ISSN Print: 2394-7500
ISSN Online: 2394-5869
Impact Factor: 5.2
IJAR 2020; 6(9): 492-495
www.allresearchjournal.com
Received: 11-07-2020
Accepted: 15-08-2020

Amit Kumar Singh
Research Scholar, Dept. of Mathematics, J.P. Univ. Chapra, Bihar, India

Dr. Ashok Kumar
Associate Professor, Dept. of Mathematics, DAV PG College, Siwan, Bihar, India

Corresponding Author:
Amit Kumar Singh
Research Scholar, Dept. of Mathematics, J.P. Univ. Chapra, Bihar, India

Study on the adaptive functional linear regression

Amit Kumar Singh and Dr. Ashok Kumar

Abstract
In this paper, we consider adaptive estimation in functional linear regression. The goal is to construct a single data-driven procedure that achieves optimality results simultaneously over a collection of parameter spaces. Such an adaptive procedure automatically adjusts to the smoothness properties of the underlying slope and covariance functions.

Keywords: Functional linear, data-driven, smoothness properties

Introduction

There has been extensive recent research on functional data analysis. Much progress has been made on developing methodologies for analyzing functional data. The two monographs [1 2]. provide comprehensive discussions on the methods and applications [3] Among many problems involving functional data, functional linear regression has received substantial attention. Consider a functional linear model where one observes a random sample $\{(X_i, Y_i) : i = 1,\ldots, n\}$ with

$$Y_i = a + \int_0^1 X_i(t)b(t)dt + Z_i, \quad (1)$$

where the response Y_i and the intercept a are scalars, the predictor X_i and slope function b are functions in $L_2([0, 1])$, and the errors Z_i are independent and identically distributed $N(0, \sigma^2)$ variables. The goal is to estimate the slope function b(t) and the intercept a based on the sample $\{(X_i, Y_i) : i = 1,\ldots, n\}$. Note that once an estimator $\hat{b}$ of b is constructed, the intercept a can be estimated easily by $\hat{a} = \bar{Y} - \int_0^1 \bar{X}(t)\hat{b}(t)dt,$

where $\bar{Y}$ and $\bar{X}$ are the averages of Y_i and X_i respectively. We shall thus focus our discussion in this paper on estimating the slope function b. The slope function is of significant interest on its own right. For example, knowing where b takes large or small values provides information about where a future observation x of X will have greatest leverage on the conditional mean of y given X = x.

The problem of slope-function estimation is intrinsically nonparametric and the convergence rate under the mean integrated squared error (MISE)

$$R(\hat{b}, b) = \mathrm{E}\,\|\hat{b} - b\|_2^2 = \mathrm{E}\int_0^1 (\hat{b}(t) - b(t))^2 dt \quad (2)$$

is typically slower than n^{-1}. Rates of convergence of an estimator $\hat{b}$ to b have been studied in[4–8] showed that the minimax rate of convergence for estimating b under the MISE (2) is determined by the smoothness of the slope function, and of the covariance function for the distribution of explanatory variables [9]. considered a related prediction problem [10] studied generalized functional linear models.

The theory on slope function estimation has so far focused on the minimax estimation where these smoothness parameters are assumed to be known. The estimators typically depend on the smoothness parameters. Although minimax risk provides a useful uniform benchmark for the comparison of estimators, minimax estimators often require full knowledge of the parameter space which is unknown in practice. A minimax estimator designed for a specific parameter space typically performs poorly over another parameter space. This makes adaptation essential for functional linear regression. In the paper we consider adaptive estimation of the slope function b. The goal is to construct a single data-driven procedure that achieves optimality results simultaneously over a collection of parameter spaces.

Methodology

Estimating the slope function b in function linear regression involves solving an ill-posed inverse problem. The main difference with the conventional linear inverse problems is that the operator is not given in the functional linear regression. A major technical step in the construction of the slope function estimator is to estimate the eigenvalues and eigenfunctions of the unknown linear operator and to bound the errors between the estimates and the estimands.

Spectral decomposition

Suppose we observe a random sample $\{(X_i, Y_i) : i = 1,..., n\}$ as in (2.1). Let (X, Y, Z) denote a generic (X_i, Y_i, Z_i). Define the covariance function and the empirical covariance function respectively as

K(u, v) = cov{X(u);X(v)}

$$\hat{K}(u, v) = \frac{1}{n}\sum_{i=1}^{n}\{X_i(u)-\bar{X}(u)\}-\{X_i(v)-\bar{X}(v)\}$$

where $\bar{X}$ = (1=n) ΣX_i. The covariance function K defines a linear operator which maps a function f to Kf given by (Kf)(u) = ∫ K (u, v) f (v)dv. We shall assume that the linear operator with kernel K is positive definite. Write the spectral decompositions of the covariance functions K and $\hat{K}$ as

$$K(u,v)=\sum_{j=1}^{\infty}\theta_j\phi_j(u)\phi_j(v),\quad \hat{K}(u,v)=\sum_{j=1}^{\infty}\hat{\theta}_j\hat{\phi}_j(u)\hat{\phi}_j(v), \tag{3}$$

where

$$\theta_1>\theta_2>.....>0,\ and\ \hat{\theta}_1\geq\hat{\theta}_2\geq..\hat{\theta}_2\geq...\geq\hat{\theta}_{n+1}=...=0 \tag{4}$$

are respectively the ordered eigenvalue sequences of the linear operators with kernels K and $\hat{K}$, and {ϕ} and { $\hat{\phi}_j$ } are the corresponding orthonormal eigenfunction sequences. The sequences {ϕ} and { $\hat{\phi}_j$ } each forms an orthonormal basis in L_2([0, 1]). The functional linear model (1) can be rewritten as

$$Y_i=\mu+\int[X_i-E(X)b+Z_i,\ i=1,2,....,n] \tag{5}$$

where μ = E(Y_i) = a + E ∫ Xb. The Karhunen-Loeve expansion of the random function X_i – EX is given by

$$X_i-EX=\sum_{j=1}^{\infty}x_{i,j}\phi_j \tag{6}$$

where the random variable $x_{i,j}$ =f (X_i–EX)ϕ_j has mean zero and variance Var($x_{i,j}$) = θ_j. In addition, the random variables $x_{i,j}$ are uncorrelated. Expand the slope function b in the orthonormal basis {ϕ_j} as b = $\sum_{j=1}^{\infty}b_j\phi_j$. Then the model (5) can be written as

$$Y_i=\mu+\sum_{j=1}^{\infty}x_{i,j}b_j+Z_i, i=1,2,...,n \tag{7}$$

and the problem of estimating the slope function b is transformed into the one of estimating the coefficients [b_j] as well as the eigenfunctions {ϕ_j}. Note that in (6) μ and $x_{i,j}$ are unknown, and thus need to be estimated from the data.

The mean μ of Y can be estimated easily by the sample mean $\hat{\mu}=\bar{Y}$. To estimate the $x_{i,j}$, we expand $X_i-\bar{X}$ in the orthonormal basis { $\hat{\phi}_j$ } as

$$X_i - \overline{X} = \sum_{j=1}^{n} \hat{x}_{i,j} \hat{\phi}_j \text{ for } i = 1, 2, \ldots, n \tag{8}$$

where the random variables $\hat{x}_{i,j} = \int (X_i - \overline{X}) \hat{\phi}_j$. Note that

$$\sum_{i=1}^{n} \hat{x}_{i,j} = \sum_{i=1}^{n} \int (X_i - \overline{X}) \hat{\phi}_j = \int \left[\sum_{i=1}^{n} (X_i - \overline{X}) \right] \hat{\phi}_j = 0$$

and

$$\frac{1}{n} \sum_{i=1}^{n} \hat{x}_{i,j} \hat{x}_{i,k} = \int \int \hat{K}(u,v) \hat{\phi}_j(u) \hat{\phi}_k(v) = \hat{\theta}_j \delta_{j,k} \tag{9}$$

for all j and k, where $\delta_{j,k}$ is the Kronecker delta with $\delta_{j,k} = 1$ if j = k and 0 otherwise. Since $\overline{Y} = a + \int_0^1 \overline{X}(t) b dt + \overline{Z}$, we have

$$Y_i - \overline{Y} = \int [X_i - \overline{X}] b + Z_i - \overline{Z}, i = 1, 2, \ldots, n.$$

Hence

$$Y_i - \overline{Y} = \sum_{j=1}^{n} \hat{x}_{i,j} \hat{b}_j + Z_i - \overline{Z}, i = 1, 2, \ldots, n \tag{10}$$

where $\breve{b}_j = \int b \hat{\phi}_j$, and consequently b = $\sum_{j=1}^{\infty} \hat{b}_j \hat{\phi}_j$. Since the slope function b is unknown, the coefficients $\hat{b}_j$ are also unknown and need to be estimated. A typical principal components regression approach is to replace "n" in equation (9) by a constant m < n and estimate $\hat{b}_j$ by ordinary least squares. Since the "predictors" $(\hat{x}_{i,j})_{1 \le j \le n}$ in equation (9) are orthogonal to each other and $\sum_{i=1}^{n} \hat{x}_{ij}^2 = \hat{\theta}_j n$ from (8), for $\hat{\theta}_j \neq 0$ we may estimate $\hat{b}_j$ (or b_j) by

$$\begin{aligned} \tilde{b}_j &= \hat{\theta}_j^{-1} n^{-1} \sum_{i=1}^{n} (Y_i - \overline{Y}) \hat{x}_{i,j} = \hat{\theta}_j^{-1} n^{-1} \sum_{i=1}^{n} (Y_i - \overline{Y}) \int [X_i(u) - \overline{X}(u)] \hat{\phi}_j(u) \\ &= \hat{\theta}_j^{-1} \int \hat{g}(u) \hat{\phi}_j(u) = \hat{\theta}_j^{-1} \hat{g}_j \end{aligned} \tag{11}$$

Where

$$\hat{g}(u) = n^{-1} \sum_{i=1}^{n} (Y_i - \overline{Y}) [X_i(u) - \overline{X}(u)] \text{ and } \hat{g}_j = \int \hat{g}_j = \int \hat{g} \hat{\phi}_j . \tag{12}$$

It is expected that $\hat{g}$ is approximately

$$g(u) = E[Y - \mu(X(u) - E(X(u))] = \int \int K(u,v) b(v) dv . \tag{13}$$

Write g = $\sum_{j=1}^{\infty} g_j \phi_j$. It is easy to check that $b_j = \theta_j^{-1} g_j$. So the estimator $\tilde{b}_j = \hat{\theta}_j^{-1} \hat{g}_j$ in (10) can be regarded as an empirical version of the true coefficient b_j.

References

1. Ramsay JO, Silverman BW. Applied Functional Data Analysis: Methods and Case Studies. Springer, New York 2012.
2. Ramsay JO, Silverman BW. Functional Data Analysis, 2nd Edition. Springer, New York 2015.
3. Ferraty F, Vieu P. Nonparametric Functional Data Analysis. Springer, New York 2016.
4. Ferraty F, Vieu P. Fractal dimensionality and regression estimation in semi-normed vectorial spaces. C. R. Acad. Sci. Paris Ser. 2000;I(330):139-142.

5. Cuevas A, Febrero M, Fraiman R. Linear functional regression: the case of fixed design and functional response. Canad. J. Statist 2002;30:285-300
6. Cardot, Sarda P. Linear regression models for functional data. In The Art of Semiparametrics, Physica-Verlag HD 2006 49-66.
7. Li Y, Hsing T. On rates of convergence in functional linear regression. J. Multivariate Analysis 2017;98:1782-1804.
8. Hall P, Horowitz JL. Methodology and convergence rates for functional linear regression. Ann. Statist 2016;35:70-91.
9. Cai TT, Hall P. Prediction in functional linear regression. Ann. Statist 2006;34:2159-2179.
10. Muller HG, Stadtmiiller U. Generalized functional linear models. Ann. Statist 2005;33:774-805.