Computational Statistics and Data Analysis 70 (2014) 183–197

Contents lists available at ScienceDirect

Computational Statistics and Data Analysis

journal homepage: www.elsevier.com/locate/csda

Model detection for functional polynomial regression

Tao Zhang [a], Qingzhao Zhang [b], Qihua Wang [b,c,*]

[a] College of Science, Guangxi University of Technology, Liuzhou, China
[b] Academy of Mathematics and Systems Sciences, Chinese Academy of Sciences, Beijing 100190, China
[c] Institute of Statistical Science, Shenzhen University, Shenzhen 518060, China

ARTICLE INFO

Article history:
Received 26 December 2012
Received in revised form 7 September 2013
Accepted 9 September 2013
Available online 16 September 2013

Keywords:
Functional polynomial regression model
Functional principal components
Adaptive group Lasso
Consistency

ABSTRACT

A functional polynomial regression model which includes the functional linear model and functional quadratic model as two special cases is considered. In functional polynomial regression, one must balance the costs and benefits of using more parameters in the model. The method of model detection to determine which orders of the polynomial are significant in functional polynomial regression is developed. The proposed methods can identify the true model consistently and have good prediction performances. Numerical studies clearly confirm our theories.

1. Introduction

Recent technological advances in collecting and storing data have confronted statisticians with situations where the datasets are of a functional nature (curves, images, etc.) with the need to build new models and develop new methods. This field of research, known as Functional Data Analysis (FDA), has been popularized by Ramsay and Silverman (2005). The first advances in nonparametric FDA are described in Ferraty and Vieu (2006) (see also Oxford Handbook of FDA by Ferraty and Romain, 2011).

In functional regression, special attention has been paid to functional linear models (Cardot et al., 2003; Shen and Faraway, 2004; Cai and Hall, 2006; Hall and Horowitz, 2007). However, it is pointed out in Yao and Müller (2010) that this model imposes a constraint on the regression relationship that may not be appropriate in some scenarios. Fully nonparametric methods have been studied recently in functional-data regression and related problems. See, e.g., Ferraty and Vieu (2003, 2006) and Müller and Stadtmüller (2005). However, the problems of nonparametric regression and prediction are intrinsically difficult from a statistical viewpoint. In particular, convergence rates can be slower than the inverse of any polynomial in sample size, and so relatively large samples may be needed in order to ensure adequate performance.

Because the functional linear model is too restrictive on the regression relationship and nonparametric methods have slow convergence rates, it is of interest to consider a compromise of the functional linear model and nonparametric model that has better properties, but at the same time enhance its flexibility. Hence, the single-index functional regression model was studied by Ait-Saïdi et al. (2008) using a cross-validated method. Chen et al. (2011) also considered the single-index functional regression model using a method similar to Ait-Saïdi et al. (2008) and extended it to multiple index functional

* Corresponding author at: Academy of Mathematics and Systems Science, Chinese Academy of Science, Beijing 100190, China. Tel.: +86 10 62651336; fax: +86 62541689.
E-mail address: qhwang@amss.ac.cn (Q. Wang).

0167-9473/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.csda.2013.09.007

regression models. Also, a polynomial rate of convergence was obtained. Li et al. (2010) used a semiparametric single index structure to model the potential interaction between the functional predictor and other covariates. Yao and Müller (2010) generalized functional linear models to a functional polynomial model, which has greater flexibility.

In addition, it is well known that prediction accuracy can sometimes be improved by decreasing the model complexity. That is, one can reduce the variance of the predicted values by sacrificing a little bit of bias, and hence may improve the overall prediction performance. Variable selection plays a fundamental role in this research field for regular regression analysis. Traditional methods such as AIC, BIC and best subset selection suffer from a huge computational burden (Fan and Li, 2001). Some shrinkage methods have been developed, which not only have less computational cost but can also simultaneously select the significant variables and estimate the unknown regression coefficients. Tibshirani (1996) developed LASSO, which became more popular after the LARS algorithm was proposed by Efron et al. (2004). The smoothly clipped absolute deviation (SCAD) penalty was studied by Fan and Li (2001), and achieves the oracle properties. Zou (2006) argued the possible inconsistency in selection by LASSO and proposed adaptive LASSO. Inspired by variable selection problems in ANOVA and nonparametric additive regression, Yuan and Lin (2006) proposed the group LASSO method. Wang and Xia (2009) borrowed the idea of group LASSO and developed KLASSO, which can consistently distinguish the predictors with varying effects and the irrelevant predictors.

In functional data analysis, we also face the problem of which orders of the polynomial are significant in functional polynomial regression analysis. For example, it is expected to develop a method which can detect the quadratic term and all the higher order terms in a functional polynomial model to be ignorable if a functional linear model can adequately describe the regression relationship. It is noted that the shrinkage methods seem a feasible approach to deal with this issue. However, up to now, these shrinkage methods have been investigated mostly in the situation where all the explanatory variables take real values. Extending these shrinkage ideas to the functional polynomial regression is challenging due to the intrinsic infinite dimension of functions. To the best of our knowledge, the model detection problem for functional polynomial regression has not been considered.

Therefore, we are motivated to develop a detection method to identify which orders in a functional polynomial are significant. To achieve this goal, we first project the predictors on a suitable basis of the underlying space. Then the functional polynomial regression model can be alternatively expressed as a function of the principal component scores of predictor processes, which lead to parsimonious representations after truncating principal components at a reasonable number. As a result, each order of the functional polynomial model can be represented as a linear combination of some products of principal component scores. Accordingly, these product terms with the same order can be regarded as a group. This motivates us to utilize adaptive group lasso to solve the problem of identifying significant orders of a functional polynomial. It is remarkable that the proposed methodology provides a general framework of model structure detection for the functional polynomial regression model.

This paper is organized as follows. In Section 2, we introduce functional polynomial models, while all of the estimating procedures are developed in Section 3. The asymptotic theory of the proposed procedures is established in Section 4, while Section 5 is devoted to a report on simulation results. In Section 6, we analyze a real dataset to illustrate the proposed procedures. Some discussions are contained in Section 7. All proofs are delayed to Section 8.

2. Functional polynomial regression model

Suppose we have independent and identically distributed (i.i.d.) observations $\{(X_i, Y_i)_{i=1,\dots,n}\}$. We consider the pth order ($p \geq 2$) functional polynomial regression model with scale response Y_i and functional predictor X_i due to Yao and Müller (2010),

$$Y_i = \alpha + \int_T \gamma_1(s) X_i^c(s) ds + \cdots + \int_{T^p} \gamma_p(t_1, \dots, t_p) X_i^c(t_1) \cdots X_i^c(t_p) dt_1 \cdots dt_p + \varepsilon_i, \tag{2.1}$$

where $E(\varepsilon_i|X_i) = 0, i = 1, \dots, n$ and α is an intercept. $\gamma_j, 1 \leq j \leq p$, are jth order regression parameter functions, respectively. $X_i^c(s) = X_i(s) - \mu_X(s)$ denotes the centered predictor process with $\mu_X(s) = EX_i(s)$. The regression parameter functions are assumed to be smooth and square integrable. Yao and Müller (2010) mainly studied the estimating problem for the functional quadratic regression model, which is a special case of model (2.1) with $p = 2$. However, we focus on the model detection and estimation for model (2.1).

The predictor processes $X_i, i = 1, \dots, n$ are assumed to be non-stationary smooth random functions in $L_2(T)$ with smooth auto-covariance function $\text{cov}(X_i(s_1), X_i(s_2)) = G_X(s_1, s_2)$. It is assumed that processes $X_i, i = 1, \dots, n$ possess Karhunen–Loève expansions with representations

$$X_i(s) = \mu_X(s) + \sum_{j=1}^{\infty} \eta_{ij} \phi_j(s), \tag{2.2}$$

where the coefficients η_{ij} are a sequence of uncorrelated random variables with means $E(\eta_{ij}) = 0$ and variances $\text{var}(\eta_{ij}) = \nu_j$, and ϕ_j is a sequence of orthogonal eigenfunctions of the auto-covariance function $G_X(s_1, s_2)$ with corresponding non-increasing eigenvalues ν_j.