Variable selection for functional linear models with functional predictors and a functional response

Matsui, Hidetoshi
Department of Mathematical Sciences, Graduate School of Engineering, Osaka Prefecture University

出版情報：MI Preprint Series. 2013-14, 2013-11-22. Faculty of Mathematics, Kyushu University
バージョン：
権利関係：

MI Preprint Series

Mathematics for Industry
Kyushu University

Variable selection for functional linear models with functional predictors and a functional response

Hidetoshi Matsui

MI 2013-14

(Received November 22, 2013)

Institute of Mathematics for Industry
Graduate School of Mathematics
Kyushu University
Fukuoka, JAPAN

Variable selection for functional linear models with functional predictors and a functional response

Hidetoshi Matsui

Faculty of Mathematics, Kyushu University
744 Motooka, Nishi-ku, Fukuoka 819-0395, Japan.

hmatsui@math.kyushu-u.ac.jp

Abstract: We consider a variable selection problem for functional linear models where both multiple predictors and a response are functions. Especially we assume that variables are given as functions of time and then construct the historical functional linear model which takes the relationship of dependences of predictors and a response into consideration. Unknown parameters included in the model are estimated by the maximum penalized likelihood method with the L_1 penalty. We can simultaneously estimate and select variables given as functions using the L_1 type penalty. A regularization parameter involved in the regularization method is decided by a model selection criterion. The effectiveness of the proposed method is investigated by simulation studies and real data analysis.

Key Words and Phrases: Functional Data Analysis, Longitudinal Data, Model selection, Sparse regularization

1 Introduction

Functional data analysis (FDA) has received considerable attentions in several fields such as meteorology, ergonomics and medicine, and there still are so many studies in both theoretical and applicative aspects (see, e.g. Ramsay and Silverman, 2005; Horváth and Kokoszka, 2012). The basic concept behind FDA is to represent repeated measurement data for individual as smooth functions and then treat them as if they themselves were the observed data. In this paper we consider a variable selection problem for functional data in the regression model.

There have been many works for functional regression models where predictors are functions while a response(s) is a scalar. The functional linear model was extended to several frameworks such as the generalized linear model (James, 2002; Müller and Stadtmüller, 2005), the additive model (Müller and Yao, 2008) and the adaptive model (James and Silverman, 2005). Furthermore, theoretical evaluations for the functional linear model were developed by Cai and Hall (2006); Hall and Horowitz (2007); Cai and Yuan (2012).

On the other hand, when both predictors and a response are given as functions there are two conceivable cases. One is that arguments of both the predictors and the response are the same (denoted by $X(t)$ and $Y(t)$ respectively) and the other vice versa ($X(s)$ and $Y(t)$). For the former case the varying coefficient model (Hastie and Tibshirani, 1993; Hoover et al., 1998) can be applied for the modeling of the relationship. On the other hand, the latter can consider the case where certain interval of the domain of the functional predictor affect the response and it is more natural than the former case. Therefore we consider the latter case throughout this paper. Ramsay and Dalzell (1991) first constructed the functional linear model for a functional predictor and a response as follows. Let $x_i(s)$ and $y_i(t)$ be respectively a predictor and a response given as functions with $s \in [0, S]$ and $t \in [0, T]$ for i-th subject. Then the functional linear model for $x_i(s)$ and $y_i(t)$ is

$$y_i(t) = \alpha(t) + \int_0^T x_i(s)\beta(s,t)ds + \varepsilon_i(t), \tag{1}$$

where $\alpha(t)$ is an intercept function, $\beta(s,t)$ is a coefficient function and $\varepsilon_i(t)$ are an error functions. Matsui et al. (2009) proposed estimating the model by the maximum penalized likelihood method and also derived a model selection criterion for evaluating the estimated model, and Yao et al. (2005) obtained a consistent estimates for the model as the theoretical development.

When s and t in the model (1) represent times, the response depends on future information of the predictor, which leads to paradoxical and inappropriate results, except that they have a periodicity. In order to solve this problem Malfait and Ramsay (2003) took the relationship of dependences of $x(s)$ and $y(t)$ into consideration and proposed a historical functional linear model (HFLM) as a special case of (1), and they also investigated how to estimate it. Furthermore, Harezlak et al. (2007) estimated the HFLM by the penalized least squares method with the L_2 or the L_1 penalty. Şentürk and Müller (2008, 2010) also discussed the similar situations for the frameworks of varying-coefficient models.

While these studies treat functional linear model with one predictor, in other word, the functional simple regression models, we consider the variable selection problem for multiple functional predictors in the multiple functional regression model, with the help of the sparse regularization. Sparse regularization is one of the most useful tools for variable selection problems and has come to be used in various situations. It can simultaneously estimate parameters and select variables by imposing L_1 type penalties. There have been proposed several L_1 type penalties (Tibshirani, 1996; Fan and Li, 2001; Zou and Hastie, 2005; Zhang, 2010). Matsui and Konishi (2011) proposed selecting functional predictors using the sparse regularization in the functional linear model with a scalar response.

We propose a method for the strategy for the problem of variable selection for the functional linear model with functional predictors and a functional response. Functional data and coefficients are represented by basis expansions. Since it is difficult to analytically evaluate functions in the model, an approximate calculation is introduced. Then parameters included in the model are estimated by the maximum penalized likelihood method via the sparse regularization. In order to choose the degrees of regularization we apply a model selection criterion derived for evaluating the functional linear model. Monte Carlo simulations are conducted to see the effectiveness of the proposed modeling strategy. Then we apply the proposed method to the analysis of typhoon data, trying to select functional variables which have effects on the path of typhoons.

This paper is organized as follows. Section 2 introduces a HFLM that models the relationship between multiple predictors and a response both of which are given by functions of time. Section 3 provides how to estimate and evaluate the model. Numerical examples are investigated in Section 4 and real data analysis is described in Section 5. Finally we summarize the main results in Section 6.

2 Functional linear model with functional predictors and a response

Suppose we have n sets of p functional predictors and a functional response $\{(x_{ij}(s), y_i(t)); s,t \in [0,T], i = 1,\dots,n, j = 1,\dots,p\}$. In order to model the relationship between predictors and a response, we consider the following historical functional linear model (HFLM, Malfait and Ramsay, 2003; Ramsay and Silverman, 2005):

$$y_i(t) = \alpha(t) + \sum_{j=1}^{p} \int_{s_j(t)}^{t} x_{ij}(s)\beta_j(s,t)ds + \varepsilon_i(t), \tag{2}$$

where $\alpha(t)$ is an intercept function, $\beta_j(s,t)$ are bivariate coefficient functions which impose varying weights on $x_{im}(s)$ at $s \in [s_j(t), t]$ rather than $s \in [0,T]$, $s_j(t) = \max\{0, t-\delta_j\}$ with a lag parameter $\delta_j > 0$ which decide how long the time is included in the model, and $\varepsilon_i(t)$ are error functions. There are several relationships between other models for longitudinal data analysis. If intervals of the integration with respect to s are shrunk to $s_j(t) = t$, that is, the arguments of the predictors and the response are the same, the HFLM corresponds to a varying-coefficient model of Hastie and Tibshirani (1993); Hoover et al. (1998):

$$y_i(t) = \alpha(t) + \sum_{j=1}^{p} x_{ij}(t)\beta_j(t) + \varepsilon_i(t).$$

On the other hand, if $[s_j(t), t]$ are discretized to be t_l, $l = 1,\dots,R_j$ so that $t_l = t$ and $t_{l-(R_j-1)} = s_j(t)$ for fixed t. then it correspond to a generalized varying-coefficient model by Şentürk and Müller (2008) with multiple predictors:

$$y_i(t_l) = \alpha(t_l) + \sum_{j=1}^{p}\sum_{r=1}^{R_j} x_{ij}(t_{l-(r-1)})\beta_{jr}(t_l) + \varepsilon_i(t_l).$$

Now we return to the HFLM (2). From the normal equation the intercept function is given by

$$\alpha(t) = \bar{y}(t) - \sum_{j=1}^{p} \int_{s_j(t)}^{t} \bar{x}_m(s)\beta_j(s,t)ds + \bar{\varepsilon}(t)$$

with $\bar{x}(s) = \sum_i x_i(s)/n$, $\bar{y}(t) = \sum_i y_i(t)/n$ and $\bar{\varepsilon}(t) = \sum_i \varepsilon_i(t)/n$. Therefore the HFLM (2) becomes

$$y_i^c(t) = \sum_{j=1}^{p} \int_{s_0(t)}^{t} x_{ij}^c(s)\beta_j(s,t)ds + \varepsilon_i^c(t), \tag{3}$$

where $x_i^c(s) = x_i(s) - \bar{x}(s)$, $y_i^c(t) = y_i(t) - \bar{y}(t)$ and $\varepsilon_i^c(t) = \varepsilon_i(t) - \bar{\varepsilon}(t)$. For simplicity we drop the suffix c for the rest of this paper.

Suppose that the coefficient functions $\beta_j(s,t)$ are approximated by basis expansions:

$$\tilde{\beta}_j(s,t) = \sum_{k=1}^{K_j} b_{jk}\phi_{jk}(s,t), \tag{4}$$

where b_{jk} are unknown parameters and $\beta_j(s,t)$ are basis functions. Typical bases include radial basis functions or thin-plate splines. Defining the residual $\tilde{\varepsilon}_{(j)}(s,t) = \beta_j(s,t) - \tilde{\beta}_j(s,t)$, the HFLM (2) becomes

$$\begin{aligned} y_i(t) &= \sum_{j=1}^{p} \int_{s_j(t)}^{t} x_{im}(s) \left\{ \sum_{k=1}^{K_j} b_{jk}\phi_{jk}(s,t) + \tilde{\varepsilon}_{(j)}(s,t) \right\} ds + \varepsilon_i(t) \\ &= \sum_{j=1}^{p} \sum_{k=1}^{K_j} b_{jk} \int_{s_j(t)}^{t} x_{ij}(s)\phi_{jk}(s,t)ds + \varepsilon_i(t) \\ &= \sum_{j=1}^{p} \sum_{k=1}^{K_j} b_{jk}\psi_{ijk}(t) + \varepsilon_i(t), \end{aligned}$$

where

$$\psi_{ijk}(t) = \int_{s_j(t)}^{t} x_{ij}(s)\phi_{jk}(s,t)ds.$$

Note that we rewrote $\sum_j \int x_{ij}(s)\tilde{\varepsilon}_{(j)}(t)ds + \varepsilon_i(t)$ as $\varepsilon_i(t)$. Using notations $\boldsymbol{y}(t) = (y_1(t), \ldots, y_n(t))'$, $\boldsymbol{\varepsilon}(t) = (\varepsilon_1(t), \ldots, \varepsilon_n(t))'$, $\Psi_j(t) = (\psi_{ijk}(t))_{ik}$, $\Psi(t) = (\Psi_1(t), \ldots, \Psi_p(t))$, $\boldsymbol{b} = (\boldsymbol{b}_1', \ldots, \boldsymbol{b}_p')'$, and $\boldsymbol{b}_j = (b_{j1}, \ldots, b_{jK_j})'$ the model (3) has the form of

$$\begin{aligned} \boldsymbol{y}(t) &= \sum_{j=1}^{p} \Psi_j(t)\boldsymbol{b}_j + \boldsymbol{\varepsilon}(t) \\ &= \Psi(t)\boldsymbol{b} + \boldsymbol{\varepsilon}(t), \end{aligned} \tag{5}$$

which is similar to the standard linear model with a design matrix $\Psi(t)$, a response vector $\boldsymbol{y}(t)$ and a coefficient vector $\boldsymbol{b}$, except that some of vectors are functions of t.

3 Estimation and evaluation

We consider estimating the functional linear model described in the previous section. As the standard and natural approach for this problem, Ramsay and Silverman (2005), Chapter 16 used the integrated sum of squared error as the criterion to be minimized as the standard and natural approach for this problem, that is, they considered minimizing

$$\text{LMSSE}(\boldsymbol{b}) = \int_0^T \sum_{i=1}^{n} \varepsilon_i^2(t)dt = \int_0^T \{\boldsymbol{y}(t) - \Psi(t)\boldsymbol{b}\}' \{\boldsymbol{y}(t) - \Psi(t)\boldsymbol{b}\} \, dt.$$

However, it is difficult to directly obtain estimates of $\boldsymbol{b}$ in (5) by two reasons. First, it is difficult to calculate ψ_{ijk} analytically because of the complexity of the integral. Second, the least squares method sometimes gives unstable or degenerate estimates. In order to solve these problems, we respectively apply the Finite Elements Method (FEM) that approximate the integration numerically and the regularization method, especially the sparse regularization to select functional variables.

3.1 Finite element method

In order to calculate ψ_{ijk} numerically we here apply the finite element method (FEM) to estimate the coefficient vector $\boldsymbol{b}$. Malfait and Ramsay (2003) used the FEM for the HFLM and described the details of it.

Consider a two dimensional coordinate of s, t which include the domain of integration in (2) (Figure 1 left). First we divide the intervals $[0, T]$ for s and t directions into N intervals, each of which have equal length of μ, and then construct triangular elements by further dividing each square grid into two triangles (Figure 1 right). The lag parameter δ_j can be approximated by $M_j\mu$ $(0 \leq M_j \leq N)$ for each j. When $M_j = N$ the shape of the domain becomes a triangle, $0 < M_j < N$ corresponds to a trapezoid, and when $M_j = 0$ the domain is only $s = t$ and it corresponds to the varying-coefficient model. As a result, there are $M_j(2N - M_j)$ triangular elements and $(M_j + 1)(N + 1 - M_j/2)$ nodes on the domain of $\beta_j(s, t)$. Each node is assigned by one basis function which has the shape of a hexagonal pyramid and has the value of 1 at the node and 0 at the adjacent nodes. Figure 2 shows a contour plot of an example of the basis function. These bases correspond to $\phi_{jk}(t)$ $(j = 1, \ldots, p,\ k = 1, \ldots, (M_j + 1)(N + 1 - M_j/2)(= K_j))$ of (4). We used Matlab functions which connect nodes with indices, available from the website of Ramsay and Silverman (2002).

We consequently discretize the time point t into finite time points Q. Malfait and Ramsay (2003) says that the number of $Q = 4N$ gives sufficiently accuracy for the approximation. Using this discretization, $y_i(t)$, $\psi_{ijk}(t)$ and $\varepsilon_i(t)$ are respectively expressed as vectors $\boldsymbol{y}_i = (y_{i1}, \ldots, y_{iQ})'$, $\boldsymbol{\psi}_{ijk} = (\psi_{i1jk}, \ldots, \psi_{iQjk})'$ and $\boldsymbol{\varepsilon}_i = (\varepsilon_{i1}, \ldots, \varepsilon_{iQ})'$. Then $\boldsymbol{y}(t)$, $\Psi(t)$ and $\boldsymbol{\varepsilon}(t)$ are respectively represented as vector or matrix forms

$$\boldsymbol{y} = \begin{pmatrix} \boldsymbol{y}_1 \\ \vdots \\ \boldsymbol{y}_n \end{pmatrix}, \Psi = \begin{pmatrix} \boldsymbol{\psi}_{111} & \cdots & \boldsymbol{\psi}_{11K_1} & \cdots & \boldsymbol{\psi}_{1p1} & \cdots & \boldsymbol{\psi}_{1pK_p} \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \boldsymbol{\psi}_{n11} & \cdots & \boldsymbol{\psi}_{n1K_1} & \cdots & \boldsymbol{\psi}_{np1} & \cdots & \boldsymbol{\psi}_{npK_p} \end{pmatrix} \text{ and } \boldsymbol{\varepsilon} = \begin{pmatrix} \boldsymbol{\varepsilon}_1 \\ \vdots \\ \boldsymbol{\varepsilon}_n \end{pmatrix},$$

and finally we have a form of linear model

$$\boldsymbol{y} = \Psi \boldsymbol{b} + \boldsymbol{\varepsilon}. \tag{6}$$

3.2 Penalized likelihood method via the sparse regularization

We here assume that error vectors $\boldsymbol{\varepsilon}_i$ are identically and independently normally distributed with mean vector $\mathbf{0}$ and variance covariance matrix Σ_0 and that Σ_0 has an autocorrelation since $\varepsilon_{i1}, \ldots, \varepsilon_{iQ}$ are discretized realization at continuous time points. That

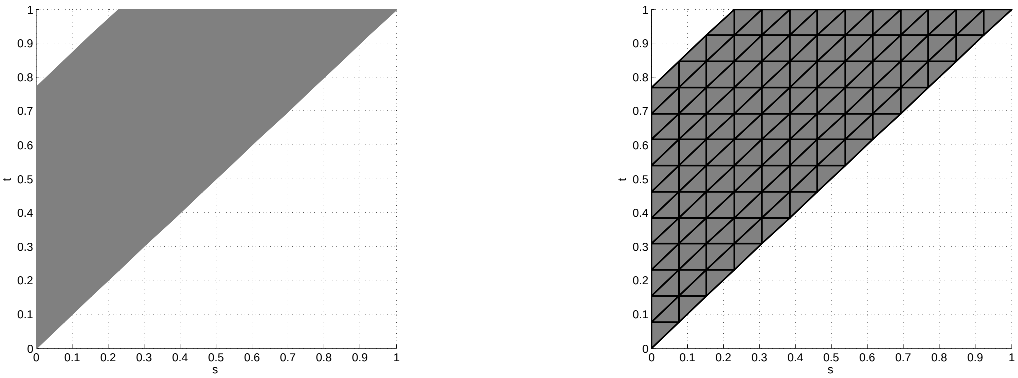

Figure 1: Illustration of the region of integral of the HFLM (left) and its triangulation (right).

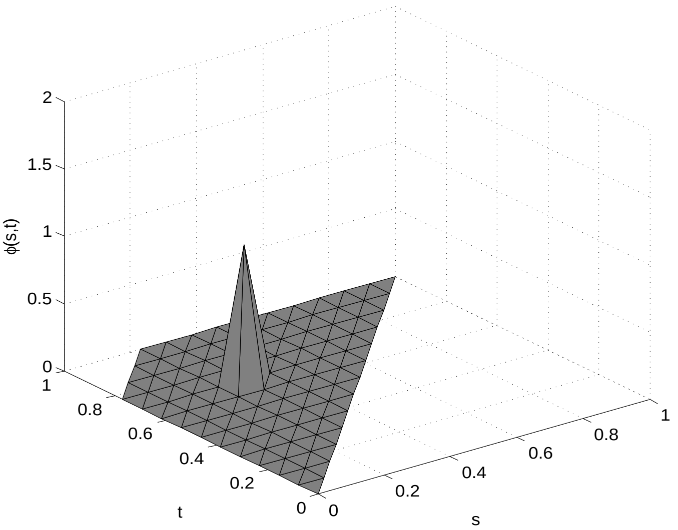

Figure 2: An example of basis functions.

is, we assume that Σ_0 has the form of

$$\Sigma_0 = \frac{\sigma^2}{1-\rho^2}\begin{pmatrix} 1 & \rho & \cdots & \rho^{Q-1} \\ \rho & 1 & \cdots & \rho^{Q-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho^{Q-1} & \rho^{Q-2} & \cdots & 1 \end{pmatrix}, \tag{7}$$

where σ^2 and $\rho \in [-1, 1]$ are variance and correlation parameters respectively. Here we referred to Fahrmeir et al. (2013) pp.192 about the autocorrelation. Then the variance covariance matrix Σ of $\boldsymbol{\varepsilon}$ is $\Sigma = I_n \otimes \Sigma_0$, and hence we have a probability density function

$$f(\boldsymbol{y}, \boldsymbol{\theta}) = \frac{1}{(2\pi)^{nQ/2} \log|\Sigma|} \exp\left\{-\frac{1}{2}(\boldsymbol{y} - \Psi\boldsymbol{b})'\Sigma^{-1}(\boldsymbol{y} - \Psi\boldsymbol{b})\right\}, \tag{8}$$

where $\boldsymbol{\theta} = (\boldsymbol{b}', \sigma^2, \rho)'$, and then the log-likelihood function is given by

$$\ell(\boldsymbol{\theta}) = -\frac{nQ}{2}\log(2\pi) - \frac{1}{2}\log|\Sigma|^{1/2} - \frac{1}{2}(\boldsymbol{y} - \Psi\boldsymbol{b})'\Sigma(\boldsymbol{y} - \Psi\boldsymbol{b}).$$

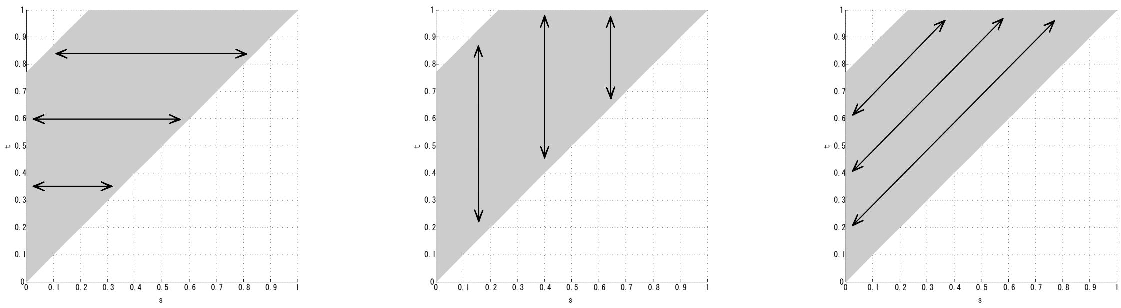

Figure 3: Directions for penalizations about $D_j^{(H)}$ (left), $D_j^{(V)}$ (center) and $D_j^{(P)}$ (right).

We estimate parameters $\boldsymbol{b}$, σ^2 and ρ by maximizing the following penalized likelihood function

$$\ell_\lambda(\boldsymbol{\theta}) = \ell(\boldsymbol{\theta}) - nQ \sum_{j=1}^{p} P_\lambda(\|\boldsymbol{b}_j\|_{\Omega_j}), \tag{9}$$

where $P_\lambda(\cdot)$ is a penalty function, $\|\boldsymbol{b}_j\|_{\Omega_j} = \sqrt{\boldsymbol{b}_j'\Omega_j\boldsymbol{b}_j}$ and Ω_j is a positive semi-definite matrix. Harezlak et al. (2007) has penalized for the fluctuation of the coefficient function in the direction parallel to s-axis (horizontal), t-axis (vertical) and $s = t$ (parallel), each of which corresponds to penalizing for neighboring basis functions for fixed t, s and $s - t$ (Figure 3). We use a first difference for the neighboring bases, constructing penalty matrices for horizontal, vertical and parallel direction for j-th variable, denoted by $D_j^{(H)}$, $D_j^{(V)}$ and $D_j^{(P)}$. For example, the elements of $D_j^{(H)}$ are given by

$$\left\{ \begin{array}{l} \left(D_j^{(H)}\right)_{lm} = 1 \\ \left(D_j^{(H)}\right)_{lm'} = -1 \\ \left(D_j^{(H)}\right)_{lm''} = 0 \end{array} \right. \quad \left(\text{if} \quad \begin{array}{c} s_m - s_{m'} = 1 \\ t_m - t_{m'} = 0 \\ m'' \neq m, m' \end{array} \right), \tag{10}$$

where s_m and t_m are respectively s and t coordinates of the m-th node, $m = 1, \ldots, K_j$, $l = 1, \ldots, L$ and L is the number of combinations that the condition in (10) occurs. $D_j^{(V)}$ and $D_j^{(P)}$ are given in similar ways. Then the matrices Ω_j are given by $\Omega_j = \gamma_j^{(H)}(D_j^{(H)})'D_j^{(H)} + \gamma_j^{(V)}(D_j^{(V)})'D_j^{(V)} + \gamma_j^{(P)}(D_j^{(P)})'D_j^{(P)}$ with tuning parameters $\gamma_j^{(H)}$, $\gamma_j^{(V)}$ and $\gamma_j^{(P)}$ that control degrees of regularizations for each direction. Although it can be considered that all of these tuning parameters are selected by model selection criteria such as AIC or BIC, the computational load can be very expensive. Alternatively we decide these values as the following rule, using the idea of Fan and Li (2004). First we obtain the maximum likelihood estimator of $\boldsymbol{b}$, denoted by $\hat{\boldsymbol{b}}^{(ML)}$, by maximizing (9) with $\lambda = 0$, and then $\gamma_j^{(H)}$ is given as the standard deviation of $D_j^{(H)}\hat{\boldsymbol{b}}_j^{(ML)}$. $\gamma_j^{(V)}$ and $\gamma_j^{(P)}$ are obtained in same way, then the overall degrees of the regularization is controlled by λ, which is selected by model selection criteria.

For the penalty $P_\lambda(\cdot)$ we apply a smoothly clipped absolute deviation (SCAD) penalty by Fan and Li (2001) whose derivative is defined by

$$P'_\lambda(|\theta|) = \lambda \left\{ I(|\theta| \le \lambda) + \frac{(a\lambda - |\theta|)_+}{(a-1)\lambda} I(|\theta| > \lambda) \right\}$$

for arbitrary θ, where λ is a regularization parameter and a is a tuning parameter. Here we set $a = 3.7$ since Fan and Li (2001) says that this value gives minimum Bayes rule.

Parameters are estimated via the local quadratic approximation which iteratively update parameters and has been applied for the lasso (Tibshirani, 1996) and the SCAD (Fan and Li, 2001). Denote an initial value of $\boldsymbol{b}$ in the update as $\boldsymbol{b}^{(0)}$, then the Taylor expansion of $P_\lambda(\|\boldsymbol{b}_j\|_{\Omega_j})$ around $\boldsymbol{b}_j^{(0)}$ gives

$$P_\lambda(\|\boldsymbol{b}_j\|_{\Omega_j}) \approx P_\lambda(\|\boldsymbol{b}_j^{(0)}\|_{\Omega_j}) + \frac{1}{2}\frac{P'_\lambda(\|\boldsymbol{b}_j^{(0)}\|_{\Omega_j})}{\|\boldsymbol{b}_j^{(0)}\|_{\Omega_j}} \left\{ \boldsymbol{b}_j'\boldsymbol{b}_j - \left(\boldsymbol{b}_j^{(0)}\right)' \boldsymbol{b}_j^{(0)} \right\}.$$

Using this approximation and assuming a fixed variance σ^2 the penalized log-likelihood function (9) can also be approximated by

$$\ell_\lambda(\boldsymbol{b}) \approx \ell(\boldsymbol{b}^{(0)}) + \frac{\partial \ell(\boldsymbol{b}^{(0)})}{\partial \boldsymbol{b}'}(\boldsymbol{b} - \boldsymbol{b}^{(0)}) + \frac{1}{2}(\boldsymbol{b} - \boldsymbol{b}^{(0)})' \frac{\partial \ell(\boldsymbol{b}^{(0)})}{\partial \boldsymbol{b} \partial \boldsymbol{b}'}(\boldsymbol{b} - \boldsymbol{b}^{(0)}) + \frac{nQ}{2}\boldsymbol{b}'\Omega(\boldsymbol{b})\boldsymbol{b},$$

where $\Omega(\boldsymbol{b}) = \mathrm{blockdiag}\{P'_\lambda(\|\boldsymbol{b}_1\|_{\Omega_1})/\|\boldsymbol{b}_1\|_{\Omega_1}, \ldots, P'_\lambda(\|\boldsymbol{b}_p\|_{\Omega_p})/\|\boldsymbol{b}_p\|_{\Omega_p}\}$. Then if k-th updated values of $\boldsymbol{b}$ and Σ, respectively denoted by $\boldsymbol{b}^{(k)}$ and $\Sigma^{(k)}$, are obtained, the $(k+1)$-th updated value of $\boldsymbol{b}$ is given by

$$\boldsymbol{b}^{(k+1)} = \left\{ \Psi' \left(\Sigma^{(k)}\right)^{-1} \Psi + nQ\Omega(\boldsymbol{b}^{(k)}) \right\}^{-1} \Psi' \left(\Sigma^{(k)}\right)^{-1} \boldsymbol{y},$$

and subsequently the correlation and variance parameter are respectively updated by

$$\rho^{(k+1)} = \frac{s_{q1}}{s_q}, \quad \left(\sigma^2\right)^{(k+1)} = \frac{1}{nQ}\left(\boldsymbol{y} - \Psi \boldsymbol{b}^{(k+1)}\right)' \left(\boldsymbol{y} - \Psi \hat{\boldsymbol{b}}^{(k+1)}\right),$$

where

$$s_{q1} = \frac{1}{nQ}\sum_{i=1}^{n}\sum_{q=2}^{Q}\left(y_{iq} - \sum_{j=1}^{p}\sum_{k=1}^{K_j}\psi_{iqjk}b_{jk}\right)\left(y_{i(q-1)} - \sum_{j=1}^{p}\sum_{k=1}^{K_j}\psi_{i(q-1)jk}b_{jk}\right),$$

$$s_q = \frac{1}{nQ}\sum_{i=1}^{n}\sum_{q=1}^{Q}\left(y_{iq} - \sum_{j=1}^{p}\sum_{k=1}^{K_j}\psi_{iqjk}b_{jk}\right)^2.$$

The updated variance covariance matrix $\Sigma^{(k+1)}$ is obtained by substituting $\rho^{(k+1)}$ and $(\sigma^2)^{(k+1)}$ into (7). This update is continued until the convergence criterion is satisfied, then we obtain the estimated parameters $\hat{\boldsymbol{b}}$, $\hat{\sigma}^2$ and $\hat{\rho}$. Substituting $\hat{\boldsymbol{\theta}} = (\hat{\boldsymbol{b}}', \hat{\sigma}^2, \hat{\rho})'$ into (8) we obtain a statistical model

$$f(\boldsymbol{y}, \hat{\boldsymbol{\theta}}) = \frac{1}{(2\pi)^{nQ/2}\log|\hat{\Sigma}|}\exp\left\{-\frac{1}{2}\left(\boldsymbol{y} - \Psi\hat{\boldsymbol{b}}\right)' \hat{\Sigma}^{-1}\left(\boldsymbol{y} - \Psi\hat{\boldsymbol{b}}\right)\right\}. \tag{11}$$

3.3 Model selection criterion

Since the estimated model (11) strongly depends on the value of tuning parameters such as regularization parameter λ. Although the cross-validation is widely used for the selection of tuning parameters, Leng et al. (2006) showed that the criteria based on the minimum prediction error such as cross-validation or generalized cross-validation do not select models consistently. On the other hand, Wang et al. (2007) and Zhang et al. (2010) showed that the BIC type criterion with the effective degrees of freedom select the true model consistently for the SCAD regularization. We use a BIC type model selection criterion for evaluating the historical functional linear model (2) estimated by the maximum penalized likelihood method with the group SCAD regularization. The BIC is given by

$$\mathrm{BIC} = -2\ell(\hat{\boldsymbol{\theta}}) + df \log(nQ),$$

where df is an effective degrees of freedom given by

$$df = \mathrm{tr}\left\{\Psi\left(\Psi'\hat{\Sigma}^{-1}\Psi + nQ\Omega(\hat{\boldsymbol{b}})\right)^{-1}\Psi'\hat{\Sigma}^{-1}\right\}.$$

We select the regularization parameter λ which minimizes the BIC and treat it as an optimal model.

4 Simulation

We conducted Monte Carlo simulations to show the effectiveness of the proposed method. We simulated functional predictors $X_j(s)$ $(j = 1, \ldots, 5)$ and a response $Y(t)$, where $j = 1, \ldots, 5$ and $s, t \in [0, 1]$, as the following rule on the model of the HFLM. First, we constructed functional predictors and coefficients by

$$x_{ij}(s) = \sum_{k=1}^{K}(u_{jk} + w_{ijk})\xi_{jk}(s), \quad \beta_j(s,t) = \sum_{k=1}^{K^2} v_{jk}\zeta_{jk}(s,t)$$

respectively, where $K = 7$ and $\xi_{jk}(s)$ and $\zeta_{jk}(s,t)$ are respectively one and two dimensional basis functions, for which we used the radial basis functions of Kawano and Konishi (2007), and u_{jk}, w_{ijk} in $x_{ij}(s)$ are respectively given as follows:

$$u_{1k} = 0.1k, \quad u_{2k} = \sin k, \quad u_{3k} = \cos k, \quad u_{4k} = \exp(2k/7), \quad u_{5k} = -k^2,$$
$$\boldsymbol{w}_{ij} \sim N_K(\mathbf{0}, \Sigma_w), \quad \boldsymbol{w}_{ij} = (w_{ij1}, \ldots, w_{ijK})', \quad \Sigma_w = (\sigma_w \rho_w^{|k-l|})_{kl}$$

with $\sigma_w = 0.3$ and $\rho_w = 0.5$. Moreover, v_{jk} in $\beta_j(s,t)$ are set to be

$$v_{1k} = 0.5^{(a-4)^2+(b-4)^2}/10, \quad v_{2k} = 0.1a, \quad v_{3k} = -\sin((a-b)^2), \quad v_{4k} = 0, \quad v_{5k} = 0,$$

where $a = 1, \ldots, K$ and $b = 1, \ldots, K$ correspond to $k = (a-1)K + b$. It means that X_4 and X_5 are unnecessary for the model. Furthermore, the error function in the HFLM was generated by the basis expansion with random coefficients:

$$\varepsilon_i(t) = \sum_{k=1}^{K} e_k \xi_k(t), \quad e_k \sim N(0, \sigma_e^2 R_i^{e2}),$$

where $\xi_k(t)$ are basis functions same as $\xi_{jk}(t)$, $\sigma_e = 0.1, 0.3$ and $R_i^e = \mathrm{sd}(y_i(t))$. Then the response is given by

$$y_i(t) = \sum_{j=1}^{p} \int_{s_j(t)}^{t} x_{ij}(s)\beta_j(s,t)ds + \varepsilon_i(t),$$

where we set the lag parameter δ_j included in $s_j(t)$ to be $\delta_j = 0.5$ for all j. Since it is natural that the observed longitudinal data themselves have noises, we added noises to above predictors and responses as follows:

$$x_{ijl} = x_{ij}(s_l) + \varepsilon_{ijl}^{(x)}, \quad y_{il} = y_i(t_l) + \varepsilon_{il}^{(y)},$$

where $l = 1, \ldots, 51$ and $\varepsilon_{ijl}^{(x)}$ and $\varepsilon_{il}^{(y)}$ respectively follow $N(0, \sigma_x^2 R_i^{x2})$ and $N(0, \sigma_y^2 R_i^{y2})$ with $\sigma_x = \sigma_y = 0.3$, $R_i^x = \mathrm{sd}(x_{ij}(s))$, $R_i^y = \mathrm{sd}(y_i(t))$.

We treated x_{ijl} and y_{il} as observed data, then converted them into functional data using B-spline basis. Numbers of basis are set to be 8 for each variable. Next we constructed a design matrix and a response vector in (6) using the FEM. There are several tuning parameters involved in the FEM, described in Section 3.1. We set parameters $N = 13$ and $\mu = 4$, and M_j are set so that $\delta_j = 0.25, 0.50, 0.75$ for all j (true is 0.5). Then parameter $\boldsymbol{\theta}$ in the model is estimated by maximum likelihood method and the penalized likelihood method with the group SCAD regularization. We conducted this strategy for 100 repetitions and for all combinations of $n = 50, 100$, $\sigma_e = 0.1, 0.3$ and $\delta_j = 0.25, 0.50, 0.75$, and then investigated average values of MSE= $\sum_i^n \sum_j^{51} \{y_i(t_j) - \hat{y}_i(t_j)\}^2$ with estimated response functions $\hat{y}_i(t)$, regularization parameter, and numbers of selected variables.

Table 1 shows the results of simulation study. It includes averaged MSEs for the maximum likelihood method (ML) and the group SCAD regularization (gSCAD), averaged value of selected regularization parameters (λ) and ratios of variables selected for 100 repetitions. This table shows that the group SCAD regularization minimized MSEs for all cases. Among δ_j, $\delta_j = 0.25$ and $\delta_j = 0.50$ minimized the MSE, although the true δ_j is 0.5. For the accuracy of variable selection, all cases had tendencies to select correct variables.

5 Real data analysis

We applied the proposed method to the analysis of typhoon data. We investigated which sets of data about the typhoon have an influence on the path of them, using the functional linear model.

The data are available on the website "Digital Typhoon.*" We picked up 88 typhoons which passed around Japan, i.e. passed between north latitude of 30 and 50 and between east longitude of 130 and 150) since 2001 to 2012 from the website. The data contains the position (latitude Y_1 and longitude Y_2), the center atmospheric pressure (X_1), velocity of the wind around the center (X_2) and radii of minor and major storm axes (winds higher than 25 m/s, respectively X_3 and X_4) and gale axes (winds higher than 15 m/s, respectively X_5 and X_6) of typhoons. They are observed every six hours from the generation

*http://agora.ex.nii.ac.jp/digital-typhoon/index.html.en

Table 1: Results on 100 repetitions in simulation studies.

			MSE (SD $\times 10$)		λ	Ratio of selection				
n	σ_e	δ	ML	gSCAD	($\times 10^2$)	X_1	X_2	X_3	X_4	X_5
50	0.1	0.25	3.17 (1.53)	2.93 (1.80)	0.82	0.82	1.00	1.00	0.31	0.01
	0.1	0.50	3.22 (2.01)	2.76 (3.16)	0.71	0.82	1.00	0.99	0.24	0.01
	0.1	0.75	3.49 (2.22)	2.98 (2.80)	0.67	0.88	1.00	1.00	0.29	0.04
	0.3	0.25	5.88 (5.56)	5.18 (6.90)	1.52	0.81	1.00	0.91	0.42	0.01
	0.3	0.50	6.67 (5.47)	5.71 (7.80)	1.35	0.83	0.99	0.84	0.36	0.07
	0.3	0.75	7.24 (6.21)	6.25 (8.64)	1.26	0.90	0.97	0.73	0.40	0.01
100	0.1	0.25	2.80 (1.20)	2.64 (1.24)	0.46	0.76	1.00	1.00	0.10	0.01
	0.1	0.50	2.56 (1.26)	2.30 (1.35)	0.34	0.93	1.00	1.00	0.27	0.02
	0.1	0.75	2.71 (1.43)	2.43 (1.63)	0.30	0.94	1.00	1.00	0.38	0.06
	0.3	0.25	4.46 (3.08)	4.10 (5.66)	0.89	0.69	1.00	0.83	0.18	0.00
	0.3	0.50	4.85 (3.55)	4.44 (5.91)	0.57	0.88	1.00	0.83	0.48	0.01
	0.3	0.75	5.17 (3.54)	4.73 (5.67)	0.52	0.95	1.00	0.80	0.51	0.00

to the disappearance of the typhoons. We aligned the generation and disappearance time of all typhoons by scaling the time points at which the data were observed on [0, 1]. Figure 4 shows some examples of typhoon data. Since the survival times differ for each typhoon numbers and positions of time points also differ, and therefore it is difficult to apply the traditional linear model directly. In this example we examined whether which kind of information of the typhoon affect the location of it. In order to do it we treated the positions as responses and other variables as predictors.

First of the analysis we converted the observed data into functions by the basis expansion, and then centered them. Then we constructed the HFLM (2) by treating the position of typhoons as the functional response and the remaining data as functional predictors. Since the model (2) contains only one response whereas there are two data sets of positions (latitude and longitude), we separately constructed the HFLM with each of the response. Unknown parameters included in the model is estimated by the maximum penalized likelihood method with the group SCAD regularization and a regularization parameter λ involved in the penalized log-likelihood function was selected by BIC. We investigated which variables are selected and surface of the estimated coefficients.

Figures 5 to 10 show estimated coefficient functions for Y_1, Y_2 and $\delta_j = 0.25, 0.50, 0.75$. From these figures we can find that some of coefficient surfaces are estimated to be zero functions, which lead to eliminations of corresponding variables from the model. When the response is latitude only one variable, the velocity of the wind (X_2) is removed from the model for all cases of δ_j. On the other hand when the response is longitude the pressure (X_1), the velocity of the wind (X_2) and the minor gale axis (X_6), and in addition the major gale axis (X_5) for cases $\delta_j = 0.50$ and $\delta_j = 0.75$ are removed from the model. Next we focus on features of coefficient surfaces of selected variables. For example, $\beta_1(s, t)$ in Figure 9 shows that there is a positive weight to the variable around $t = s$ direction, and this weight decreases for the $t = 1 - s$ direction with s goes to 0. It indicates that the last information of the predictor has positive weight to the response, whereas this influence gradually vanishes as it becomes a thing of the past. Same implications are

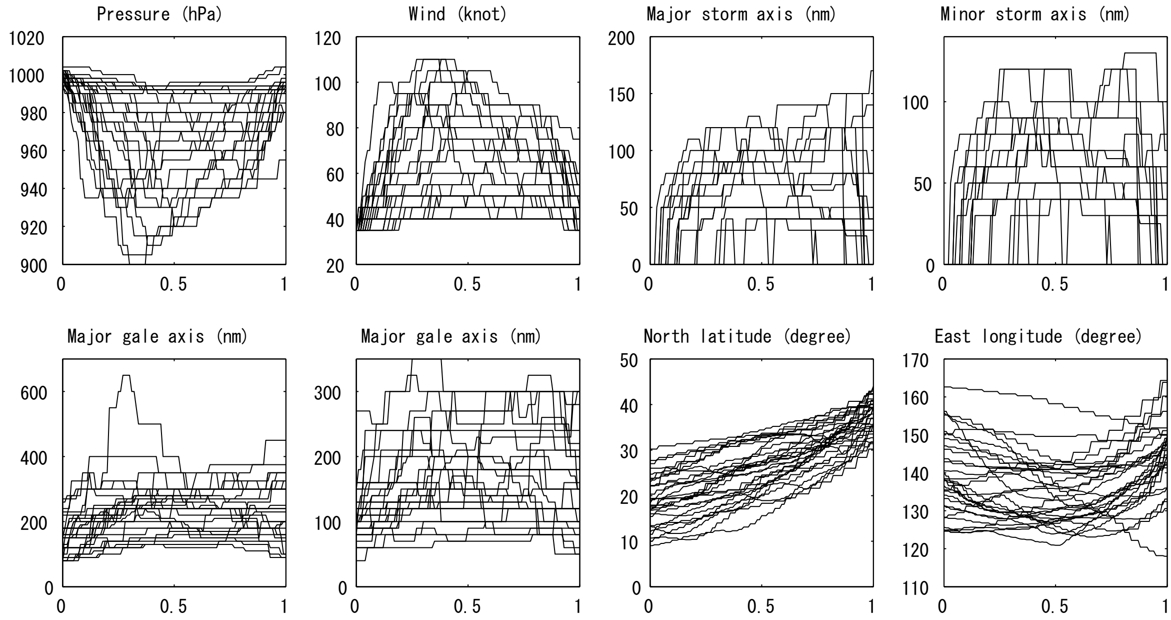

Figure 4: Examples of typhoon data. Bottom right 2 plots (north latitude and east longitude) are responses and remaining data are predictors.

obtained from other coefficient surfaces. Note that, however, for the $\delta_j = 0.25$ case there are many fluctuations on the surfaces and therefore it is difficult to obtain insights from them than other cases of δ_j.

6 Concluding remarks

We have proposed the method for variable selection of variable selection where predictors and a response are given as functions. Especially when they are functions of time we need to take account of the dependency between predictors and a response, and therefore we applied the historical functional linear model. Unknown parameters included in the historical functional linear model are estimated by the maximum penalized likelihood method with the group SCAD regularization, and a regularization parameter included in the model is selected by a BIC type model selection criterion. In addition to the estimation of parameters we can also select functional predictors in the regression model due to the property of the L_1 type regularization. Simulation and real data analysis revealed that the proposed method appropriately selected variables.

In this work we assumed that several tuning parameters except for the regularization parameter λ are given. Especially for the lag parameter δ_j, though, we think it is a crucial issue to select these values objectively, since it decides how long the time is included in the model. However, we cannot directly apply model selection criteria since the "sample size" in model (6) changes as the lag parameter δ_j changes. Therefore future works include the selection of them. Furthermore we can consider extending the HFLM to that with multiple response. Then we can treat two responses, such as the longitude and the latitude of typhoon data used in Section 5, together in one model to take the correlation among responses into consideration.

Acknowledgment

This work was supported by Grant-in-Aid for Young Scientists (B) No. 25730017 of JSPS.

References

Cai, T. and Hall, P. (2006), "Prediction in functional linear regression," *Ann. Statist.*, 34, 2159–2179.

Cai, T. and Yuan, M. (2012), "Minimax and Adaptive Prediction for Functional Linear Regression," *J. Amer. Statist. Assoc.*, 107, 1201–1216.

Fahrmeir, L., Kneib, T., Lang, S., and Marx, B. (2013), *Regression*, Springer.

Fan, J. and Li, R. (2001), "Variable selection via nonconcave penalized likelihood and its oracle properties," *J. Amer. Statist. Assoc.*, 96, 1348–1360.

— (2004), "New estimation and model selection procedures for semiparametric modeling in longitudinal data analysis," *J. Amer. Statist. Assoc.*, 99, 710–723.

Hall, P. and Horowitz, J. L. (2007), "Methodology and convergence rates for functional linear regression," *The Annals of Statistics*, 35, 70–91.

Harezlak, J., Coull, B., Laird, N., Magari, S., and Christiani, D. (2007), "Penalized solutions to functional regression problems," *Comput. Statist. Data Anal.*, 51, 4911–4925.

Hastie, T. and Tibshirani, R. (1993), "Varying-coefficient models," *J. Roy. Statist. Soc. Ser. B*, 55, 757–796.

Hoover, D., Rice, J., Wu, C., and Yang, L. (1998), "Nonparametric smoothing estimates of time-varying coefficient models with longitudinal data," *Biometrika*, 85, 809–822.

Horváth, L. and Kokoszka, P. (2012), *Inference for functional data with applications*, New York: Springer.

James, G. (2002), "Generalized linear models with functional predictors," *J. Roy. Statist. Soc. Ser. B*, 64, 411–432.

James, G. and Silverman, B. (2005), "Functional adaptive model estimation," *J. Amer. Statist. Assoc.*, 100, 565–576.

Kawano, S. and Konishi, S. (2007), "Nonlinear regression modeling via regularized Gaussian basis functions," *Bull. Inform. Cybern.*, 39, 83–96.

Leng, C., Lin, Y., and Wahba, G. (2006), "A note on the lasso and related procedures in model selection," *Statist. Sinica*, 16, 1273–1284.

Malfait, N. and Ramsay, J. (2003), "The historical functional linear model," *Canad. J. Statist.*, 31, 115–128.

Matsui, H., Kawano, S., and Konishi, S. (2009), "Regularized functional regression modeling for functional response and predictors," *Journal of Mathematics for Industry*, 1, 17–25.

Matsui, H. and Konishi, S. (2011), "Variable selection for functional regression models via the L1 regularization," *Comput. Statist. Data Anal.*, 55, 3304–3310.

Müller, H. and Stadtmüller, U. (2005), "Generalized functional linear models," *Ann. Statist.*, 33, 774–805.

Müller, H. and Yao, F. (2008), "Functional additive models," *J. Amer. Statist. Assoc.*, 103, 1534–1544.

Ramsay, J. and Dalzell, C. (1991), "Some tools for functional data analysis," *J. Roy. Statist. Soc. Ser. B*, 53, 539–572.

Ramsay, J. and Silverman, B. (2002), *Applied functional data analysis: methods and case studies*, New York: Springer Verlag.

— (2005), *Functional data analysis 2nd ed.*, New York: Springer Verlag.

Şentürk, D. and Müller, H. (2010), "Functional varying coefficient models for longitudinal data," *J. Amer. Statist. Assoc.*, 105, 1256–1264.

Şentürk, D. and Müller, H.-G. (2008), "Generalized varying coefficient models for longitudinal data," *Biometrika*, 95, 653–666.

Tibshirani, R. (1996), "Regression shrinkage and selection via the lasso," *J. Roy. Statist. Soc. Ser. B*, 58, 267–288.

Wang, H., Li, R., and Tsai, C. (2007), "Tuning parameter selectors for the smoothly clipped absolute deviation method," *Biometrika*, 94, 553–568.

Yao, F., Müller, H., and Wang, J. (2005), "Functional linear regression analysis for longitudinal data," *Ann. Statist.*, 33, 2873–2903.

Zhang, C. (2010), "Nearly unbiased variable selection under minimax concave penalty," *Ann. Statist.*, 38, 894–942.

Zhang, Y., Li, R., and Tsai, C. (2010), "Regularization parameter selections via generalized information criterion," *J. Amer. Statist. Assoc.*, 105, 312–323.

Zou, H. and Hastie, T. (2005), "Regularization and variable selection via the elastic net," *J. Roy. Statist. Soc. Ser. B*, 67, 301–320.

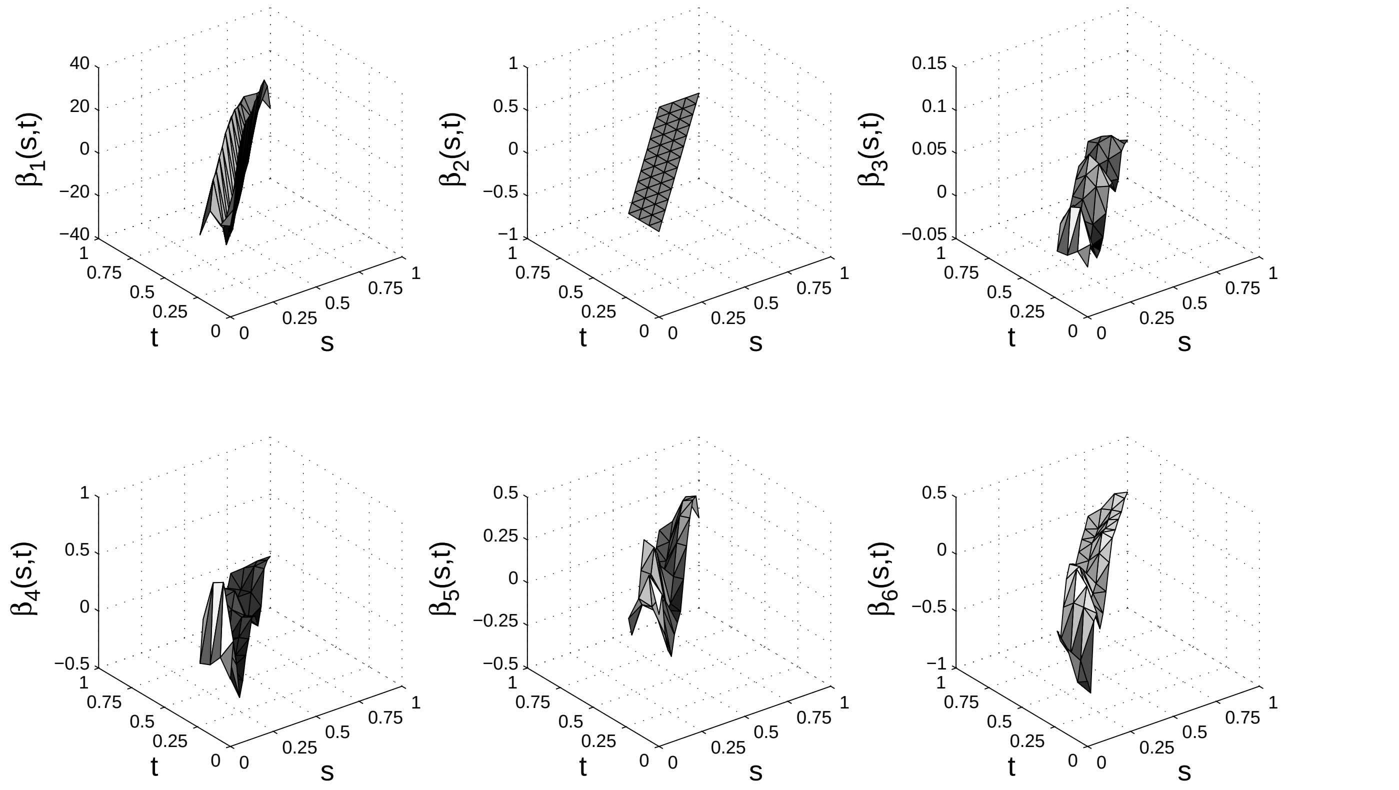

Figure 5: Coefficient functions for Y_1 and $\delta_j = 0.25$.

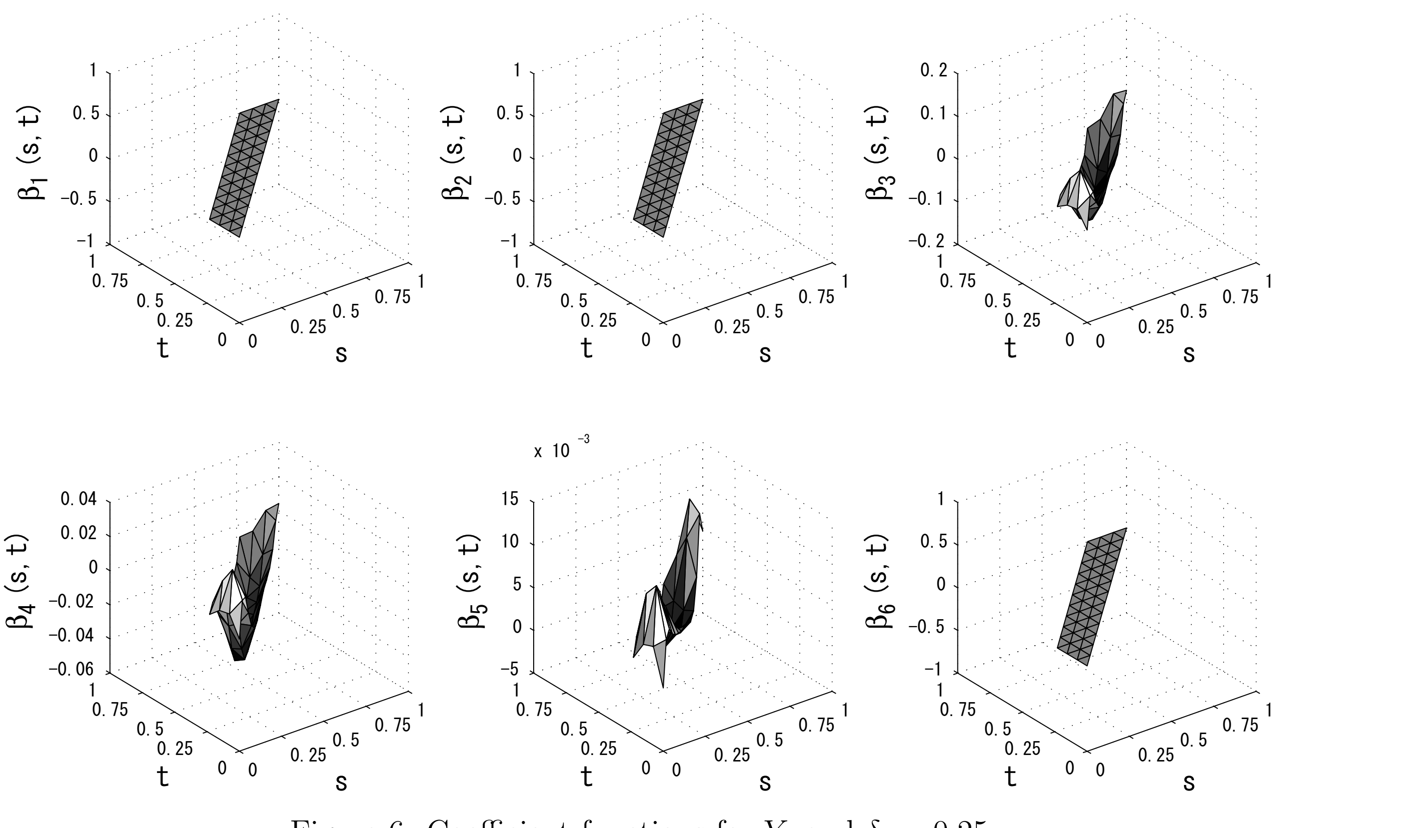

Figure 6: Coefficient functions for Y_2 and $\delta_j = 0.25$.

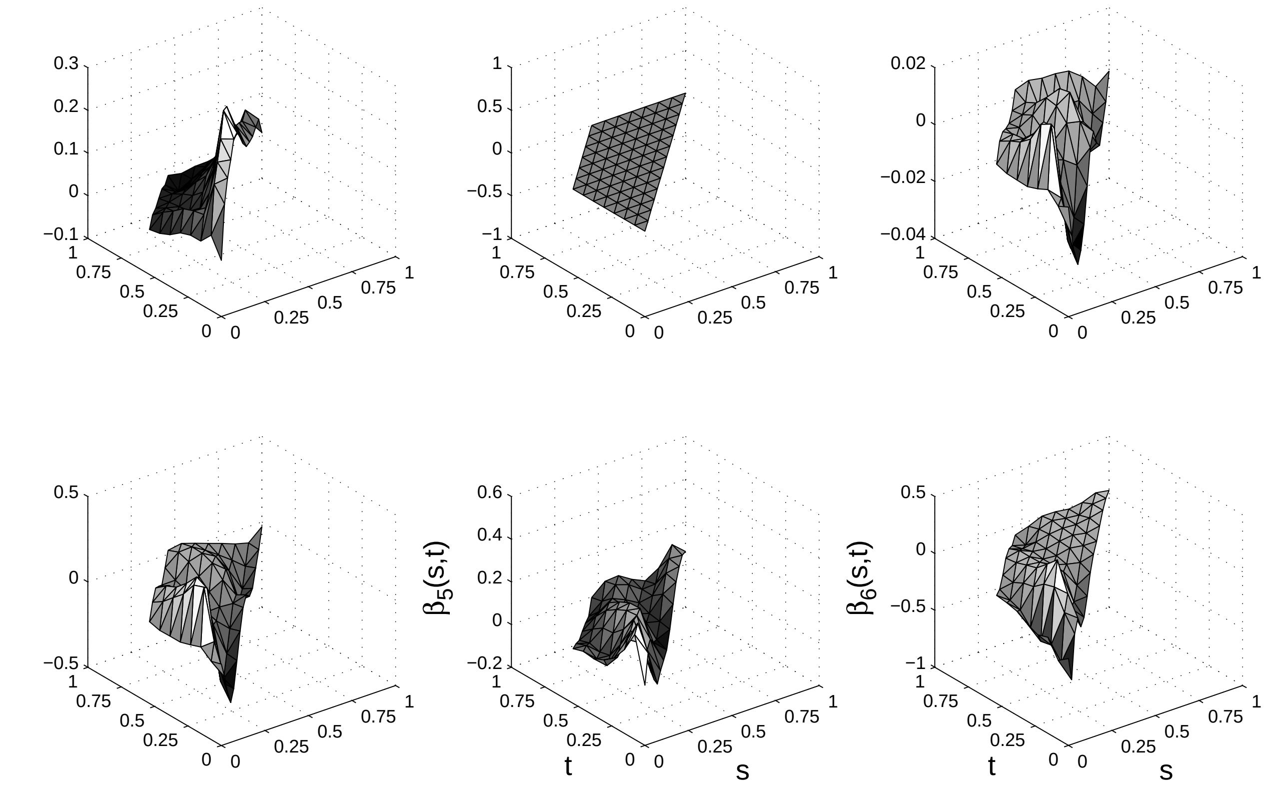

Figure 7: Coefficient functions for Y_1 and $\delta_j = 0.50$.

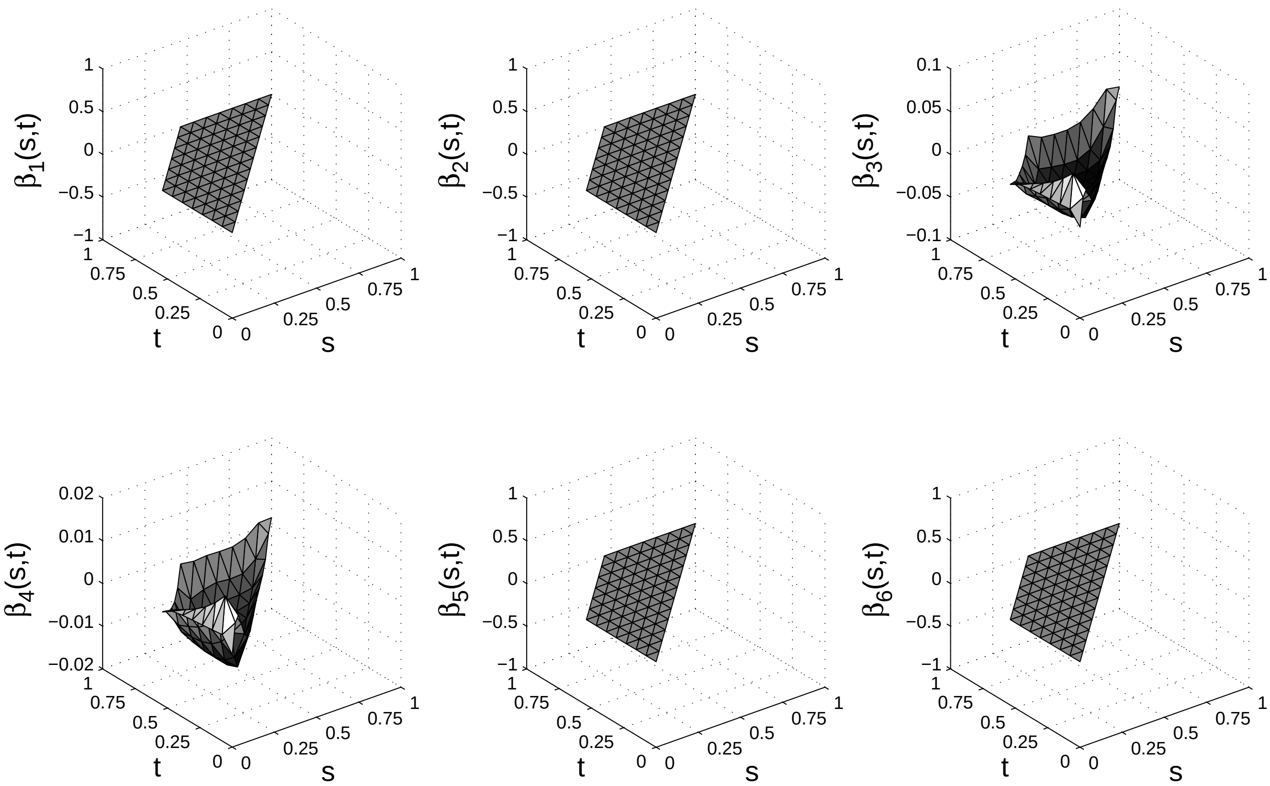

Figure 8: Coefficient functions for Y_2 and $\delta_j = 0.50$.

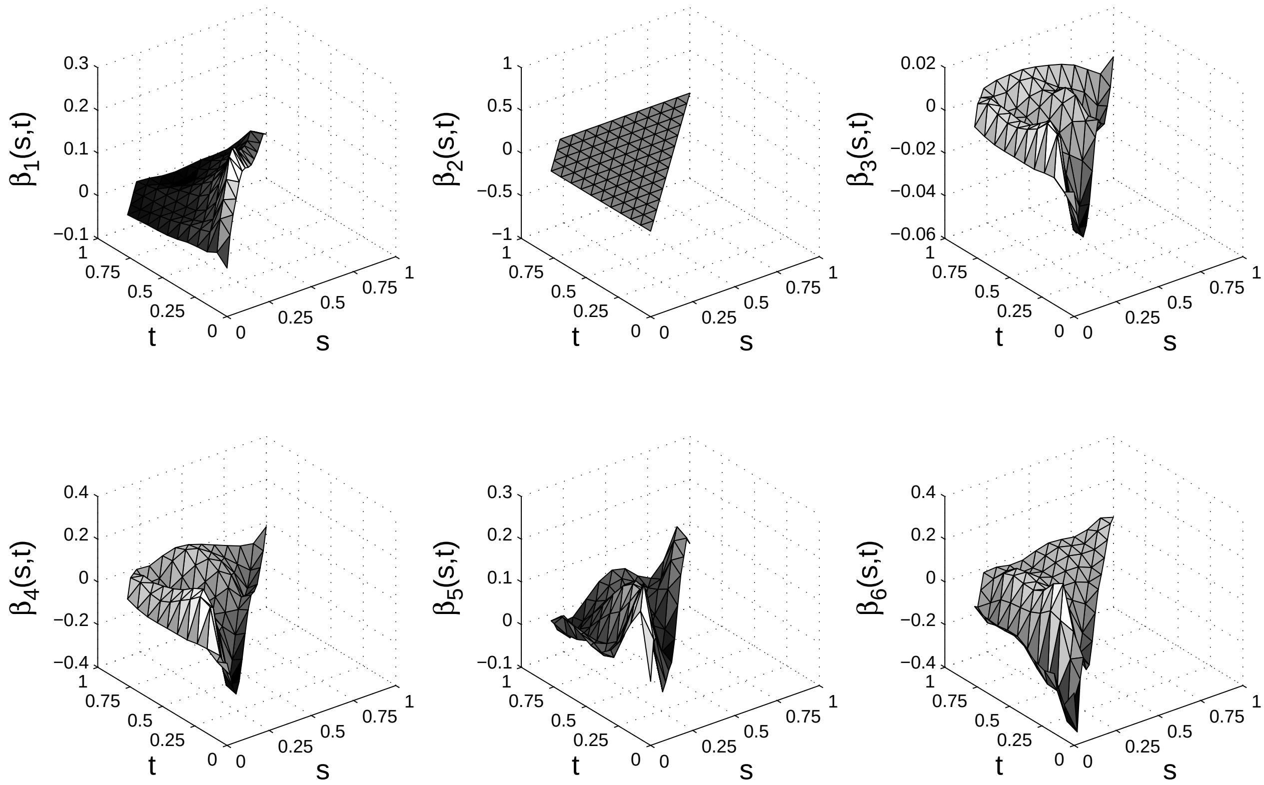

Figure 9: Coefficient functions for Y_1 and $\delta_j = 0.75$.

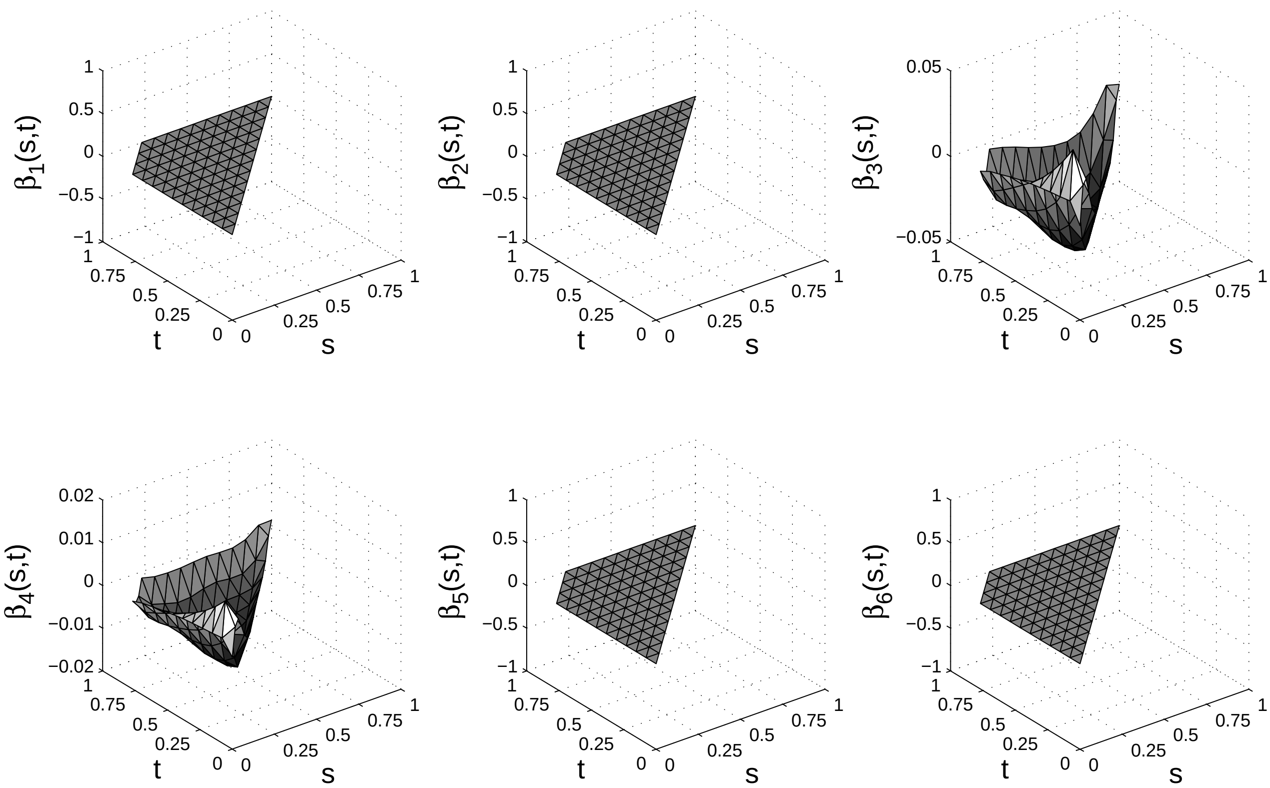

Figure 10: Coefficient functions for Y_2 and $\delta_j = 0.75$.

List of MI Preprint Series, Kyushu University

The Global COE Program
Math-for-Industry Education & Research Hub

MI

MI2008-1 Takahiro ITO, Shuichi INOKUCHI & Yoshihiro MIZOGUCHI
Abstract collision systems simulated by cellular automata

MI2008-2 Eiji ONODERA
The intial value problem for a third-order dispersive flow into compact almost Hermitian manifolds

MI2008-3 Hiroaki KIDO
On isosceles sets in the 4-dimensional Euclidean space

MI2008-4 Hirofumi NOTSU
Numerical computations of cavity flow problems by a pressure stabilized characteristic-curve finite element scheme

MI2008-5 Yoshiyasu OZEKI
Torsion points of abelian varieties with values in nfinite extensions over a p-adic field

MI2008-6 Yoshiyuki TOMIYAMA
Lifting Galois representations over arbitrary number fields

MI2008-7 Takehiro HIROTSU & Setsuo TANIGUCHI
The random walk model revisited

MI2008-8 Silvia GANDY, Masaaki KANNO, Hirokazu ANAI & Kazuhiro YOKOYAMA
Optimizing a particular real root of a polynomial by a special cylindrical algebraic decomposition

MI2008-9 Kazufumi KIMOTO, Sho MATSUMOTO & Masato WAKAYAMA
Alpha-determinant cyclic modules and Jacobi polynomials

MI2008-10 Sangyeol LEE & Hiroki MASUDA
Jarque-Bera Normality Test for the Driving Lévy Process of a Discretely Observed Univariate SDE

MI2008-11 Hiroyuki CHIHARA & Eiji ONODERA
A third order dispersive flow for closed curves into almost Hermitian manifolds

MI2008-12 Takehiko KINOSHITA, Kouji HASHIMOTO and Mitsuhiro T. NAKAO
On the L^2 a priori error estimates to the finite element solution of elliptic problems with singular adjoint operator

MI2008-13 Jacques FARAUT and Masato WAKAYAMA
Hermitian symmetric spaces of tube type and multivariate Meixner-Pollaczek polynomials

MI2008-14 Takashi NAKAMURA
Riemann zeta-values, Euler polynomials and the best constant of Sobolev inequality

MI2008-15 Takashi NAKAMURA
Some topics related to Hurwitz-Lerch zeta functions

MI2009-1 Yasuhide FUKUMOTO
Global time evolution of viscous vortex rings

MI2009-2 Hidetoshi MATSUI & Sadanori KONISHI
Regularized functional regression modeling for functional response and predictors

MI2009-3 Hidetoshi MATSUI & Sadanori KONISHI
Variable selection for functional regression model via the L_1 regularization

MI2009-4 Shuichi KAWANO & Sadanori KONISHI
Nonlinear logistic discrimination via regularized Gaussian basis expansions

MI2009-5 Toshiro HIRANOUCHI & Yuichiro TAGUCHII
Flat modules and Groebner bases over truncated discrete valuation rings

MI2009-6 Kenji KAJIWARA & Yasuhiro OHTA
Bilinearization and Casorati determinant solutions to non-autonomous 1+1 dimensional discrete soliton equations

MI2009-7 Yoshiyuki KAGEI
Asymptotic behavior of solutions of the compressible Navier-Stokes equation around the plane Couette flow

MI2009-8 Shohei TATEISHI, Hidetoshi MATSUI & Sadanori KONISHI
Nonlinear regression modeling via the lasso-type regularization

MI2009-9 Takeshi TAKAISHI & Masato KIMURA
Phase field model for mode III crack growth in two dimensional elasticity

MI2009-10 Shingo SAITO
Generalisation of Mack's formula for claims reserving with arbitrary exponents for the variance assumption

MI2009-11 Kenji KAJIWARA, Masanobu KANEKO, Atsushi NOBE & Teruhisa TSUDA
Ultradiscretization of a solvable two-dimensional chaotic map associated with the Hesse cubic curve

MI2009-12 Tetsu MASUDA
Hypergeometric τ-functions of the q-Painlevé system of type $E_8^{(1)}$

MI2009-13 Hidenao IWANE, Hitoshi YANAMI, Hirokazu ANAI & Kazuhiro YOKOYAMA
A Practical Implementation of a Symbolic-Numeric Cylindrical Algebraic Decomposition for Quantifier Elimination

MI2009-14 Yasunori MAEKAWA
On Gaussian decay estimates of solutions to some linear elliptic equations and its applications

MI2009-15 Yuya ISHIHARA & Yoshiyuki KAGEI
Large time behavior of the semigroup on L^p spaces associated with the linearized compressible Navier-Stokes equation in a cylindrical domain

MI2009-16 Chikashi ARITA, Atsuo KUNIBA, Kazumitsu SAKAI & Tsuyoshi SAWABE
Spectrum in multi-species asymmetric simple exclusion process on a ring

MI2009-17 Masato WAKAYAMA & Keitaro YAMAMOTO
Non-linear algebraic differential equations satisfied by certain family of elliptic functions

MI2009-18 Me Me NAING & Yasuhide FUKUMOTO
Local Instability of an Elliptical Flow Subjected to a Coriolis Force

MI2009-19 Mitsunori KAYANO & Sadanori KONISHI
Sparse functional principal component analysis via regularized basis expansions and its application

MI2009-20 Shuichi KAWANO & Sadanori KONISHI
Semi-supervised logistic discrimination via regularized Gaussian basis expansions

MI2009-21 Hiroshi YOSHIDA, Yoshihiro MIWA & Masanobu KANEKO
Elliptic curves and Fibonacci numbers arising from Lindenmayer system with symbolic computations

MI2009-22 Eiji ONODERA
A remark on the global existence of a third order dispersive flow into locally Hermitian symmetric spaces

MI2009-23 Stjepan LUGOMER & Yasuhide FUKUMOTO
Generation of ribbons, helicoids and complex scherk surface in laser-matter Interactions

MI2009-24 Yu KAWAKAMI
Recent progress in value distribution of the hyperbolic Gauss map

MI2009-25 Takehiko KINOSHITA & Mitsuhiro T. NAKAO
On very accurate enclosure of the optimal constant in the a priori error estimates for H_0^2-projection

MI2009-26 Manabu YOSHIDA
Ramification of local fields and Fontaine's property (Pm)

MI2009-27 Yu KAWAKAMI
Value distribution of the hyperbolic Gauss maps for flat fronts in hyperbolic three-space

MI2009-28 Masahisa TABATA
Numerical simulation of fluid movement in an hourglass by an energy-stable finite element scheme

MI2009-29 Yoshiyuki KAGEI & Yasunori MAEKAWA
Asymptotic behaviors of solutions to evolution equations in the presence of translation and scaling invariance

MI2009-30 Yoshiyuki KAGEI & Yasunori MAEKAWA
On asymptotic behaviors of solutions to parabolic systems modelling chemotaxis

MI2009-31 Masato WAKAYAMA & Yoshinori YAMASAKI
Hecke's zeros and higher depth determinants

MI2009-32 Olivier PIRONNEAU & Masahisa TABATA
Stability and convergence of a Galerkin-characteristics finite element scheme of lumped mass type

MI2009-33 Chikashi ARITA
Queueing process with excluded-volume effect

MI2009-34 Kenji KAJIWARA, Nobutaka NAKAZONO & Teruhisa TSUDA
Projective reduction of the discrete Painlevé system of type$(A_2 + A_1)^{(1)}$

MI2009-35 Yosuke MIZUYAMA, Takamasa SHINDE, Masahisa TABATA & Daisuke TAGAMI
Finite element computation for scattering problems of micro-hologram using DtN map

MI2009-36 Reiichiro KAWAI & Hiroki MASUDA
Exact simulation of finite variation tempered stable Ornstein-Uhlenbeck processes

MI2009-37 Hiroki MASUDA
On statistical aspects in calibrating a geometric skewed stable asset price model

MI2010-1 Hiroki MASUDA
Approximate self-weighted LAD estimation of discretely observed ergodic Ornstein-Uhlenbeck processes

MI2010-2 Reiichiro KAWAI & Hiroki MASUDA
Infinite variation tempered stable Ornstein-Uhlenbeck processes with discrete observations

MI2010-3 Kei HIROSE, Shuichi KAWANO, Daisuke MIIKE & Sadanori KONISHI
Hyper-parameter selection in Bayesian structural equation models

MI2010-4 Nobuyuki IKEDA & Setsuo TANIGUCHI
The Itô-Nisio theorem, quadratic Wiener functionals, and 1-solitons

MI2010-5 Shohei TATEISHI & Sadanori KONISHI
Nonlinear regression modeling and detecting change point via the relevance vector machine

MI2010-6 Shuichi KAWANO, Toshihiro MISUMI & Sadanori KONISHI
Semi-supervised logistic discrimination via graph-based regularization

MI2010-7 Teruhisa TSUDA
UC hierarchy and monodromy preserving deformation

MI2010-8 Takahiro ITO
Abstract collision systems on groups

MI2010-9 Hiroshi YOSHIDA, Kinji KIMURA, Naoki YOSHIDA, Junko TANAKA & Yoshihiro MIWA
An algebraic approach to underdetermined experiments

MI2010-10 Kei HIROSE & Sadanori KONISHI
Variable selection via the grouped weighted lasso for factor analysis models

MI2010-11 Katsusuke NABESHIMA & Hiroshi YOSHIDA
Derivation of specific conditions with Comprehensive Groebner Systems

MI2010-12 Yoshiyuki KAGEI, Yu NAGAFUCHI & Takeshi SUDOU
Decay estimates on solutions of the linearized compressible Navier-Stokes equation around a Poiseuille type flow

MI2010-13 Reiichiro KAWAI & Hiroki MASUDA
On simulation of tempered stable random variates

MI2010-14 Yoshiyasu OZEKI
Non-existence of certain Galois representations with a uniform tame inertia weight

MI2010-15 Me Me NAING & Yasuhide FUKUMOTO
Local Instability of a Rotating Flow Driven by Precession of Arbitrary Frequency

MI2010-16 Yu KAWAKAMI & Daisuke NAKAJO
The value distribution of the Gauss map of improper affine spheres

MI2010-17 Kazunori YASUTAKE
On the classification of rank 2 almost Fano bundles on projective space

MI2010-18 Toshimitsu TAKAESU
Scaling limits for the system of semi-relativistic particles coupled to a scalar bose field

MI2010-19 Reiichiro KAWAI & Hiroki MASUDA
Local asymptotic normality for normal inverse Gaussian Lévy processes with high-frequency sampling

MI2010-20 Yasuhide FUKUMOTO, Makoto HIROTA & Youichi MIE
Lagrangian approach to weakly nonlinear stability of an elliptical flow

MI2010-21 Hiroki MASUDA
Approximate quadratic estimating function for discretely observed Lévy driven SDEs with application to a noise normality test

MI2010-22 Toshimitsu TAKAESU
A Generalized Scaling Limit and its Application to the Semi-Relativistic Particles System Coupled to a Bose Field with Removing Ultraviolet Cutoffs

MI2010-23 Takahiro ITO, Mitsuhiko FUJIO, Shuichi INOKUCHI & Yoshihiro MIZOGUCHI
Composition, union and division of cellular automata on groups

MI2010-24 Toshimitsu TAKAESU
A Hardy's Uncertainty Principle Lemma in Weak Commutation Relations of Heisenberg-Lie Algebra

MI2010-25 Toshimitsu TAKAESU
On the Essential Self-Adjointness of Anti-Commutative Operators

MI2010-26 Reiichiro KAWAI & Hiroki MASUDA
On the local asymptotic behavior of the likelihood function for Meixner Lévy processes under high-frequency sampling

MI2010-27 Chikashi ARITA & Daichi YANAGISAWA
Exclusive Queueing Process with Discrete Time

MI2010-28 Jun-ichi INOGUCHI, Kenji KAJIWARA, Nozomu MATSUURA & Yasuhiro OHTA
Motion and Bäcklund transformations of discrete plane curves

MI2010-29 Takanori YASUDA, Masaya YASUDA, Takeshi SHIMOYAMA & Jun KOGURE
On the Number of the Pairing-friendly Curves

MI2010-30 Chikashi ARITA & Kohei MOTEGI
Spin-spin correlation functions of the q-VBS state of an integer spin model

MI2010-31 Shohei TATEISHI & Sadanori KONISHI
Nonlinear regression modeling and spike detection via Gaussian basis expansions

MI2010-32 Nobutaka NAKAZONO
Hypergeometric τ functions of the q-Painlevé systems of type $(A_2 + A_1)^{(1)}$

MI2010-33 Yoshiyuki KAGEI
Global existence of solutions to the compressible Navier-Stokes equation around parallel flows

MI2010-34 Nobushige KUROKAWA, Masato WAKAYAMA & Yoshinori YAMASAKI
Milnor-Selberg zeta functions and zeta regularizations

MI2010-35 Kissani PERERA & Yoshihiro MIZOGUCHI
Laplacian energy of directed graphs and minimizing maximum outdegree algorithms

MI2010-36 Takanori YASUDA
CAP representations of inner forms of $Sp(4)$ with respect to Klingen parabolic subgroup

MI2010-37 Chikashi ARITA & Andreas SCHADSCHNEIDER
Dynamical analysis of the exclusive queueing process

MI2011-1 Yasuhide FUKUMOTO& Alexander B. SAMOKHIN
Singular electromagnetic modes in an anisotropic medium

MI2011-2 Hiroki KONDO, Shingo SAITO & Setsuo TANIGUCHI
Asymptotic tail dependence of the normal copula

MI2011-3 Takehiro HIROTSU, Hiroki KONDO, Shingo SAITO, Takuya SATO, Tatsushi TANAKA & Setsuo TANIGUCHI
Anderson-Darling test and the Malliavin calculus

MI2011-4 Hiroshi INOUE, Shohei TATEISHI & Sadanori KONISHI
Nonlinear regression modeling via Compressed Sensing

MI2011-5 Hiroshi INOUE
Implications in Compressed Sensing and the Restricted Isometry Property

MI2011-6 Daeju KIM & Sadanori KONISHI
Predictive information criterion for nonlinear regression model based on basis expansion methods

MI2011-7 Shohei TATEISHI, Chiaki KINJYO & Sadanori KONISHI
Group variable selection via relevance vector machine

MI2011-8 Jan BREZINA & Yoshiyuki KAGEI
Decay properties of solutions to the linearized compressible Navier-Stokes equation around time-periodic parallel flow
Group variable selection via relevance vector machine

MI2011-9 Chikashi ARITA, Arvind AYYER, Kirone MALLICK & Sylvain PROLHAC
Recursive structures in the multispecies TASEP

MI2011-10 Kazunori YASUTAKE
On projective space bundle with nef normalized tautological line bundle

MI2011-11 Hisashi ANDO, Mike HAY, Kenji KAJIWARA & Tetsu MASUDA
An explicit formula for the discrete power function associated with circle patterns of Schramm type

MI2011-12 Yoshiyuki KAGEI
Asymptotic behavior of solutions to the compressible Navier-Stokes equation around a parallel flow

MI2011-13 Vladimír CHALUPECKÝ & Adrian MUNTEAN
Semi-discrete finite difference multiscale scheme for a concrete corrosion model: approximation estimates and convergence

MI2011-14 Jun-ichi INOGUCHI, Kenji KAJIWARA, Nozomu MATSUURA & Yasuhiro OHTA
Explicit solutions to the semi-discrete modified KdV equation and motion of discrete plane curves

MI2011-15 Hiroshi INOUE
A generalization of restricted isometry property and applications to compressed sensing

MI2011-16 Yu KAWAKAMI
A ramification theorem for the ratio of canonical forms of flat surfaces in hyperbolic three-space

MI2011-17 Naoyuki KAMIYAMA
Matroid intersection with priority constraints

MI2012-1 Kazufumi KIMOTO & Masato WAKAYAMA
Spectrum of non-commutative harmonic oscillators and residual modular forms

MI2012-2 Hiroki MASUDA
Mighty convergence of the Gaussian quasi-likelihood random fields for ergodic Levy driven SDE observed at high frequency

MI2013-6 Hiroshi INOUE
Improved bounds on Restricted isometry for compressed sensing

MI2013-7 Hidetoshi MATSUI
Variable and boundary selection for functional data via multiclass logistic regression modeling

MI2013-8 Hidetoshi MATSUI
Variable selection for varying coefficient models with the sparse regularization

MI2013-9 Naoyuki KAMIYAMA
Packing Arborescences in Acyclic Temporal Networks

MI2013-10 Masato WAKAYAMA
Equivalence between the eigenvalue problem of non-commutative harmonic oscillators and existence of holomorphic solutions of Heun's differential equations, eigenstates degeneration, and Rabi's model

MI2013-11 Masatoshi OKITA
Optimal decay rate for strong solutions in critical spaces to the compressible Navier-Stokes equations

MI2013-12 Shuichi KAWANO, Ibuki HOSHINA, Kazuki MATSUDA & Sadanori KONISHI
Predictive model selection criteria for Bayesian lasso

MI2013-13 Hayato CHIBA
The First Painleve Equation on the Weighted Projective Space

MI2013-14 Hidetoshi MATSUI
Variable selection for functional linear models with functional predictors and a functional response