Accepted Manuscript

Reduced rank modeling for functional regression with functional responses

Hongmei Lin, Xuejun Jiang, Heng Lian, Weiping Zhang

PII: S0047-259X(18)30460-3
DOI: https://doi.org/10.1016/j.jmva.2018.09.004
Reference: YJMVA 4414

To appear in: *Journal of Multivariate Analysis*

Received date : 6 September 2018

Reduced rank modeling for functional regression with functional responses

Hongmei Lin[a], Xuejun Jiang[b], Heng Lian[c,], Weiping Zhang[d]

[a]*School of Statistics and Information, Shanghai University of International Business and Economics, Shanghai, China*
[b]*Department of Mathematics, Southern University of Science and Technology, Shenzhen, China*
[c]*Department of Mathematics, City University Hong Kong, Hong Kong*
[d]*Department of Statistics and Finance, University of Science and Technology of China, Hefei, China, 230026*

Abstract

This article considers regression problems where both the predictor and the response are functional in nature. Driven by the desire to build a parsimonious model, we consider functional reduced rank regression in the framework of reproducing kernel Hilbert spaces, which can be formulated in the form of linear factor regression with estimated multivariate factors, and achieves dimension reduction in both the predictor and the response spaces. The convergence rate of the estimator is derived. Simulations and real data sets are used to demonstrate the competitive performance of the proposed method.

Keywords: Dimension reduction, Functional data, Functional response, Reproducing kernel Hilbert space.

1. Introduction

It is increasingly common to deal with regression problems in which the predictor, the response or both are functional in nature, with recent contributions include but not limited to [2, 6, 9, 12, 13, 16, 17, 20, 23, 24, 27, 29]. In this article we consider the following functional linear regression model with functional response

$$Y(t) = \mu(t) + \int_0^1 \beta(t, s)X(s)\, ds + \epsilon(t), \tag{1}$$

where $Y, X, \epsilon \in L^2([0, 1])$ and $\mathrm{E}(\epsilon \mid X) = 0$. This problem has been studied in [1, 4, 10, 25, 27]. We assume that the entire functional predictor and response are observed. As the mean function can be trivially estimated in this situation, for simplicity and without loss of generality we assume $\mathrm{E}(Y) = \mathrm{E}(X) = 0$ and thus do not model the intercept $\mu(t)$.

With either scalar or functional responses, there are several approaches to fit the functional linear model given a random sample $(X_1, Y_1), \ldots, (X_n, Y_n)$. The traditionally most popular one is to approximate functional variables via basis expansion. Polynomial splines were used in [25] whereas Fourier bases are more appropriate for periodic functional data. Wavelet bases in functional data analysis were popularized because of their suitability for modeling spatially heterogeneous functions [21, 22]. Random basis functions obtained from functional principal component analysis (PCA) are also used, with particular advantages in theoretical analysis [8, 13].

The framework adopted here is based on reproducing kernel Hilbert spaces (RKHS), which were studied in [7, 28], and extended in [18] to functional responses. Assuming that for all $t \in [0, 1]$, $\beta(t, \cdot)$ is in an RKHS $\mathcal{H}_K$ with kernel K, and denoting the L^2 norm and the RKHS norm by $\|\cdot\|$ and $\|\cdot\|_{\mathcal{H}_K}$, respectively, we can estimate β by

$$\hat{\beta} = \arg\min_{\beta} \frac{1}{n}\sum_{i=1}^{n} \left\| Y_i - \int_0^1 \beta(\cdot, s)X_i(s)\, ds \right\|^2 + \lambda \int_0^1 \|\beta(t, \cdot)\|^2_{\mathcal{H}_K}\, dt. \tag{2}$$

Email address: `henglian@cityu.edu.hk` (Heng Lian)