) = T_i d_i + V_i c_i \tag{18}
\]

Then from the same way, it was obtained the result for \(i = 2, 3, \ldots, r\)

\[
f_i(x_i) = T_i d_i + V_i c_i \tag{19}
\]

\(T_i\) is the matrix of order \(n \times m\), \(d_i\) is the vector of order \(m\), \(V_i\) is the matrix of order \(n \times n\), \(c_i\) is the vector of order \(n\). Thus, the form of penalized regression multi-predictors model \(f(x)\) is as follows:

\[
f(x) = T d + V c. \tag{20}
\]

Estimation of Weighted \(W = (W^{-1})^{-1}\) coming from the inverse of variance-covariance error \(W^{-1}\). The estimation of \(W^{-1}\)

\[
\hat{W}^{-1} = \begin{bmatrix}
    \hat{\Sigma}_{11,11} & 0 & \cdots & 0 & \hat{\Sigma}_{12,11} & \hat{\Sigma}_{12,12} & \cdots & \hat{\Sigma}_{12,1r} \\
    0 & \hat{\Sigma}_{11,12} & \cdots & 0 & \hat{\Sigma}_{12,11} & \hat{\Sigma}_{12,12} & \cdots & \hat{\Sigma}_{12,1r} \\
    \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
    0 & 0 & \cdots & \hat{\Sigma}_{11,r} & \hat{\Sigma}_{12,r} & \hat{\Sigma}_{12,r+1} & \cdots & \hat{\Sigma}_{12,2r} \\
    \hat{\Sigma}_{12,11} & \hat{\Sigma}_{12,21} & \cdots & \hat{\Sigma}_{12,r} & 0 & \cdots & 0 \\
    \hat{\Sigma}_{12,12} & \hat{\Sigma}_{12,22} & \cdots & \hat{\Sigma}_{12,r+1} & 0 & \cdots & 0 \\
    \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
    \hat{\Sigma}_{12,r} & \hat{\Sigma}_{12,r+1} & \cdots & \hat{\Sigma}_{12,2r} & 0 & \cdots & 0 \\
\end{bmatrix}
\]

using Maximum Likelihood Estimation (MLE) approach, with likelihood function:
Reproducing Kernel Hilbert space

\[ L(f, W|y) = \prod_{i=1}^{n} \frac{1}{(2\pi)^{n/2}|W^{-1}|^{n/2}} \exp \left( -\frac{1}{2} (y_i - f_i)^T \Sigma_{ii} (y_i - f_i) \right) \]

\[ = \frac{1}{(2\pi)^{n/2}|W^{-1}|^{n/2}} \exp \left( -\frac{1}{2} \sum_{i=1}^{n} \left( (y_i - f_i)^T \Sigma_{ii} (y_i - f_i) + \sum_{j=1}^{n} (y_{ij} - f_{ij})^T \Sigma_{jj} (y_{ij} - f_{ij}) \right) \right) \]

Then the likelihood:

\[ L(f, W|y) = \prod_{i=1}^{n} \frac{1}{(2\pi)^{n/2}|W^{-1}|^{n/2}} \exp \left( -\frac{1}{2} \sum_{i=1}^{n} \left( (y_{i1} - f_{i1})^T \Sigma_{11} (y_{i1} - f_{i1}) + \sum_{j=2}^{n} (y_{ij} - f_{ij})^T \Sigma_{jj} (y_{ij} - f_{ij}) \right) \right) \]

Estimation the variance-covariance matrix \( \hat{W} \) from the optimization of \( \hat{W} \) from the optimization of \( \hat{W} \). Solving the variance-covariance \( \hat{W} \) is equals to solving the \( \hat{\Sigma}_{11,1}, \hat{\Sigma}_{12,1}, ..., \hat{\Sigma}_{22,r}, \hat{\Sigma}_{12,11}, ..., \hat{\Sigma}_{12,rr} \) as follow:
\[
\hat{\Sigma}_{11,11} = \frac{\hat{e}_{11}^2}{n} = \frac{(y_{11} - \hat{f}_{11,k_1})(y_{11} - \hat{f}_{11,k_1})'}{n},
\]
\[
\vdots \quad \vdots 
\]
\[
\hat{\Sigma}_{22,rr} = \frac{\hat{e}_{22}^2}{n} = \frac{(y_{2r} - \hat{f}_{2r,k_2})(y_{2r} - \hat{f}_{2r,k_2})'}{n},
\]
\[
\hat{\Sigma}_{12,11} = \frac{\hat{e}_{11}^2}{n} = \frac{(y_{11} - \hat{f}_{11,k_2})(y_{21} - \hat{f}_{21,k_2})'}{n},
\]
\[
\vdots \quad \vdots 
\]
\[
\hat{\Sigma}_{12,rr} = \frac{\hat{e}_{22}^2}{n} = \frac{(y_{1r} - \hat{f}_{1r,k_2})(y_{2r} - \hat{f}_{2r,k_2})'}{n}.
\]

For the purposes of \( \hat{f} \) estimation in (20), RKHS approach with completes the Penalized Weighted Least Square (PWLS) criterion as follow:

\[
\text{Min} \left\{ \left\| \hat{W} \xi \right\|^2 \right\} = \text{Min} \left\{ \left\| \hat{W}(y - f) \right\|^2 \right\}, \tag{21}
\]

With constraints : \( \left\| f_i \right\|^2 < \gamma_i, \quad \gamma_i \geq 0 \)

Then function \( H = W^2[a_{k_i}, b_{k_i}] \) used was Sobolev order-2 space which is defined as follows:

\[
W^2[a_{k_i}, b_{k_i}] = \left\{ f : \int_{a_{k_i}}^{b_{k_i}} (f^{(m)}(x_i))^2 \, dx_i < \omega \right\},
\]

with \( a_{k_i} \leq x_i \leq b_{k_i} \). Based on that space, norm for each function \( f_{k_i} \in W^2[a_{k_i}, b_{k_i}] \), defined is \( \left\| f_{k_i} \right\|^2 = \int_{a_{k_i}}^{b_{k_i}} (f_{k_i}^{(m)}(x_i))^2 \, dx_i \).

Optimization with constrain on the equation (21) can be written as :

\[
\text{Min} \left\{ \left\| \hat{W} \xi \right\|^2 \right\} = \text{Min} \left\{ \left\| \hat{W}(y - f) \right\|^2 \right\}, \tag{22}
\]

With constrain \( \int_{a_{k_i}}^{b_{k_i}} (f_{k_i}^{(m)}(x_i))^2 \, dx_i < \gamma_{k_i}, \quad \gamma_{k_i} \geq 0 \).

Weighed optimization with constrain (22) is equivalent to completing PWLS optimization with the equation (6). To complete the optimization, first penalty component must be described:
Reproducing Kernel Hilbert space

\[ \int_{a_1}^{b_1} \left[ f^{(m)}_{i1}(x_i) \right]^2 dx_i = \|P_i f_{i1}\|^2 = (P_i f_{i1}, P_i f_{i1}) \]

\[ = \left( \xi'_i \xi_{i1} \right) \left( \xi'_i \xi_{i1} \right)' = \xi'_i (\xi'_i \xi'_i) \xi_{i1} = \xi'_i V_{i1} \xi_{i1} \cdot \]

As a result

\[ \lambda_{i1} \int_{a_1}^{b_1} \left[ f^{(m)}_{i1}(x_i) \right]^2 dx_i = \lambda_{i1} \xi_{i1} V_{i1} \xi_{i1}. \] (23)

Using the same way, it was obtained:

\[ \lambda_{i2} \int_{a_2}^{b_2} \left[ f^{(m)}_i(x_i) \right]^2 dx_i = \lambda_{i2} \xi'_i V_{i2} \xi_{i2}. \] (24)

Based on the equation (21), penalty value gained:

\[ \sum_{k=1}^{2} \sum_{i=1}^{r} \lambda_{ik} \int_{a_{ik}}^{b_{ik}} \left[ f^{(m)}_{ki}(x_i) \right]^2 dx_i = \xi' \lambda V \xi, \] (25)

with \( \lambda = d_i a \lambda_{i1}, \lambda_{i2}, \lambda_{i3}, \lambda_{i4} \)

Based on the equation (24), the Goodness of fit on the PWLS optimization (24) can be written as:

\[ N^{-1} \left( y - f \right) \hat{W} \left( y - f \right) = N^{-1} \left( y - T_d - V \xi \right) \hat{W} \left( y - T_d - V \xi \right). \] (26)

By combining the goodness of fit (26) and the penalty (25), the PWLS optimization can be presented in the form of:

\[ \min_{c \in \mathbb{R}^{2m}, d \in \mathbb{R}^{2m}} \left\{ N^{-1} \left( y - T_d - V \xi \right) \hat{W} \left( y - T_d - V \xi \right) + \xi' \lambda V \xi \right\} \]

\[ = \min_{c \in \mathbb{R}^{2m}, d \in \mathbb{R}^{2m}} \left\{ (y' \hat{W} y - 2d' \hat{W} y - 2c' \hat{V} \hat{W} y + d' \hat{W} T_d + d' \hat{W} V \xi) + \right\} \]

\[ \xi' \hat{V} T d \xi' \left( \hat{V} \hat{W} V + N \lambda V \right) \}

\[ = \min_{c \in \mathbb{R}^{2m}, d \in \mathbb{R}^{2m}} \left\{ Q(c, d) \right\}. \] (22)

The completion of the optimization (22), was obtained by partially deriving Q(\( \xi, d \)) against \( \xi \) then the result was equated to zero, and gave the result:
\[-2\hat{V}\hat{W}_y + 2\hat{V}\hat{W}d + 2\left(\hat{V}\hat{W}\hat{V} + N\lambda V\right)\hat{\zeta} = 0\]
\[-\hat{W}_y + \hat{W}d + [\hat{W}\hat{V} + N\lambda I]\hat{\zeta} = 0.\] (23)

Suppose given matrix \( \mathbf{M} = \hat{W}\hat{V} + N\lambda I \)

Then the equation (23) can be written as:
\[-\hat{W}_y + \hat{W}d + \mathbf{M}\hat{\zeta} = 0\]
\[\mathbf{M}\hat{\zeta} = \hat{W}(y-Td).\] (24)

The equation (24) was multiply with \( \mathbf{M}^{-1} \) it was obtained the equation:
\[\hat{\zeta} = \mathbf{M}^{-1}\hat{W}(y-Td)\] (25)

Then partial derivatif against \( d \) then the result was equated to zero, it gave the result of:
\[-T\hat{W}_y + T\hat{W}d + T[\hat{W}\mathbf{V}\mathbf{M}^{-1}]\hat{W}(y-Td)=0.\] (26)

Because \( \mathbf{M} = \hat{W}\hat{V} + N\lambda I \) so \( \mathbf{V} = [\hat{W}]^{-1}(\mathbf{M} - N\lambda I) \).

As a result, it was obtained the equation:
\[\mathbf{V}\mathbf{M}^{-1} = \hat{W}^{-1}(\mathbf{M} - N\lambda I) \mathbf{M}^{-1} = \hat{W}^{-1}(\mathbf{I} - N\lambda \mathbf{M}^{-1})\]

By doubling the equation above with \( \mathbf{W} \) it was obtained:
\[\hat{W}\mathbf{V}\mathbf{M}^{-1} = \mathbf{I} - N\lambda \mathbf{M}^{-1}\]

This equation was substituted in the equation (26) it was obtained the equation:
\[-T^\top\hat{W}_y + T^\top\hat{W}d + (T^\top[I - N\lambda \mathbf{M}^{-1}]\hat{W}(y-Td)=0.\]

If the equation above was described it was gained the equation:
\[\hat{d} = (T^\top\mathbf{M}^{-1}\hat{W}^\top)^{-1} T^\top\mathbf{M}^{-1}\hat{W}.\] (27)
Reproducing Kernel Hilbert space

The equation (26) was substituted into the equation (24) it was gained:

\[ \hat{\xi} = M^{-1}\hat{\mathbf{W}}[I - T(T'M^{-1}\hat{\mathbf{W}})^{-1}T'M^{-1}\hat{\mathbf{W}}]y. \]  

(28)

Based on the Equation (27) and Equation (28), it was obtained the estimator for nonparametric regression curve for longitudinal data involving a single predictor variable as follows:

\[ \hat{f}_x = T\hat{\xi} + V\hat{\xi} \]

\[ = \{ T(T'M^{-1}\hat{\mathbf{W}})^{-1}T'M^{-1}\hat{\mathbf{W}} + VM^{-1}\hat{\mathbf{W}}[I - T(T'M^{-1}\hat{\mathbf{W}})^{-1}T'M^{-1}\hat{\mathbf{W}}] \}y. \]

(29)

with:

\[ A(\hat{\xi}) = T(T'M^{-1}\hat{\mathbf{W}})^{-1}T'M^{-1}\hat{\mathbf{W}} + VM^{-1}\hat{\mathbf{W}}[I - T(T'M^{-1}\hat{\mathbf{W}})^{-1}T'M^{-1}\hat{\mathbf{W}}]. \]

4 Conclusion

Based on the results of the study, several things can be concluded that Base of RKHS, the curve nonparametric regression form is \( f(x) = T\hat{\xi} + V\hat{\xi} \), with

\[
\begin{pmatrix}
T_1 & 0 & \cdots & 0 \\
0 & T_2 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & T_r
\end{pmatrix}
\text{ and }
\begin{pmatrix}
V_1 & 0 & \cdots & 0 \\
0 & V_2 & \cdots & 0 \\
0 & 0 & \cdots & V_r
\end{pmatrix}
\]

(18)

Solving the weighted of PLWS coming from variance-covariance \( \hat{\mathbf{W}} \) is equals to solving the \( \hat{\Sigma}_{11,11}, \hat{\Sigma}_{11,2}, \ldots, \hat{\Sigma}_{22,r}, \hat{\Sigma}_{12,11}, \ldots, \hat{\Sigma}_{12,rr} \) as follow:

\[
\hat{\Sigma}_{11,11} = \frac{\hat{\xi}_{11,11}^\prime}{n} = \frac{(y_{11} - \hat{\xi}_{11,1})(y_{11} - \hat{\xi}_{11,1})^\prime}{n},
\]

\[
\vdots \\
\hat{\Sigma}_{22,rr} = \frac{\hat{\xi}_{22,rr}^\prime}{n} = \frac{(y_{22,rr} - \hat{\xi}_{22,rr})(y_{22,rr} - \hat{\xi}_{22,rr})^\prime}{n},
\]

\[
\hat{\Sigma}_{12,11} = \frac{\hat{\xi}_{12,11}^\prime}{n} = \frac{(y_{11} - \hat{\xi}_{12})(y_{21} - \hat{\xi}_{21})^\prime}{n},
\]

\[
\vdots \\
\hat{\Sigma}_{12,rr} = \frac{\hat{\xi}_{12,rr}^\prime}{n} = \frac{(y_{1r} - \hat{\xi}_{1r})(y_{2r} - \hat{\xi}_{2r})^\prime}{n}.
\]
For the purposes of \( \hat{f} \) estimation, RKHS approach with completes the Penalized Weighted Least Square (PWLS) criterion is 

\[
\hat{f}_\alpha = A^* (\hat{\lambda}) y, \text{ with } \\
A (\hat{\lambda}) = T (T' M^{-1} W T)^{-1} T' M^{-1} W + V M^{-1} \hat{W} [1 - T (T' M^{-1} W T)^{-1} T' M^{-1} \hat{W}]
\]

References


Received: August 15, 2014