Variable selection for varying coefficient models with the sparse regularization

Matsui, Hidetoshi Faculty of Mathematics, Kyushu University

Misumi, Toshihiro Astellas Pharma Inc. | Graduate School of Science and Engineering, Chuo University

https://hdl.handle.net/2324/26215

出版情報:Computational Statistics. 29, 2013-03-28. 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

Variable selection for varying coefficient models with the sparse regularization

Hidetoshi Matsui

MI 2013-8

(Received March 28, 2013)

Faculty of Mathematics Kyushu University Fukuoka, JAPAN

Variable selection for varying coefficient models with the sparse regularization

Hidetoshi Matsui^a Toshihiro Misumi^{b,c}

 ^a Faculty of Mathematics, Kyushu University, 744 Motooka, Nishi-ku, Fukuoka 819-0395, Japan,
 ^b Astellas Pharma Inc., 3-17-1, Hasune, Itabashi-ku, Tokyo 174-8612, Japan,
 ^c Graduate School of Science and Engineering, Chuo University, 1-13-27, Kasuga, Bunkyo-ku, Tokyo 112-8551, Japan.

hmatsui@math.kyushu-u.ac.jp (H. Matsui)

Abstract

Varying-coefficient models are useful tools for analyzing longitudinal data. They can effectively describe a relationship between predictors and responses repeatedly measured. We consider the problem of selecting variables in the varying-coefficient models via the adaptive elastic net regularization. Coefficients given as functions are expressed by basis expansions, and then parameters involved in the model are estimated by the penalized likelihood method using the coordinate descent algorithm derived for solving the problem of sparse regularization. We examine the effectiveness of our modeling procedure through Monte Carlo simulations and real data analysis.

Key Words and Phrases: Basis expansion, Elastic net, Group lasso, Varyingcoefficient model.

1 Introduction

Longitudinal data analysis has been used in various fields such as bioscience, ergonomics and meteorology. In longitudinal studies the data are measured repeatedly over time for individual and they have possibly different time points, which makes it difficult to directly apply the traditional multivariate analysis. Enormous works have contributed to the development of longitudinal data analysis during the past few decades (see, e.g. Diggle, 2002). In this article we focus on the problem of selecting variables which seem to have a relation with a response which is repeatedly measured.

Several methods have been proposed for the analysis of longitudinal data. Laird and Ware (1982) introduced random effects models and their estimation procedures and Zeger and Diggle (1994) applied semiparametric models as the regression models for longitudinal data. Ramsay and Silverman (2005) considered treating longitudinal data as smooth functions, and then established the Functional Data Analysis (FDA). On the other hand, Hoover et al. (1998) generalized the semiparametric model and applied the time varying-coefficient model. It is a special case of the varying-coefficient models (Hastie and Tibshirani, 1993) and obtained for analyzing longitudinal data. They approached the modeling by use of kernel smoothing or smoothing splines with an L_2 -type penalty. Huang et al. (2002) also used smoothing splines for time varying-coefficient models and suggested a non-iterative solution for coefficient estimates. The effectiveness of varying coefficient models are reported in several works (Fan and Zhang, 1999; Cai et al., 2000; Eubank et al., 2004).

More recently, a class of sparse regularization has come to be used for the varyingcoefficient modeling. Sparse regularization including the lasso (Tibshirani, 1996), SCAD (Fan and Li, 2001) and the elastic net (Zou and Hastie, 2005) simultaneously shrinks parameters and selects variables which seem to be relevant to a response by estimating coefficients of rest of the variables to be exactly zeros, and therefore it is one of the most attractive methods in recent years. More details of sparse regularization can be seen in Hastie et al. (2009). We can apply the sparse regularization to the varying-coefficient models by expressing coefficient functions by basis expansions and then imposing L_1 type penalties to coefficient vectors. Since each variable has multiple parameters in this case, Wang et al. (2007b) treated this problem as the grouped variable selection and then applied the idea of the group lasso by Yuan and Lin (2006). They also selected the transcriptional factors involved in gene regulation during a biological process. Wei et al. (2011) applied the adaptive lasso penalty (Zou, 2006) and then estimated the model by the coordinate descent algorithm (Friedman et al., 2007) which is derived for the sparse regularization problem. Many other works concerning the varying-coefficient modeling with the sparse regularization have been reported (Wang et al., 2008; Noh and Park, 2010; Xue and Qu, 2012; Wang et al., 2013).

In this paper we propose a method for estimating and selecting models simultaneously by the varying-coefficient modeling along with the elastic net regularization by extending the method by Wei et al. (2011). Especially we use an adaptive elastic net penalty proposed by Zou and Zhang (2009). They combined the idea of the elastic net and the adaptive lasso in the framework of the general linear model. The adaptive elastic net regularization can select variables even if the number of variables is much larger than the sample size and can take into account highly correlated predictors. Furthermore its estimates has an preferable property called the "oracle property." By applying the elastic net regularization to the varying-coefficient model we can prevent its estimation from giving unstable or degenerate results owing to both L_1 and L_2 terms of the penalty. Coefficient functions of varying-coefficient models are expressed by basis expansions, and then parameters are estimated by the penalized maximum likelihood method with the help of the coordinate descent algorithm. In order to select tuning parameters involved in the adaptive elastic net penalty we use a Bayesian model selection criterion derived for evaluating the varying-coefficient model. The proposed modeling strategy is applied to Monte Carlo simulations and clinical investigation data to investigate the effectiveness of our method.

This paper is organized as follows. In Section 2 we introduce a varying-coefficient model which describe the relationship between multiple predictors and a response repeatedly measured. Section 3 provides a method for estimating the varying-coefficient model with the adaptive elastic net regularization and selecting tuning parameters. Monte Carlo simulations are conducted in Section 4 in order to evaluate the effectiveness of the proposed modeling strategy and real data analysis are described in Section 5. Finally concluding remarks are given in Section 6.

2 Varying-coefficient models

Suppose we have p sets of predictors X_k (k = 1, ..., p) and a response Y, each of them are repeatedly observed at possibly different time points. Denote *i*-th of n individuals at *j*-th of n_i time points of X_k and Y as x_{ijk} and y_{ij} respectively. The varying-coefficient model (Hoover et al., 1998) that represents the relationship between X_k s and Y is given by

$$y_{ij} = \beta_0(t_{ij}) + x_{ij1}\beta_1(t_{ij}) + \dots + x_{ijp}\beta_p(t_{ij}) + \varepsilon_{ij}, \qquad (1)$$

where $\beta_0(\cdot)$ is an intercept and $\beta_1(\cdot), \ldots, \beta_p(\cdot)$ are coefficients, both of which are given as functions to be estimated. Moreover, ε_{ij} are random noises whose vector $\boldsymbol{\varepsilon}_i = (\varepsilon_{i1}, \ldots, \varepsilon_{in_i})'$ are normally distributed with mean vector **0** and variance covariance matrix $\sigma^2 S_i$ with unknown scalar σ^2 and known $n_i \times n_i$ symmetric matrix S_i .

We assume that coefficient functions $\beta_k(\cdot)$ (k = 0, ..., p) are expressed by basis expansions as follows;

$$\beta_k(t_{ij}) = \sum_{m=1}^{M_k} \gamma_{km} \phi_m^{(k)}(t_{ij}) = \boldsymbol{\gamma}'_k \boldsymbol{\phi}^{(k)}(t_{ij}),$$

where $\boldsymbol{\gamma}_k = (\gamma_{k1}, \ldots, \gamma_{kM_k})'$ are coefficient vectors and $\boldsymbol{\phi}^{(k)}(t_{ij}) = (\phi_1^{(k)}(t_{ij}), \ldots, \phi_{M_k}^{(k)}(t_{ij}))'$ are basis functions. There are several basis functions available for $\boldsymbol{\phi}^{(k)}(t)$ such as wavelets or radial basis functions. Here we apply cubic *B*-spline bases, whose details are referred to de Boor (2001); Imoto and Konishi (2003). Using this assumption and denoting $\boldsymbol{y}_i = (y_{i1}, \ldots, y_{in_i})'$, $D_{ik} = \text{diag}(x_{i1k}, \ldots, x_{in_ik})$ $(k = 1, \ldots, p)$, $D_{i0} = I_{n_i}$ and $\Phi_{ik} = (\boldsymbol{\phi}^{(k)}(t_{i1}), \ldots, \boldsymbol{\phi}^{(k)}(t_{in_i}))'$, the varying-coefficient model (1) can be rewritten as

$$\boldsymbol{y}_i = \sum_{k=0}^p D_{ik} \Phi_{ik} \boldsymbol{\gamma}_k + \boldsymbol{\varepsilon}_i, \qquad \boldsymbol{\varepsilon}_i \sim N_{n_i}(\boldsymbol{0}, \sigma^2 S_i)$$

Then we have a probability density function

$$f(\boldsymbol{y}_{i}|\boldsymbol{\theta}) = \frac{1}{(2\pi\sigma^{2})^{n_{i}/2}|S_{i}|^{1/2}} \exp\left\{-\frac{1}{2\sigma^{2}}\left(\boldsymbol{y}_{i} - \sum_{k=0}^{p} D_{ik}\Phi_{ik}\boldsymbol{\gamma}_{k}\right)'S_{i}^{-1}\left(\boldsymbol{y}_{i} - \sum_{k=0}^{p} D_{ik}\Phi_{ik}\boldsymbol{\gamma}_{k}\right)\right\},$$
(2)

where we denote a vector of unknown parameters by $\boldsymbol{\theta} = \{\boldsymbol{\gamma}_0, \dots, \boldsymbol{\gamma}_p, \sigma^2\}$. After obtaining an estimator of $\boldsymbol{\theta}$, denoted by $\hat{\boldsymbol{\theta}}$, by the method described in the next section, we have a statistical model $f(\boldsymbol{y}_i|\hat{\boldsymbol{\theta}})$.

3 Estimation and evaluation via the sparse regularization

Unknown parameters; coefficient vectors $\boldsymbol{\gamma}_k$ and variance parameter σ^2 , involved in the varying coefficient model are estimated by the maximum penalized likelihood method which maximizes a penalized log-likelihood function. The penalized log-likelihood function is given in the form of

$$l_{\lambda}(\boldsymbol{\theta}) = \sum_{i=1}^{n} \log f(\boldsymbol{y}_{i}|\boldsymbol{\theta}) - n\lambda \sum_{k=1}^{p} P_{\alpha}(\|\boldsymbol{\gamma}_{k}\|),$$

where $P_{\alpha}(\cdot)$ is a penalty function, $\|\cdot\|$ is a standard L_2 norm and $\lambda > 0$ is a regularization parameter which controls the degree of the penalty. We impose the penalty composed by a sum of penalty functions of L_2 norms of the coefficient vectors $\boldsymbol{\gamma}_k$ instead of its components separately. Then we can shrink all elements of the vector $\boldsymbol{\gamma}_k$ towards exactly zeros when the corresponding predictor seems to be less relevant to the response. It is exactly the idea of the group lasso by Yuan and Lin (2006). For the penalty function P_{α} we use an adaptive elastic net penalty (Zou and Zhang, 2009) given by

$$P_{\alpha}(\theta) = \frac{1}{2}(1-\alpha)\theta^2 + \alpha \hat{w}_k |\theta|$$
(3)

for all $\theta \in \mathbb{R}$, where $\alpha \in [0, 1]$ tunes the type of the penalty between the ridge ($\alpha = 0$) and the lasso ($\alpha = 1$). Furthermore, $\hat{w}_k > 0$ is an adaptive weight which is given in the form of $\hat{w}_k = (\sqrt{M_k} \| \boldsymbol{\gamma}_k \|)^{-\beta} (\| \boldsymbol{\gamma}_k \| \neq 0), = \infty (\| \boldsymbol{\gamma}_k \| = 0)$ with a positive constant β , where we use $\beta = 1$ in this paper.

We have difficulty in analytical expressions for the estimates of coefficients of regression models via the L_1 -type regularization. Wang et al. (2007b) used a local quadratic approximation (Fan and Li, 2001) for the varying-coefficient modeling via the group SCAD regularization, and Wei et al. (2011) applied a coordinate descent procedure (Friedman et al., 2007) to the group adaptive lasso regularization. The coordinate descent method has similar algorithm to the backfitting algorithm, used in the varying coefficient modeling (Hastie and Tibshirani, 1993), in that it updates coefficients of each variable in turn. Using this method, we obtain updated values of k-th coefficients (k = 1, ..., p) as

$$\hat{\boldsymbol{\gamma}}_{k} = \left(\sum_{i=1}^{n} \Phi_{ik}^{\prime} D_{ik} S_{i}^{-1} D_{ik} \Phi_{ik} + n(1-\alpha)\lambda \hat{\sigma}^{2} I\right)^{-1} \left(\|\boldsymbol{\zeta}_{k}\| - n\alpha\lambda \hat{w}_{k} \hat{\sigma}^{2}\right)_{+} \frac{\boldsymbol{\zeta}_{k}}{\|\boldsymbol{\zeta}_{k}\|}, \quad (4)$$

where

$$\boldsymbol{\zeta}_{k} = \sum_{i=1}^{n} \Phi_{ik}^{\prime} D_{ik} S_{i}^{-1} \left(\boldsymbol{y}_{i} - \sum_{l \neq k}^{p} D_{il} \Phi_{il} \hat{\boldsymbol{\gamma}}_{l} \right)$$

and $z_{+} = \max\{z, 0\}$ for $z \in \mathbb{R}$. Both of two terms of the adaptive elastic net penalty (3) directly have effects to the estimates (4) of the varying-coefficient model. The L_1 term of the penalty shrinks some of coefficients γ_k towards zero vectors, while the L_2 term prevents the degeneracy of the inverse, especially when there are few time points for individual and a large number of basis functions. Although the first term in the inverse matrix of (4) does not always degenerate even if n < p unlike the ordinal linear model, the L_2 term of the elastic net penalty provides stable estimates. Note that when k = 0there is no regularization and therefore $\hat{\gamma}_0$ has the form of (4) with a generalized inverse and $\lambda = 0$. The variance parameter is then estimated in the following form:

$$\hat{\sigma}^2 = \frac{1}{\sum_i n_i} \sum_{i=1}^n \left(\boldsymbol{y}_i - \sum_{k=0}^p D_{ik} \Phi_{ik} \hat{\boldsymbol{\gamma}}_k \right)' S_i^{-1} \left(\boldsymbol{y}_i - \sum_{k=0}^p D_{ik} \Phi_{ik} \hat{\boldsymbol{\gamma}}_k \right).$$
(5)

Since $\hat{\gamma}_k$ and $\hat{\sigma}^2$ depend on each other, they are updated until a convergence condition is achieved. Consequently, the algorithm is given as follows:

- 1. Initialize parameters $\gamma_0, \gamma_1, \ldots, \gamma_p$ and σ^2 by 1 for all elements.
- 2. For k = 0, 1..., p, update γ_k by (4) in turn.
- 3. Update σ^2 by (5).
- 4. Iterate 2 and 3 until convergence.

Thus the statistical model is obtained by substituting the estimators $\hat{\boldsymbol{\theta}} = \{\hat{\boldsymbol{\gamma}}_0, \dots, \hat{\boldsymbol{\gamma}}_p, \hat{\sigma}^2\}$ given above into the probability density function (2), that is,

$$f(\boldsymbol{y}_{i}|\hat{\boldsymbol{\theta}}) = \frac{1}{(2\pi\hat{\sigma}^{2})^{n_{i}/2}|S_{i}|^{1/2}} \exp\left\{-\frac{1}{2\hat{\sigma}^{2}}\left(\boldsymbol{y}_{i} - \sum_{k=0}^{p} D_{ik}\Phi_{ik}\hat{\boldsymbol{\gamma}}_{k}\right)'S_{i}^{-1}\left(\boldsymbol{y}_{i} - \sum_{k=0}^{p} D_{ik}\Phi_{ik}\hat{\boldsymbol{\gamma}}_{k}\right)\right\}.$$

The statistical model estimated by the adaptive elastic net regularization depends on the regularization parameter, the tuning parameter and the number of basis functions, therefore we need to decide these values objectively. The decision problem is regarded as a model selection or evaluation problem. Although a cross validation is widely used for the model selection problem, it often selects unstable estimates and has high computational burden. Wang et al. (2007a) showed that GCV does not select the true model consistently and that BIC consistently select the true model for the SCAD regularization. We apply the BIC for selecting the regularization parameter λ , tuning parameter α and the number of basis functions M_k . Model selection criterion BIC has the form of

$$BIC = -2\sum_{i=1}^{n} \log f(Y|\hat{\theta}) + edf \log n$$

= $-\log(2\pi\hat{\sigma}^2)\sum_{i=1}^{n} n_i - \sum_{i=1}^{n} \log |S_i| - \sum_{i=1}^{n} n_i + edf \log n,$ (6)

where edf is an effective degrees of freedom for the varying-coefficient model. Matsui et al. (2013) derived an effective degrees of freedom for the varying coefficient models estimated by the regularization method with the L_2 penalty. Applying this result, the effective degrees of freedom for the varying coefficient models estimated by the adaptive elastic net regularization is given by

$$edf = \sum_{i=1}^{n} \sum_{k=1}^{p} \operatorname{tr} \left\{ D_{ik} \Phi_{ik} \left(\sum_{i=1}^{n} \Phi_{ik}' D_{ik} S_{i}^{-1} D_{ik} \Phi_{ik} + n(1-\alpha) \lambda \hat{\sigma}^{2} I_{M_{k}} \right)^{-1} \\ \left(\| \boldsymbol{\zeta}_{k} \| - n\alpha \lambda \hat{w}_{k} \hat{\sigma}^{2} \right)_{+} \frac{1}{\| \boldsymbol{\zeta}_{k} \|} \Phi_{ik}' D_{ik} S_{i}^{-1} \right\}.$$

4 Simulation study

Monte Carlo simulations are conducted in order to examine the effectiveness of our modeling procedure. We generated a repeated measurement data set $\{(\boldsymbol{x}_{ij}, y_{ij}); i = 1, \ldots, n, j = 1, \ldots, n_i\}$, where $\boldsymbol{x}_{ij} = (x_{ij0}, x_{ij1}, \ldots, x_{ijp})'$ with $x_{ij0} = 1$, from a true model

$$y_{ij} = f(\boldsymbol{x}_{ij}) + \varepsilon_{ij}, \quad f(\boldsymbol{x}_{ij}) = \sum_{k=0}^{p} x_{ijk} \beta_k(t_{ij}),$$
$$\varepsilon_{ij} \sim N(0, \sigma^2), \quad \sigma = s \left\{ \max_{i,j} f(\boldsymbol{x}_{ij}) - \min_{i,j} f(\boldsymbol{x}_{ij}) \right\},$$

where we set the number of predictors to be p = 50, which only first 30 predictors are relevant to the response. We investigated whether the proposed model appropriately select predictors which is relevant to the response for various values of sample sizes n and variance parameters s.

First, time points t_{ij} were generated from an uniform distribution U(-0.5, 0.5) and their numbers n_i were uniformly generated as integer values between 4 and 15 for each subject *i*. Then predictors $X_k = X_k(t)$ were generated by following random numbers:

$$X_{1}(t) \sim U(t/10, 2+t/10), \qquad X_{7}(t) \sim N(3\exp(t/30), 1),$$

$$X_{13}(t) \sim N(t, 1), \qquad X_{19}(t) \sim N(\sin(2\pi t) + 2, 1),$$

$$X_{25}(t) \sim N(\cos(2\pi t) + 2, 1)$$

for k = 1, 7, 13, 19, 25 respectively. They are independent of any other variables. Other variables for $k \leq 30$, which are dependent on above variables, are generated from $N((1 + x_{ik'}(t))/(2 + x_{ik'}(t)), 1)$ where $k' = 6\lfloor (k - 1)/6 \rfloor + 1$ with a floor function $\lfloor \cdot \rfloor$ for each *i*. Remaining variables $X_k(t)$ for $k = 31, \ldots, 50$, independent of each other, are generated from a multivariate normal distribution with mean vector **0** and variance covariance matrix $\operatorname{cov}(X_k(t), X_k(s)) = 4 \exp(-|t - s|/5)$. Next, coefficient functions are set to be

$$\beta_0(t) = \exp(t^3), \qquad \beta_1(t) = 2 + 3\sin(\pi t/60), \beta_6(t) = 2 + 3\cos(\pi t/60), \qquad \beta_{11}(t) = 2 - 3\sin(\pi (t - 25)/15), \beta_{16}(t) = 2 - 3\cos(\pi (t - 25)/15), \qquad \beta_{21}(t) = 6 - 0.2t^2, \beta_{26}(t) = -4 + (20 - t)^3/2000.$$

Other coefficients for $k \leq 30$ are obtained from

$$\beta_k(t) = \beta_{k''}(t) + \boldsymbol{\eta}' \boldsymbol{\psi}(t),$$

where $k'' = 5\lfloor (k-1)/5 \rfloor + 1$ and $\psi(t) = (\psi_1(t), \ldots, \psi_5(t))'$ is a vector of cubic *B*-spline basis functions and $\eta = (\eta_1, \ldots, \eta_5)'$ is a random variable which follows multivariate normal distribution with mean vector **0** and variance covariance matrix $\operatorname{cov}(\eta_i, \eta_j) = 4 \exp(-|i-j|/5)$. The remaining coefficients are set to be $\beta_k(t) = 0$ for $k = 31, \ldots, 50$, which suggest that the variables X_k for $k = 31, \ldots, 50$ are irrelevant to the response. We considered three patterns of sample size n = 15, 25, 50 and two patterns of variance parameter s = 0.05, 0.1.

For the data set, we applied our varying-coefficient modeling procedure with cubic *B*-spline basis, here we assumed that the number of basis functions $M_k = 6$ for all *k*. Regularization parameters in the penalized log-likelihood function were selected by BIC given in (6). In order to investigate the effectiveness of the proposed method, we compared results of our modeling procedure with those of the adaptive group lasso. We repeated this procedure for 100 times, and then obtained 100 mean squared errors $MSE = \sum_i \sum_j \{f(\mathbf{x}_{ij}) - \hat{y}_{ij}\}^2 / \sum_i n_i$, where \hat{y}_{ij} is a predictive value of y_{ij} . Furthermore we examined numbers of selected variables and proportions of variables correctly estimated by zero and incorrectly estimated by zero so as to investigate the variable selection performance.

Table 1 shows results of simulation studies. It contains averages and standard deviations of 100 MSEs and selected regularization parameters and averages of tuning parameters. This table also provides numbers of selected variables and proportions that were correctly and incorrectly estimated to be zeros. These results show that our method gives fewer and stable (or competitive) MSEs than the adaptive lasso. The number of selected variables of the adaptive elastic net are larger than that of the adaptive lasso. The proportions of variables correctly estimated to be zeros are fewer than the adaptive lasso,

	Adapt	ive elas	tic net	Adaptive lasso			
n	15	25	50	15	25	50	
s = 0.05							
$ave(MSE) \times 10^{-2}$	0.91	2.43	2.60	1.66	3.49	3.37	
$sd(MSE) \times 10^{-2}$	0.76	1.27	1.01	1.60	1.55	1.26	
$\operatorname{ave}(\lambda) \times 10^2$	4.21	4.17	2.20	5.96	6.17	2.85	
$\mathrm{sd}(\lambda)\! imes\!10^2$	1.48	1.17	0.35	3.45	2.47	1.47	
$\operatorname{ave}(\alpha)$	0.85	0.80	0.80				
select	32.42	24.18	25.55	23.10	9.19	12.49	
correct	35.85	67.15	67.95	50.55	92.70	89.90	
incorrect	38.03	44.63	39.53	59.30	77.57	68.43	
s = 0.1							
$ave(MSE) \times 10^{-2}$	1.59	2.61	3.04	2.37	3.37	3.98	
$sd(MSE) \times 10^{-2}$	1.38	1.39	1.33	1.78	1.56	1.38	
$\operatorname{ave}(\lambda) \times 10^2$	3.97	4.04	2.25	5.70	5.26	3.50	
$\mathrm{sd}(\lambda) \! imes \! 10^2$	1.11	1.07	0.49	3.33	1.89	0.20	
$\operatorname{ave}(\alpha)$	0.85	0.80	0.80				
select	31.68	23.90	24.59	22.00	10.24	10.57	
correct	38.95	66.60	71.55	55.90	89.10	94.05	
incorrect	38.43	45.93	40.33	59.40	76.47	72.07	

Table 1: Result of simulation studies. Notations "ave" and "sd" represent average and standard deviation respectively.

while those incorrectly estimated to be zeros are larger. It suggests that the adaptive elastic net tends to select necessary variables than the adaptive lasso.

5 Real data example

We applied the proposed modeling strategy to the analysis of the multicenter AIDS cohort study data, available on the R package timereg. The aim of the analysis is to investigate the relationship between properties of the human who are infected with the Human Immunodeficiency Virus (HIV) and the percentages of the CD4 cells in their blood. The data set contains cigarette smoking status, age at HIV infection, pre-HIV infection CD4 cell percent and the CD4 cell percentage of each subject, observed at distinct time points after HIV infection. Several researchers have applied the varying-coefficient modeling to the analysis of this data (Fan and Zhang, 2000; Huang et al., 2004; Wang et al., 2008). We applied the varying coefficient model and the group adaptive elastic net regularization in order to investigate which combination of variables is important.

We represent the relationship of variables described above by the time varying-coefficient

Table 2: Selected numbers of each variable for 100 bootstrap samples. "k" denotes the number of the variable. The upper side of the table indicates the result of original variables, while the lower is of artificially appended variables.

k	1	2	3	4						
select	0	0	6	100						
k	5	6	7	8	9	10	11	12	13	14
select	0	0	0	22	22	4	0	0	6	4

model written by

$$y_i(t) = \sum_{k=0}^p x_{ik}\beta_k(t) + \varepsilon_i(t),$$

where predictors x_{ik} represent an overall intercept or baseline (BASE, k = 0), observed time (TIME, k = 1), age at HIV infection (AGE, k = 2), cigarette smoking status represented by 0 or 1 (SMOKE, k = 3), and pre-infection CD4 percent (PreCD4, k = 4) of the *i*-th subject, respectively, response $y_i(t)$ is the CD4 percent of the *i*-th subject observed at differing time points, $\beta_k(t)$ are time varying coefficients and $\varepsilon_i(t)$ are the error functions. We assumed that the variance covariance matrix of the vector ($\varepsilon_i(t_{i1}), \ldots, \varepsilon_i(t_{in_i})$)' were $\sigma^2 I$. In addition to these original predictors, we appended 10 artificial variables to the predictor by the following method. First, we randomly selected a variable from $X_k, k =$ 2, 3, 4 and then assigned a random sample of the selected variable without replacement to a new variable. We repeated this process for 10 times and then treated them as X_k ($k = 5, \ldots, 14$).

Coefficient functions are supposed to be expressed by cubic B-spline basis expansions. Then the model was fitted by the maximum penalized likelihood method with group adaptive elastic net penalty, and it was evaluated by BIC. In order to avoid time consuming process, we assumed that the number of basis functions was the same among variables and the number of basis functions was set to be 6. We generated 100 sets of bootstrap samples from the data, and then obtained 100 estimates of coefficient functions and numbers of selected variables.

Table 2 shows sums of numbers of selected variables except for BASE (intercept) for 100 repetition. We can find that PreCD4 was most selected and other original predictors were hardly selected. In addition, artificially appended variables were less selected. These results show that PreCD4 strongly has an influence on CD4 cell percentage, while other variables such as SMOKE and AGE have little relevance. Furthermore the proposed method can also appropriately exclude variables which are truly unnecessary.

Estimated coefficients are shown in Figure 1. Solid lines are mean coefficient functions for 100 bootstrap samples and dashed lines are pointwise 90 % confidence intervals. Figure



Figure 1: Estimated coefficient functions for original variables. Top left: BASE, top right: TIME, bottom left: AGE, bottom middle: SMOKE, bottom right: PreCD4.

1 suggests that the CD4 data have a trend that decreases with time in early time and then increases gradually, and that PreCD4 has a positive influence on CD4 cell percentage, but it gradually becomes weak with time. On the other hand, TIME, AGE and SMOKE seem to have almost no effect on the CD4 percentage all the time since most coefficients of them are estimated to be exactly zeros. It reveals that these three variables are irrelevant to the CD4 cell percentage. These results are quite similar to those of Wang et al. (2008).

6 Concluding Remarks

In this article, we have applied the sparse regularization to varying-coefficient models in order to select variables repeatedly measured at possibly different time points. We used an adaptive elastic net penalty in the penalized likelihood method, and then parameters involved in the model are estimated by the framework of the coordinate descent algorithm. In order to select regularization parameters involved in the penalized likelihood method, we derived an effective degrees of freedom for the varying coefficient model and then apply a BIC-type criterion. Simulation studies suggest that our modeling strategy worked well in the viewpoint of variable selection and prediction accuracy rather than the existing method. Furthermore we applied it to the analysis of real data and showed that some of information of the HIV patient seems to irrelevant to the CD4 cell percentage. More recently, Şentürk and Müller (2010) proposed a new type of varying-coefficient models which considers the relationship between a response and predictors with recent time points. Future work will focus on applying the sparse regularization to such models to investigate the time range that the predictors affect the response.

References

- Cai, Z., Fan, J., and Li, R. (2000), "Efficient estimation and inferences for varyingcoefficient models," J. Amer. Statist. Assoc., 95, 888–902.
- de Boor, C. (2001), A practical guide to splines, New York: Springer Verlag.
- Diggle, P. (2002), Analysis of longitudinal data, Oxford: Oxford University Press.
- Eubank, R., Huang, C., Maldonado, Y., Wang, N., Wang, S., and Buchanan, R. (2004), "Smoothing spline estimation in varying-coefficient models," J. Roy. Statist. Soc. Ser. B., 66, 653–667.
- Fan, J. and Li, R. (2001), "Variable selection via nonconcave penalized likelihood and its oracle properties," J. Amer. Statist. Assoc., 96, 1348–1360.
- Fan, J. and Zhang, J. (2000), "Two-step estimation of functional linear models with applications to longitudinal data," J. Roy. Statist. Soc. Ser. B., 62, 303–322.
- Fan, J. and Zhang, W. (1999), "Statistical estimation in varying coefficient models," Ann. Statist., 27, 1491–1518.
- Friedman, J., Hastie, T., Höfling, H., and Tibshirani, R. (2007), "Pathwise coordinate optimization," Ann. Appl. Statist., 1, 302–332.
- Hastie, T. and Tibshirani, R. (1993), "Varying-coefficient models," J. Roy. Statist. Soc. Ser. B., 55, 757–796.
- Hastie, T., Tibshirani, R., and Friedman, J. (2009), *The elements of statistical learning* 2nd ed., New York: Springer.
- 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.
- Huang, J., Wu, C., and Zhou, L. (2002), "Varying-coefficient models and basis function approximations for the analysis of repeated measurements," *Biometrika*, 89, 111–128.
- (2004), "Polynomial spline estimation and inference for varying coefficient models with longitudinal data," *Statist. Sinica*, 14, 763–788.

- Imoto, S. and Konishi, S. (2003), "Selection of smoothing parameters in B-spline nonparametric regression models using information criteria," Ann. Inst. Statist. Math., 55, 671–687.
- Laird, N. and Ware, J. (1982), "Random-effects models for longitudinal data," *Biometrics*, 963–974.
- Matsui, H., Misumi, T., and Kawano, S. (2013), "Model selection criteria for the varyingcoefficient modeling via regularized basis expansions," To appear in J. Statist. Comput. Simulation.
- Noh, H. and Park, B. (2010), "Sparse varying coefficient models for longitudinal data," *Statist. Sinica*, 20, 1183–1202.
- Ramsay, J. and Silverman, B. (2005), *Functional data analysis*, New York: Springer Verlag.
- Şentürk, D. and Müller, H. (2010), "Functional varying coefficient models for longitudinal data," J. Amer. Statist. Assoc., 105, 1256–1264.
- 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. (2007a), "Tuning parameter selectors for the smoothly clipped absolute deviation method," *Biometrika*, 94, 553–568.
- Wang, H., Zou, G., and Wan, A. (2013), "Adaptive LASSO for varying-coefficient partially linear measurement error models," J. Statist. Plann. Inference, 143, 40–54.
- Wang, L., Chen, G., and Li, H. (2007b), "Group SCAD regression analysis for microarray time course gene expression data," *Bioinformatics*, 23, 1486–1494.
- Wang, L., Li, H., and Huang, J. (2008), "Variable selection in nonparametric varyingcoefficient models for analysis of repeated measurements," J. Amer. Statist. Assoc., 103, 1556–1569.
- Wei, F., Huang, J., and Li, H. (2011), "Variable selection and estimation in highdimensional varying-coefficient models," *Statist. Sinica*, 21, 1515–1540.
- Xue, L. and Qu, A. (2012), "Variable Selection in High-dimensional Varying-coefficient Models with Global Optimality," J. Mach. Learn. Res., 13, 1973–1998.
- Yuan, M. and Lin, Y. (2006), "Model selection and estimation in regression with grouped variables," J. Roy. Statist. Soc. Ser. B., 68, 49–67.

- Zeger, S. and Diggle, P. (1994), "Semiparametric models for longitudinal data with application to CD4 cell numbers in HIV seroconverters," *Biometrics*, 50, 689–699.
- Zou, H. (2006), "The adaptive lasso and its oracle properties," J. Amer. Statist. Assoc., 101, 1418–1429.
- Zou, H. and Hastie, T. (2005), "Regularization and variable selection via the elastic net," J. Roy. Statist. Soc. Ser. B., 67, 301–320.
- Zou, H. and Zhang, H. (2009), "On the adaptive elastic-net with a diverging number of parameters," Ann. Statist., 37, 1733.

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 characteristiccurve 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 -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 threespace
- 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 highfrequency 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
- MI2012-12 Masato WAKAYAMA On simplicity of the lowest eigenvalue of non-commutative harmonic oscillators
- MI2012-13 Masatoshi OKITA On the convergence rates for the compressible Navier- Stokes equations with potential force
- MI2013-1 Abuduwaili PAERHATI & Yasuhide FUKUMOTO A Counter-example to Thomson-Tait-Chetayev's Theorem
- MI2013-2 Yasuhide FUKUMOTO & Hirofumi SAKUMA A unified view of topological invariants of barotropic and baroclinic fluids and their application to formal stability analysis of three-dimensional ideal gas flows
- MI2013-3 Hiroki MASUDA Asymptotics for functionals of self-normalized residuals of discretely observed stochastic processes
- MI2013-4 Naoyuki KAMIYAMA On Counting Output Patterns of Logic Circuits
- MI2013-5 Hiroshi INOUE RIPless Theory for Compressed Sensing

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