Ann. Appl. Math.
doi: 10.4208/aam.OA-2021-0016

Vol. **38**, No. 3, pp. 280-295
August 2022

A Gradient Iteration Method for Functional Linear Regression in Reproducing Kernel Hilbert Spaces

Hongzhi Tong[1] and Michael Ng[2,*]

[1] *School of Statistics, University of International Business and Economics, Beijing 100029, China*

[2] *Institute of Data Science and Department of Mathematics, The University of Hong Kong, Pokfulam, Hong Kong, China*

Received 1 December 2021; Accepted (in revised version) 31 May 2022

Abstract. We consider a gradient iteration algorithm for prediction of functional linear regression under the framework of reproducing kernel Hilbert spaces. In the algorithm, we use an early stopping technique, instead of the classical Tikhonov regularization, to prevent the iteration from an overfitting function. Under mild conditions, we obtain upper bounds, essentially matching the known minimax lower bounds, for excess prediction risk. An almost sure convergence is also established for the proposed algorithm.

AMS subject classifications: 60K35, 62J05

Key words: Gradient iteration algorithm, functional linear regression, reproducing kernel Hilbert space, early stopping, convergence rates.

---

1 Introduction

Due to advance in technology, data are increasingly collected in the form of random functions or curves, as opposed to scalars or vectors. Functional data analysis (FDA) is developed to handle this situation, has drawn considerable attention in recent decades. Various approaches for the analysis of functional data have been developed and proposed in the literature [6, 8, 12, 13, 17], offering a comprehensive overview.

*Corresponding author.
Emails: `tonghz@uibe.edu.cn` (H. Tong), `mng@maths.hku.hk` (M. Ng)

Among many problems involving functional data, functional linear regression is widely used to model the prediction of a functional predictor. Consider the following functional linear regression model

$$Y=\alpha_0+\int_{\mathbb{I}}X(s)\beta_0(s)ds+\epsilon, \tag{1.1}$$

where Y is a scalar response, $X(\cdot)$ is a square integrable random function (with respect to Lesbesgue measure) on a bounded interval $\mathbb{I}$, α_0 is the intercept, $\beta_0(\cdot)$ is an unknown slope function and ϵ is a centered noise random variable. Without loss of much generality, throughout the paper we assume $\mathbb{E}(X)=0$ and the intercept $\alpha_0=0$, since the intercept can be easily estimated.

The goal of the prediction problems is to estimate the functional

$$\eta_0(X):=\int_{\mathbb{I}}X(s)\beta_0(s)ds$$

based on a set of training data $\{(X_i,Y_i):i=1,\cdots,n\}$ consisting of n independent copies of (X,Y). Define the risk for a prediction η as

$$\mathcal{E}(\eta):=\mathbb{E}^*[Y^*-\eta(X^*)]^2,$$

where (X^*,Y^*) is a copy of (X,Y) independent of the training data, and $\mathbb{E}^*$ represents expectations taken over X^* and Y^* only. Let $\hat{\eta}$ be a prediction constructed from the training data. Then, its accuracy can be naturally measured by the excess risk:

$$\mathcal{E}(\hat{\eta})-\mathcal{E}(\eta_0)=\mathbb{E}^*[\hat{\eta}(X^*)-\eta_0(X^*)]^2. \tag{1.2}$$

In the context of functional linear regression, most of existing methods are based upon functional principal analysis (FPCA), see, e.g., [2, 7, 13, 19]. The FPCA approach expands the unknown function β_0 using the eigenfunction of the predictor covariance operator. For such a strategy to work well, it is usually necessary to assume that such a basis provides a good presentation of the slope function, which may not have anything to do with the predictor in terms of basis representation accuracy. A more general assumption for slope function may be on its smoothness, it makes a reproducing kernel Hilbert space (RKHS) [9] an interesting alternative, see [3, 11, 20]. In particular, it has already been shown in [3] that the approach based on RKHS performs better when the slope function does not align well with the eigenfunctions of the covariance kernel.

Motivated by these observations, in this paper we will develop an iterative estimation procedure for functional linear model (1.1) within the framework of RKHS, under which the unknown slope function β_0 is assumed to reside in an RKHS $\mathcal{H}_K$.