Adaptive basis expansion via the extended fused lasso

Kim, Daeju
Graduate School of Mathematics, Kyushu University

Kawano, Shuichi
Department of Mathematical Sciences, Graduate School of Engineering, Osaka Prefecture University

Ninomiya, Yoshiyuki
Institute of Mathematics for Industry, Kyushu University

出版情報：Computational Statistics. 29 (5), pp.1005-1023, 2014-10-01. Springer Berlin Heidelberg
バージョン：
権利関係：

MI Preprint Series
Kyushu University
The Global COE Program
Math-for-Industry Education & Research Hub

Adaptive basis expansion via the extended fused lasso

Daeju Kim, Shuichi Kawano
& Yoshiyuki Ninomiya

MI 2012-11

(Received November 27, 2012)

Faculty of Mathematics
Kyushu University
Fukuoka, JAPAN

Adaptive basis expansion via the extended fused lasso

Daeju Kim[1], Shuichi Kawano[2] and Yoshiyuki Ninomiya[3]

[1] *Graduate School of Mathematics, Kyushu University,*

744 Motooka, Nishi-ku, Fukuoka 819-0395, Japan.

[2] *Department of Mathematical Sciences,*

Graduate School of Engineering, Osaka Prefecture University,

1-1 Gakuen-cho, Sakai, Osaka 599-8531, Japan.

[3] *Institute of Mathematics for Industry, Kyushu University,*

744 Motooka, Nishi-ku, Fukuoka 819-0395, Japan.

E-mail: [1]*t-kin@math.kyushu-u.ac.jp,* [2]*skawano@ms.osakafu-u.ac.jp,* [3]*nino@imi.kyushu-u.ac.jp*

Abstract

We propose a nonlinear regression model that uses basis expansion for the case where the underlying regression function has inhomogeneous smoothness. In this case, conventional nonlinear regression models tend to over- or underfit where the function is smoother or less smooth, respectively. We begin by roughly approximating the underlying regression function with a locally linear function. We then extend the fused lasso signal approximator and thereby develop a fast and efficient algorithm. We next use the residuals between the locally linear functions and the data to adaptively prepare the basis functions. Finally, using a regularization method, we construct a nonlinear regression model with these basis functions. To select the optimal value of the tuning parameter for the regularization method, we provide an explicit form of the generalized information criterion. The validity of our proposed method is then demonstrated through several numerical examples.

1 Introduction

Recently, nonlinear regression models with basis expansion have received considerable attention in various statistical and engineering fields. Basis expansion is widely used as an effective approach for analyzing data with a complex structure. The essential idea behind basis expansion is to represent the underlying regression function as a linear combination of known nonlinear functions, which are called basis functions. In constructing the statistical model, various basis functions, such as natural cubic splines, B-splines, and radial basis functions (Green and Silverman 1994; de Boor 2001; Hastie *et al*. 2009) are used according to the structure of the data or the purpose of the analysis.

While basis expansion is shown to work well in many situations, it is often inappropriate when the underlying regression function has inhomogeneous smoothness. Let us call the region where the function is smoother the strongly smooth region and that where it is less smooth the weakly smooth region. Basis expansion often leads to underfitting in the strongly smooth region and overfitting in the weakly smooth region. Loader (1999) overcame this problem using the local likelihood, which is defined as a locally weighted log-likelihood with weights determined by a kernel function and a bandwidth. Miyata and Shen (2003) also resolved the problem using free-knot splines, along with the technique of variable multiple knots and an evolutionary Monte Carlo algorithm. Although these methods have performed well in simulations, they bear intensive computational burdens.

In this study, our aim is to estimate an underlying function with inhomogeneous smoothness in a way that is not highly computationally demanding. To do so, we propose a fast and efficient nonlinear regression modeling method that uses a given number of Gaussian basis functions with a specified center and bandwidth. Our strategy is based on the idea that the variation in the data can be decomposed as global and local variations. We regard the global variation as that which is caught by a locally smooth function, and

indeed we catch it by a locally linear function. Regions in which the underlying function is weakly smooth are correlated with large local variation. Since the expression of weakly smooth underlying functions requires many basis functions, we increase the number of basis functions on such regions.

To achieve our aim, we perform three stages, as follow. In the first stage, the global variation is caught by an extended fused lasso signal approximator (eFLSA) using locally linear functions. In the second stage, we measure the magnitude of the local variation by the residual sum of squares between the global variation function and the given data. In the final stage, the magnitude obtained in the previous stage is used to adaptively define basis functions. After setting these adaptive basis functions in all regions, we construct a nonlinear regression model by using the ridge regularization method.

The remainder of this article is organized as follows. In Section 2, we describe a framework for nonlinear regression models based on basis expansion and describe the fused lasso signal approximator (FLSA; Friedman *et al.*, 2007). Section 3 describes the eFLSA and presents our procedure, which uses it. In Section 4, we investigate the performance of our procedure by performing Monte Carlo simulations. Some concluding remarks are presented in Section 5.

2 Background

2.1 Nonlinear regression model with basis expansion

Suppose that we have n independent observations $\{(x_i, y_i);\ i = 1, \ldots, n\}$ in terms of an explanatory variable x and a response variable y. We consider a regression model

$$y_i = g(x_i) + \varepsilon_i, \quad i = 1, \ldots, n, \tag{1}$$

where $g(\cdot)$ is an unknown underlying function and the ε_is are independently distributed according to $\mathrm{N}(0, \sigma^2)$. It is assumed that the function $g(\cdot)$ can be expressed as a linear combination of basis functions $\phi_k(x)$ $(k = 1, \ldots, m)$ in the form

$$g(x; \boldsymbol{w}) = \sum_{k=1}^{m} w_k \phi_k(x) = \boldsymbol{w}^{\mathrm{T}} \boldsymbol{\phi}(x), \tag{2}$$

where $\boldsymbol{\phi}(x) = (\phi_1(x), \ldots, \phi_m(x))^{\mathrm{T}}$ is a vector in which the components are known basis functions and $\boldsymbol{w} = (w_1, \ldots, w_m)^{\mathrm{T}}$ is a vector in which the components are the unknown coefficients for the basis functions. As mentioned above, the basis expansion may use a variety of basis functions such as natural cubic splines, B-splines, and radial basis functions (de Boor 2001; Green and Silverman 1994; Hastie *et al*. 2009), according to the structure of the data and the purpose of the analysis.

From (1) and (2), the probability density function of y_i is

$$f(y_i | x_i; \boldsymbol{w}, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{\left\{ y_i - \boldsymbol{w}^{\mathrm{T}} \boldsymbol{\phi}(x_i) \right\}^2}{2\sigma^2} \right], \quad i = 1, \ldots, n.$$

Note that the unknown parameters in the model are $\boldsymbol{w} = (w_1, \ldots, w_m)^{\mathrm{T}}$ and σ^2.

When the number of basis functions is large, the maximum likelihood method yields unstable estimates (Konishi and Kitagawa 2008). Therefore, many researchers estimate $\boldsymbol{w}$ and σ^2 with a regularization method, such as the ridge method (Hoerl and Kennard 1970), the lasso method (Tibshirani 1996), or their variants (e.g., Zou and Hastie 2005; Fan and Li 2001).

However, when the underlying function $g(\cdot)$ has inhomogeneous smoothness, conventional regularization methods often lead to overfitting in the region where the function is strongly smooth or underfitting where it is weakly smooth (see Figure 1). To overcome this problem, Loader (1999) used the local likelihood method, which is specified as follows.

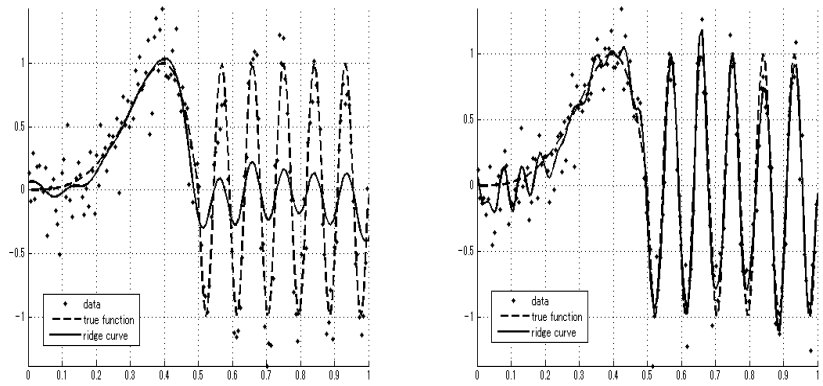

Figure 1: Estimated curves based on a conventional use of basis expansion with a strong regularization (left) and a weak regularization (right). Solid lines, broken lines, and dots depict the estimated curves, the true regression functions, and the data, respectively.

Let $k_h(x, x_0)$ be a weight function that assigns large weights to observations close to x:

$$k_h(x, x_0) = (2\pi h^2)^{-\frac{1}{2}} \exp\left\{-\frac{(x - x_0)^2}{2h^2}\right\}, \tag{3}$$

where h is a bandwidth parameter with positive value. The local likelihood method considers the weighted maximum likelihood at each point x_0 as

$$L(\boldsymbol{w}(x_0), \sigma^2(x_0)) = \sum_{i=1}^{n} k_h(x_i, x_0) \log f(y_i|x_i; \boldsymbol{w}(x_0), \sigma^2(x_0)). \tag{4}$$

Maximizing over the parameters $\boldsymbol{w}(x_0)$ and $\sigma^2(x_0)$ leads to the local likelihood estimate. This method performs well in simulations, but it bears intensive computational burdens. Miyata and Shen (2003) proposed an effective method for estimating these functions by using free-knot splines with variable multiple knots and an adaptive scheme for locating the optimal knots, but it is also computationally intensive.

2.2 Fused lasso signal approximator (FLSA)

In this section, we briefly describe the FLSA (Friedman *et al*. 2007), which we will extend in the next section. In the FLSA, there is one parameter per observation y_i, that

is, a model

$$y_i = \beta_i + \varepsilon_i, \quad (i = 1, \ldots, n) \tag{5}$$

is considered, where ε_i is an error with mean zero and variance σ^2, and β_i is the parameter to be estimated. The FLSA provides an estimator of $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_n)^{\mathrm{T}}$ as

$$\hat{\boldsymbol{\beta}} = \underset{\boldsymbol{\beta} \in \mathbb{R}^n}{\operatorname{argmin}} \left\{ \frac{1}{2} \sum_{i=1}^{n} (y_i - \beta_i)^2 + \lambda \sum_{i=2}^{n} |\beta_i - \beta_{i-1}| + \lambda' \sum_{i=1}^{n} |\beta_i| \right\}, \tag{6}$$

where λ and λ' are tuning parameters with positive values. The first penalty encourages sparsity in their differences, and the second encourages sparsity in the coefficients. As a result, $\hat{\boldsymbol{\beta}}$ becomes a step function.

It is difficult to obtain $\hat{\boldsymbol{\beta}}$ in (6) exactly when the sample size n is sufficiently large. Friedman *et al*. (2007) describe an algorithm to approximate $\hat{\boldsymbol{\beta}}$ quickly. The algorithm consists of three nested cycles: a descent cycle, a fusion cycle, and a smoothing cycle. For more details of the algorithm, see Friedman *et al.* (2007).

3 Proposed method

3.1 Extended FLSA

In this section, we consider an estimator of $\boldsymbol{\beta}$ which minimize

$$\frac{1}{2} \sum_{i=1}^{n} (y_i - \beta_i)^2 + \lambda \sum_{i=3}^{n} |\beta_i - 2\beta_{i-1} + \beta_{i-2}| + \lambda' \sum_{i=2}^{n} |\beta_i - \beta_{i-1}| + \lambda'' \sum_{i=1}^{n} |\beta_i|,$$

which leads to the extended FLSA (eFLSA), where λ (> 0), λ' (> 0), and λ'' (> 0) are tuning parameters. The first penalty term is new, is the absolute value of the second differences, and encourages their sparsity; the second and third penalty terms are the same as in (6). Our aim is to use the eFLSA to obtain a locally linear regression function that does not require that we collapse the data, i.e., we set $\lambda' = 0$ and $\lambda'' = 0$, and hence

consider

$$\hat{\boldsymbol{\beta}} = \underset{\boldsymbol{\beta}\in\mathbb{R}^n}{\operatorname{argmin}}\, h(\boldsymbol{\beta}) = \underset{\boldsymbol{\beta}\in\mathbb{R}^n}{\operatorname{argmin}} \left\{ \frac{1}{2}\sum_{i=1}^{n}(y_i-\beta_i)^2 + \lambda \sum_{i=3}^{n} |\beta_i - 2\beta_{i-1} + \beta_{i-2}| \right\}. \tag{7}$$

In order to solve (7), we propose an eFLSA algorithm that fuses $(\beta_{i+2} - \beta_{i+1})$ and $(\beta_{i+1} - \beta_i)$ for a number of i's. To describe this model, let us define the set $F = \{i_1, \ldots, i_p\}$ such that there is no j satisfying $i_j - i_{j-1} = 2$ and such that

$$\beta_{i_{j-1}+2} - \beta_{i_{j-1}+1} = \beta_{i_{j-1}+3} - \beta_{i_{j-1}+2} = \cdots = \beta_{i_j} - \beta_{i_j-1} \tag{8}$$

if $i_j - i_{j-1} \geq 3$ $(1 \leq j \leq p)$, where $i_0 = 0$ and $i_p = n$. We call F the fusion set. Note that β_i for $i_{j-1} < i \leq i_j$ can be expressed as $\beta_{i_j} - \gamma_j(i_j - i)$, where $\gamma_j = \beta_{i_j} - \beta_{i_j-1}$. The eFLSA algorithm updates F and estimates of $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_n)^{\mathrm{T}}$ through the following three nested cycles.

- Descent cycle: Update estimates of $\boldsymbol{\beta}$ for a given fusion set F by a coordinate wise descent with respect to each parameter β_{i_j} and γ_j.
- Fusion cycle: Consider the fusion of neighboring parameters to update the fusion set F, followed by the descent cycle.
- Smoothing cycle: Increase the tuning parameter λ by a small amount, and rerun the two previous cycles.

Note that the initial value of F is $\{1, \ldots, n\}$. We describe each cycle in more detail below.

3.1.1 Descent cycle

First, we consider the descent cycle. If $i_j - i_{j-1} = 1$, the derivative of (7) with respect to β_{i_j}, holding all $\beta_i = \tilde{\beta}_i$ $(i \neq i_j)$ fixed at their current estimates, is

$$\begin{aligned}
&\partial h(\boldsymbol{\beta})/\partial\beta_{i_j} \\
&= -(y_{i_j} - \beta_{i_j}) + \lambda \cdot \operatorname{sign}(\beta_{i_j} - 2\tilde{\beta}_{i_j-1} + \tilde{\beta}_{i_j-2}) \\
&\quad -2\lambda \cdot \operatorname{sign}(\tilde{\beta}_{i_j+1} - 2\beta_{i_j} + \tilde{\beta}_{i_j-1}) + \lambda \cdot \operatorname{sign}(\tilde{\beta}_{i_j+2} - 2\tilde{\beta}_{i_j+1} + \beta_{i_j}),
\end{aligned} \tag{9}$$

when $\beta_{i_j} \notin \{2\tilde{\beta}_{i_j-1} - \tilde{\beta}_{i_j-2},\ (\tilde{\beta}_{i_j+1} + \tilde{\beta}_{i_j-1})/2,\ 2\tilde{\beta}_{i_j+1} - \tilde{\beta}_{i_j+2}\}$. We have only to check for zero of (9). If no solution is found, we examine the three active-constraint values for β_i: $2\tilde{\beta}_{i-1} - \tilde{\beta}_{i-2}$, $(\tilde{\beta}_{i+1} + \tilde{\beta}_{i-1})/2$, and $2\tilde{\beta}_{i+1} - \tilde{\beta}_{i+2}$, and find the one that gives the smallest value of the objective function $h(\boldsymbol{\beta})$. On the other hand, if $i_j - i_{j-1} \geq 3$, the derivative of (7) with respect to β_{i_j}, holding the other parameters fixed at their current estimates, is

$$\begin{aligned}
&\partial h(\boldsymbol{\beta})/\partial \beta_{i_j} \\
&= -\sum_{i=i_{j-1}+1}^{i_j} \{y_i - \beta_{i_j} - \tilde{\gamma}_j \cdot (i_j - i)\} \\
&\quad +\lambda \cdot \text{sign}\{\beta_{i_j} - \tilde{\gamma}_j \cdot (i_j - i_{j-1} - 1) - 2\tilde{\beta}_{i_{j-1}} + \tilde{\beta}_{i_{j-1}-1}\} \\
&\quad -\lambda \cdot \text{sign}\{-\beta_{i_j} - \tilde{\gamma}_j \cdot (i_j - i_{j-1}) + \tilde{\beta}_{i_{j-1}}\} \\
&\quad -\lambda \cdot \text{sign}(\tilde{\beta}_{i_j+1} - \beta_{i_j} - \tilde{\gamma}_j) + \lambda \cdot \text{sign}(\tilde{\beta}_{i_j+2} - 2\tilde{\beta}_{i_j+1} + \beta_{i_j}),
\end{aligned} \tag{10}$$

and the derivative of (7) with respect to γ_j, holding all the other parameters fixed at their current estimates, is

$$\begin{aligned}
&\partial h(\boldsymbol{\beta})/\partial \gamma_j \\
&= -\sum_{i=i_{j-1}+1}^{i_j-1} (i_j - i)[y_i - \{\tilde{\beta}_{i_j} - \gamma_j \cdot (i_j - i)\}] \\
&\quad -\lambda \cdot (i_j - i_{j-1} - 1) \cdot \text{sign}\{\beta_{i_j} - \tilde{\gamma}_j \cdot (i_j - i_{j-1} - 1) - 2\tilde{\beta}_{i_{j-1}} + \tilde{\beta}_{i_{j-1}-1}\} \\
&\quad +\lambda \cdot (i_j - i_{j-1}) \cdot \text{sign}\{-\beta_{i_j} - \tilde{\gamma}_j \cdot (i_j - i_{j-1}) + \tilde{\beta}_{i_{j-1}}\} \\
&\quad -\lambda \cdot \text{sign}(\tilde{\beta}_{i_j+1} - \beta_{i_j} - \tilde{\gamma}_j).
\end{aligned} \tag{11}$$

We then check for zeros of (10) or (11), and if no solution is found, we examine some active-constraint values.

3.1.2 Fusion cycle

Friedman *et al.* (2007) reported that the solution of the descent cycle of the FLSA introduced in Section 2 often gets stuck, and the same problem occurs in our eFLSA. To

overcome this problem, we considered a fusion cycle similar to that of Friedman *et al.* (2007), and we checked if it would help to remove the index i_j from the fusion set F, that is, whether $\beta_{i_j+2} - \beta_{i_j+1} = \cdots = \beta_{i_{j+1}} - \beta_{i_{j+1}-1}$ and $\beta_{i_{j-1}+2} - \beta_{i_{j-1}+1} = \cdots = \beta_{i_j} - \beta_{i_j-1}$ should be fused. We ran the descent cycle for the fusion set $F \setminus \{i_j\}$ and computed the objective function. If it was smaller than the objective function for F, we updated the fusion set from F to $F \setminus \{i_j\}$. Note that if $i_{j+1} - i_j = i_j - i_{j-1} = 1$, $F \setminus \{i_j\}$ does not become a fusion set. Therefore, in this case, we considered the fusion set $F \setminus \{i_j, i_{j+1}\}$ in place of $F \setminus \{i_j\}$.

3.1.3 Smoothing cycle

Finally, we consider the smoothing cycle. We increased the value of the penalty λ by a small amount and reran the two previous cycles. Note that the new optimal fusion set will be a subset of the optimal fusion set before incrementation. Therefore, we use the fusion set before incrementation as the initial value for the rerun cycles, which reduces the number of computations. The smoothing cycle is then as follows:

1. Start with $\lambda = 0$, and run the descent cycle.

2. Increment $\lambda \leftarrow \lambda + \delta$, and run the descent and fusion cycles repeatedly until no further changes occur.

3. Repeat step 2 until a target value of λ is reached.

3.2 Nonlinear regression modeling via eFLSA

By applying our eFLSA algorithm, we obtain the $\hat{\boldsymbol{\beta}}$ that makes $h(\boldsymbol{\beta})$ in (7) small, from the final fusion set $F = \{i_1, \ldots, i_p\}$ and estimators $\hat{\beta}_{i_j}$ and $\hat{\gamma}_j$ $(1 \leq j \leq p)$. We can regard $\hat{\boldsymbol{\beta}}$ as the global variation, that is, we can catch the global variation with linear functions on the subintervals $[a_{j-1},\ a_j]$ $(1 \leq j \leq p)$, where $a_0 = x_1$, $a_j = (x_{i_j} + x_{i_j+1})/2$ $(1 \leq j \leq p-1)$,

and $a_p = x_n$. Here we measure the magnitude of the local variation at each subinterval $[a_{j-1}, a_j]$ by using the residuals of the data from the global variation that was caught by the eFLSA, and we propose to determine the basis functions adaptively, based on these magnitudes.

We use a Gaussian function that is one of the ones most commonly used in basis expansion,

$$\phi_k(x) = \exp\left\{-\frac{(x - c_k)^2}{2h_k^2}\right\}, \quad k = 1, \ldots, m, \tag{12}$$

where c_k and h_k are the center and bandwidth, respectively. Usually, the centers are set at regular intervals of the data region, and the bandwidths are set to be constant. Note that this procedure does not perform well if the underlying function has inhomogeneous smoothness.

Our idea is to determine the number and bandwidth of the basis functions appropriately according to the magnitude of the local variation. Here we use $d_j - d_{j-1}$ as a measure of the magnitude for the region $[a_{j-1}, a_j]$, where

$$d_j = \frac{\sum_{x_i \in [a_0, a_j]} (y_i - \hat{\beta}_i)^2}{\sum_{x_i \in [a_0, a_p]} (y_i - \hat{\beta}_i)^2}, \quad j = 1, \ldots, p.$$

Then we provide the number of basis functions m_j on each subinterval $[a_{j-1}, a_j]$ as follows:

$$m_j = [m \times d_j]_{\text{round}} - [m \times d_{j-1}]_{\text{round}}, \quad j = 1, \ldots, p,$$

where $[\,\cdot\,]_{\text{round}}$ is the round-off function. Based on m_j, we determine the center $\hat{c}_{j,\ell}$ and the bandwidth parameter $\hat{h}_j$ for each subinterval $[a_{j-1}, a_j]$, as follows:

$$\hat{c}_{j,\ell} = a_{j-1} + \left(\ell - \frac{1}{2}\right)\hat{h}_j, \quad \ell = 1, \ldots, m_j; \quad j = 1, \ldots, p,$$

and

$$\hat{h}_j = \frac{a_j - a_{j-1}}{m_j}, \quad j = 1, \ldots, p.$$

Thus, we can provide a large number of Gaussian basis functions with a small bandwidth on the region where the local variation is large, and a small number with a large bandwidth where the local variation is small. Replacing c_k and h_k^2 in (12) by $\hat{c}_{j,\ell}$ and $\hat{h}_j^2$, respectively, we obtain m basis functions

$$\phi_{j,\ell}(x) = \exp\left\{-\frac{(x-\hat{c}_{j,\ell})^2}{2\hat{h}_j^2}\right\}, \quad \ell = 1,\ldots,m_j; \quad j = 1,\ldots,p. \tag{13}$$

Finally, we estimate the model parameters $\boldsymbol{w}$ and σ^2 by the ridge method. We use the ridge penalty on the log-likelihood function

$$\ell_\eta(\boldsymbol{w},\sigma^2) = \sum_{i=1}^{n} \log f(y_i|x_i;\boldsymbol{w},\sigma^2) - \frac{n\eta}{2}\boldsymbol{w}^{\mathrm{T}}\boldsymbol{w}, \tag{14}$$

where η (> 0) is a tuning parameter that controls the smoothness of the fitted model. The maximum penalized likelihood estimates of $\boldsymbol{w}$ and σ^2 are, respectively,

$$\hat{\boldsymbol{w}} = (\Phi^{\mathrm{T}}\Phi + n\eta\hat{\sigma}^2 I)^{-1}\Phi^{\mathrm{T}}\boldsymbol{y}, \quad \hat{\sigma}^2 = \frac{1}{n}(\boldsymbol{y}-\Phi\hat{\boldsymbol{w}})^{\mathrm{T}}(\boldsymbol{y}-\Phi\hat{\boldsymbol{w}}).$$

Here, Φ is an $n \times m$ matrix $(\Phi_1, \Phi_2, \ldots, \Phi_p)$, where

$$\Phi_j = \begin{bmatrix} \phi_{j,1}(x_1) & \cdots & \phi_{j,m_j}(x_1) \\ \vdots & \ddots & \vdots \\ \phi_{j,1}(x_n) & \cdots & \phi_{j,m_j}(x_n) \end{bmatrix}; \quad j = 1,\ldots,p. \tag{15}$$

Note that $\hat{\boldsymbol{w}}$ and $\hat{\sigma}^2$ depend on each other. Therefore, we provide an appropriate initial value for σ^2 and then $\hat{\boldsymbol{w}}$ and $\hat{\sigma}^2$ are updated until convergence.

The choice of the tuning parameter η and number of basis functions m are crucial issues. To determine them, we use the generalized information criterion (GIC; Konishi and Kitagawa, 1996), which is the Akaike information criterion (AIC) generalized for the regularization method. For our problem, the GIC can be obtained as follows:

$$\mathrm{GIC} = n\left\{\log(2\pi)+1\right\} + n\log\hat{\sigma}^2 + 2\mathrm{tr}\left\{R^{-1}Q\right\}. \tag{16}$$

Here, R and Q are $(m+1)\times(m+1)$ matrices given by

$$R=\frac{1}{n\hat{\sigma}^2}\begin{bmatrix}\Phi^{\mathrm{T}}\Phi+n\eta\hat{\sigma}^2\mathrm{I}_m & \Phi^{\mathrm{T}}\Lambda\mathbf{1}_n/\hat{\sigma}^2\\ \mathbf{1}_n^{\mathrm{T}}\Lambda\Phi/\hat{\sigma}^2 & n/(2\hat{\sigma}^2)\end{bmatrix}$$

and

$$Q=\frac{1}{n\hat{\sigma}^2}\begin{bmatrix}\Phi^{\mathrm{T}}\Lambda^2\Phi/\hat{\sigma}^2-\eta\mathrm{I}_m\hat{\boldsymbol{w}}\mathbf{1}_n^{\mathrm{T}}\Lambda\Phi & \Phi^{\mathrm{T}}\Lambda^3\mathbf{1}_n/(2\hat{\sigma}^4)-\Phi^{\mathrm{T}}\Lambda\mathbf{1}_n/(2\hat{\sigma}^2)\\ \mathbf{1}_n^{\mathrm{T}}\Lambda^3\Phi/(2\hat{\sigma}^4)-\mathbf{1}_n^{\mathrm{T}}\Lambda\Phi/(2\hat{\sigma}^2) & \mathbf{1}_n^{\mathrm{T}}\Lambda^4\mathbf{1}_n/(4\hat{\sigma}^6)-n/(4\hat{\sigma}^2)\end{bmatrix},$$

where $\mathbf{1}_n=(1,\ldots,1)^{\mathrm{T}}$ and $\Lambda=\mathrm{diag}[y_1-\hat{\boldsymbol{w}}^{\mathrm{T}}\boldsymbol{\phi}(x_1),\ldots,y_n-\hat{\boldsymbol{w}}^{\mathrm{T}}\boldsymbol{\phi}(x_n)]$. We use the values of η and m that minimize the GIC in (16), since by choosing to use the GIC, they are considered optimal.

4 Numerical studies

4.1 Monte Carlo simulations

To investigate the performance of our proposed adaptive basis expansion, we applied it with simple regression functions to several types of simulated data. We generated samples $\{(x_i,y_i);i=1,\ldots,n\}$ from $y_i=g(x_i)+\varepsilon_i$ with a regression function $g(x)$ and noise ε_i. We consider the following two cases for the regression function:

$$g(x)=\frac{5}{\sqrt{2\pi}}\exp\left(\frac{x^2}{2}\right),\qquad(-10\le x\le 10),\tag{17}$$

and

$$g(x)=\begin{cases}\sin\left\{8\pi\exp(x^3)\right\}, & (0\le x<0.5),\\ \sin(22\pi x), & (0.5\le x<1).\end{cases}\tag{18}$$

Table 1: Mean, median, and standard deviation of the MSEs for the regression function in (17).

		proposed	non-adaptive
		Mean [Median] (SD)	Mean [Median] (SD)
	$n = 50$	0.0057 [0.0047] (0.0047)	0.0064 [0.0061] (0.0016)
$\tau = 0.1$	$n = 100$	0.0043 [0.0030] (0.0070)	0.0120 [0.0046] (0.0420)
	$n = 150$	0.0042 [0.0021] (0.0207)	0.0307 [0.0037] (0.0768)
	$n = 50$	0.0160 [0.0157] (0.0063)	0.0238 [0.0235] (0.0058)
$\tau = 0.2$	$n = 100$	0.0158 [0.0106] (0.0208)	0.0170 [0.0142] (0.0232)
	$n = 150$	0.0080 [0.0074] (0.0040)	0.0230 [0.0119] (0.0485)
	$n = 50$	0.0388 [0.0317] (0.0427)	0.0463 [0.0446] (0.0118)
$\tau = 0.3$	$n = 100$	0.0218 [0.0200] (0.0135)	0.0297 [0.0275] (0.0109)
	$n = 150$	0.0169 [0.0146] (0.0133)	0.0263 [0.0225] (0.0124)

Here we assume that the design points $\{x_i\}_{i=1}^n$ are uniformly spaced on $[0, 1]$ and the errors ε_i are independently distributed according to $\mathrm{N}(0, \tau^2)$. The sample size and standard deviation are, respectively, $n = 50,\ 100$, or 150 and $\tau = 0.1$, 0.2, or 0.3 for the case in (17), and they are $n = 100,\ 200$, or 300 and $\tau = 0.2$, 0.4, or 0.6 for the case in (18).

Simulation results were obtained from 100 Monte Carlo trials, and then we evaluated the mean squared errors (MSE), defined by $\sum_{i=1}^n \{g(x_i) - \hat{y}_i\}^2 / n$, to measure the goodness of fit. The proposed method was compared with a conventional Gaussian basis expansion with the ridge method, where the basis functions were represented by

$$\psi_k(x) = \exp\left\{-\frac{(x - b_k)^2}{2h}\right\}, \quad k = 1, \ldots, m. \tag{19}$$

Table 2: Mean and standard deviation of the MSEs for the regression function in (18) (Region 1: $0 < x < 0.5$, Region 2: $0.5 < x < 1$).

			Region 1	Region 2	Total
			mean (SD)	mean (SD)	mean (SD)
$\tau = 0.2$	$n = 100$	proposed	0.0039 (0.0019)	0.0085 (0.0151)	0.0063 (0.0110)
		non-adaptive	0.0179 (0.0038)	0.0185 (0.0035)	0.0182 (0.0036)
	$n = 200$	proposed	0.0023 (0.0015)	0.0063 (0.0192)	0.0043 (0.0138)
		non-adaptive	0.0060 (0.0018)	0.0063 (0.0018)	0.0061 (0.0018)
	$n = 300$	proposed	0.0014 (0.0007)	0.0048 (0.0229)	0.0031 (0.0163)
		non-adaptive	0.0050 (0.0024)	0.0053 (0.0024)	0.0051 (0.0024)
$\tau = 0.4$	$n = 100$	proposed	0.0183 (0.0083)	0.0250 (0.0090)	0.0216 (0.0093)
		non-adaptive	0.0773 (0.0153)	0.0767 (0.0125)	0.0770 (0.0140)
	$n = 200$	proposed	0.0097 (0.0040)	0.0131 (0.0052)	0.0114 (0.0049)
		non-adaptive	0.0261 (0.0068)	0.0255 (0.0075)	0.0258 (0.0071)
	$n = 300$	proposed	0.0065 (0.0027)	0.0084 (0.0027)	0.0074 (0.0028)
		non-adaptive	0.0166 (0.0047)	0.0161 (0.0035)	0.0163 (0.0042)
$\tau = 0.6$	$n = 100$	proposed	0.0442 (0.0201)	0.0499 (0.0199)	0.0471 (0.0201)
		non-adaptive	0.1030 (0.0246)	0.1022 (0.0267)	0.1026 (0.0475)
	$n = 200$	proposed	0.0253 (0.0098)	0.0266 (0.0109)	0.0259 (0.0104)
		non-adaptive	0.0564 (0.0158)	0.0550 (0.0163)	0.0557 (0.0160)
	$n = 300$	proposed	0.0144 (0.0060)	0.0176 (0.0059)	0.0160 (0.0061)
		non-adaptive	0.0358 (0.0102)	0.0348 (0.0084)	0.0353 (0.0093)

In the conventional approach, the centers of the Gaussian basis functions b_k $(k = 1, \ldots, m)$ were set at regular intervals on the data region, and the bandwidths h were set to be constant. Hereinafter, we will call this the non-adaptive method. We chose the number of basis functions m, the tuning parameter η, and the bandwidth parameters h that minimize the information criterion GIC in (16). The candidate values were specified as $m = 100 + 25q$ $(q = 0, \ldots, 8)$, $\eta = 10^{-9} + 3q \cdot 10^{-6}$ $(q = 0, \ldots, 5)$, and $h = 10^{-3} + 3q \cdot 10^{-2}$ $(q = 0, \ldots, 5)$.

Table 1 evaluates the MSE for the case of (17), where both methods are equipped with the GIC. In all cases, the proposed method gives smaller MSEs than does the non-adaptive method. Because the non-adaptive method sometimes yields inaccurate results, several values for the mean of the MSEs for the non-adaptive method are unnaturally large. We thus also evaluate the median of the MSEs. From the values for the median, we conclude that the proposed method produces the same degree of improvement regardless of the values of n and τ.

Table 2 presents the MSE for the case of (18) by Region 1 $(0 < x < 0.5)$, Region 2 $(0.5 < x < 1)$, and the entire region $(0 < x < 1)$. We observe that the mean of the MSEs of the proposed method is smaller than that of the non-adaptive method in both Regions 1 and 2. In all simulations, the improvement of the proposed method in Region 1 (a strongly smooth region) tends to be larger than that in Region 2 (a weakly smooth region). This means that the proposed method avoids overfitting more than underfitting. When τ is small (i.e., $\tau = 0.2$) and n is relatively large (i.e., $n = 200$ or $n = 300$), the performances of the two methods are competitive. This is not surprising, because in this case the standard deviation of the noise is small enough and the sample size is large enough that the non-adaptive method will work well. When $\tau = 0.4$ and $\tau = 0.6$, the proposed method produces the same degree of improvement. We can thus see the validity of the proposed method in all cases, but especially in the strongly smooth regions.

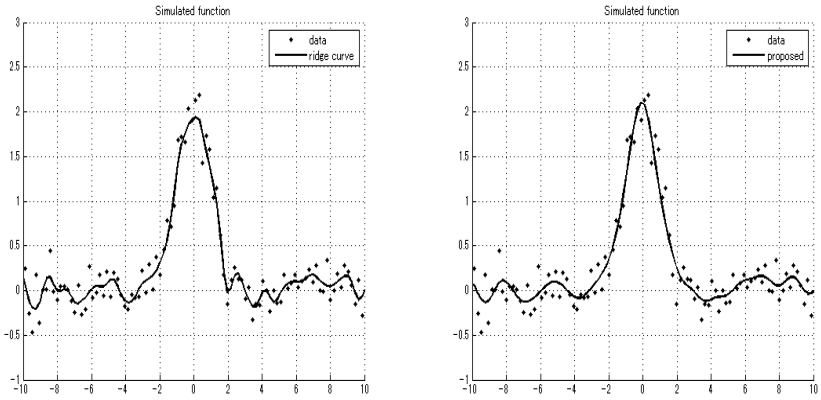

Figure 2: Estimated curves based on a conventional basis expansion (left) and the proposed adaptive basis expansion (right) for the regression function in (17).

Figures 2 and 3 give estimated curves for each method for the case of $\tau = 0.2$ and $n = 100$ and for the case of $\tau = 0.4$ and $n = 200$, respectively. These curves are both typical, and we can analyze them as above. From Figures 2 and 3, respectively, we can see that the proposed method avoids overfitting on $|x| \geq 2$ and on Region 1, and it avoids underfitting around $x = 0$ and on Region 2. From Figure 3, we see that its improvement on Region 1 is a little bit larger than that on Region 2.

4.2 Benchmark example

We next treat the Doppler data with sample size $n = 128$ presented by Donoho and Johnstone (1995). Doppler data have an underlying function that vibrates with equal amplitude but with a shorter period near zero, and so it has inhomogeneous smoothness. We applied both the proposed and the non-adaptive methods to this data. The candidate values were specified as $m = 100 + 50q$ $(q = 0, \ldots, 4)$, $\lambda = 10^{-9} + 3q \cdot 10^{-6}$ $(q = 0, \ldots, 3)$, and $h = 10^{-2} + 3q \cdot 10^{-1}$ $(q = 0, \ldots, 3)$, and we selected appropriate values from these candidates by applying the GIC. Figure 4 displays the estimated curves based on each

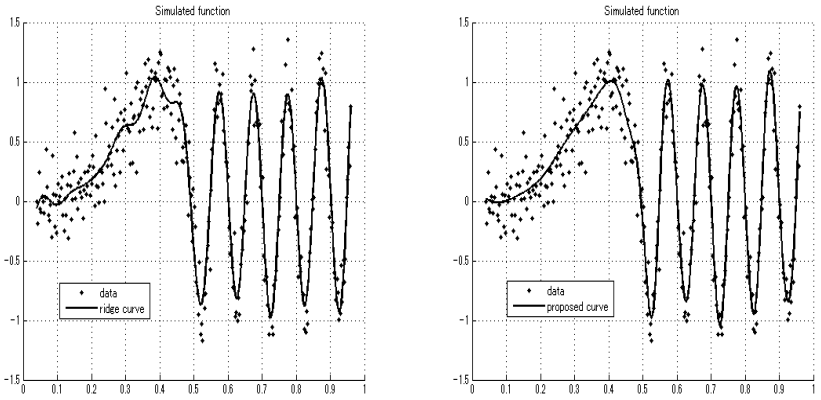

Figure 3: Estimated curves based on a conventional basis expansion (left) and the proposed adaptive basis expansion (right) for the regression function in (18).

method. The function estimated by the proposed method vibrated near zero, whereas the one estimated by the non-adaptive method did not. Hence we can say that the proposed method caught the underlying regression function in the Doppler data.

5 Concluding Remarks

In this study, an adaptive basis expansion has been proposed for estimating underlying regression functions with inhomogeneous smoothness. In the first stage, our method tries to catch the global variation of data using a locally linear function. To do so, we extended the FLSA and constructed an extended version of the descent, fusion, and smoothing cycles in order to obtain the approximator quickly and efficiently. We call this the eFLSA algorithm. Second, our method measures the magnitude of the residuals between the global variation and the data by regions, and the basis functions are determined adaptively based on these magnitudes. In particular, a large number of Gaussian basis functions with small bandwidth are set on regions with large residuals, and a smaller number with larger bandwidth are set on regions with small residuals. Finally, our method uses a ridge method

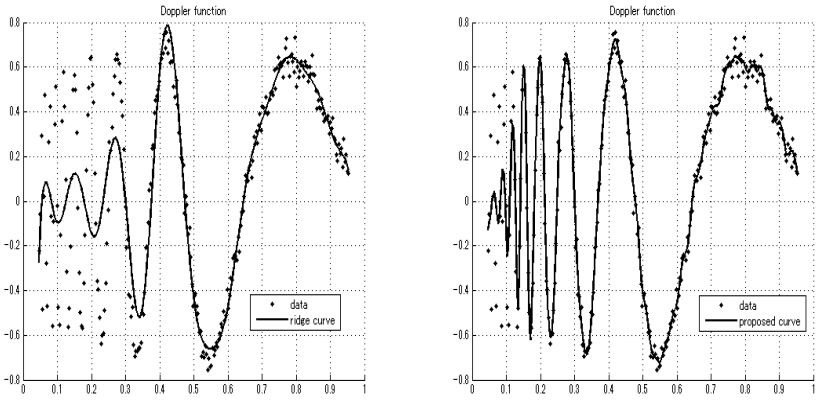

Figure 4: Estimated curves based on a conventional basis expansion (left) and the proposed adaptive basis expansion (right) for Doppler data

to estimate the parameters for the models. To choose the number of basis functions and the tuning parameters for the ridge method, we have provided an explicit form of the GIC. Numerical examples suggest that when there is inhomogeneous smoothness, our method can capture the true structure better than a conventional basis expansion, in terms of the averaged mean squared errors. The conventional method causes over- and underfitting on strongly and weakly smooth regions, respectively, but our proposed methodology considerably reduces this problem.

References

[1] de Boor, C. (2001). *A Practical Guide to Splines. Revised Edition.* Springer.

[2] Donoho, D. L. and Johnstone, J. M. (1994). Ideal spatial adaptation by wavelet shrinkage. *Biometrika*, **81**, 425–455.

[3] Fan, J. and Li, R. (2001). Variable selection via nonconcave penalized likelihood and its oracle properties. *Journal of the American Statistical Association*, **96**, 1348–1360.

[4] Friedman, J., Hastie, T., Höfling, H. and Tibshirani, R. (2007). Pathwise coordinate optimization. *The Annals of Applied Statistics*, **1**, 302–332.

[5] Green, P. J. and Silverman, B. W. (1994). *Nonparametric Regression and Generalized Linear Models.* Chapman & Hall.

[6] Hastie, T., Tibshirani, R. and Friedman, J. (2009). *The Elements of Statistical Learning.* Springer.

[7] Hoerl, A. E. and Kennard, R. W. (1970). Ridge regression: biased estimation for nonorthogonal problems. *Technometrics*, **1**, 55–67.

[8] Konishi, S. and Kitagawa, G. (1996). Generalised information criteria in model selection. *Biometrika*, **83**, 875–890.

[9] Konishi, S. and Kitagawa, G. (2008). *Information Criteria and Statistical Modeling.* Springer.

[10] Miyata, S. and Shen, X. (2003). Adaptive free-knot splines. *Journal of Computational and Graphical Statistics*, **12**, 197–213.

[11] Loader, C. (1999). *Local Regression and Likelihood.* Springer.

[12] Stone, C. J. (1974). Cross-validatory choice and assessment of statistical predictions (with discussion). *Journal of Royal Statistical Society Series B*, **64**, 583–639.

[13] Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society Series B*, **58**, 267–288.

[14] Zou, H. and Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society Series B*, **67**, 301–320.

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

MI2012-3 Hiroshi INOUE
A Weak RIP of theory of compressed sensing and LASSO

MI2012-4 Yasuhide FUKUMOTO & Youich MIE
Hamiltonian bifurcation theory for a rotating flow subject to elliptic straining field

MI2012-5 Yu KAWAKAMI
On the maximal number of exceptional values of Gauss maps for various classes of surfaces

MI2012-6 Marcio GAMEIRO, Yasuaki HIRAOKA, Shunsuke IZUMI, Miroslav KRAMAR, Konstantin MISCHAIKOW & Vidit NANDA
Topological Measurement of Protein Compressibility via Persistence Diagrams

MI2012-7 Nobutaka NAKAZONO & Seiji NISHIOKA
Solutions to a q-analog of Painlevé III equation of type $D_7^{(1)}$

MI2012-8 Naoyuki KAMIYAMA
A new approach to the Pareto stable matching problem

MI2012-9 Jan BREZINA & Yoshiyuki KAGEI
Spectral properties of the linearized compressible Navier-Stokes equation around time-periodic parallel flow

MI2012-10 Jan BREZINA
Asymptotic behavior of solutions to the compressible Navier-Stokes equation around a time-periodic parallel flow

MI2012-11 Daeju KIM, Shuichi KAWANO & Yoshiyuki NINOMIYA
Adaptive basis expansion via the extended fused lasso