Journal of Multivariate Analysis 155 (2017) 252–271

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

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

An innovations algorithm for the prediction of functional linear processes

J. Klepsch *, C. Klüppelberg

Center for Mathematical Sciences, Technical University of Munich, 85748 Garching, Boltzmannstrasse 3, Germany

ARTICLE INFO

Article history:
Received 23 August 2016
Available online 10 January 2017

AMS 2010 subject classifications:
62M10
62M15
62M20
62H25
60G25

Keywords:
Functional data analysis (FDA)
Functional linear process
Functional principal components
Functional time series
Hilbert space valued process
Innovations Algorithm
Prediction
Prediction error

ABSTRACT

When observations are curves over some natural time interval, the field of functional data analysis comes into play. Functional linear processes account for temporal dependence in the data. The prediction problem for functional linear processes has been solved theoretically, but the focus for applications has been on functional autoregressive processes. We propose a new computationally tractable linear predictor for functional linear processes. It is based on an application of the Multivariate Innovations Algorithm to finite-dimensional subprocesses of increasing dimension of the infinite-dimensional functional linear process. We investigate the behavior of the predictor for increasing sample size. We show that, depending on the decay rate of the eigenvalues of the covariance and the spectral density operator, the resulting predictor converges with a certain rate to the theoretically best linear predictor.

1. Introduction

We consider observations which are consecutive curves over a fixed time interval within the field of functional data analysis (FDA). In this paper curves are representations of a functional linear process. The data generating process is a time series $X = (X_n)_{n\in\mathbb{Z}}$ where each X_n is a random element $X_n(t)$, $t \in [0, 1]$, of a Hilbert space, often the space of square integrable functions on $[0, 1]$.

Several books contain a mathematical or statistical treatment of dependent functional data as Bosq [4], Horvàth and Kokoszka [14], and Bosq and Blanke [7]. The main source of our paper is the book Bosq [4] on linear processes in function spaces, which gives the most general mathematical treatment of linear dependence in functional data, developing estimation, limit theorems and prediction for functional autoregressive processes. In Hörmann and Kokoszka [13] the authors develop limit theorems for the larger class of weakly dependent functional processes. More recently, Hörmann et al. [12] and Panaretros and Tavakoli [22] contribute to frequency domain methods of functional time series.

Solving the prediction equations in function spaces is problematic and research to-date has mainly considered first order autoregressive models. Contributions to functional prediction go hand in hand with an estimation method for the autoregressive parameter operator. Bosq [4] suggests a Yule–Walker type moment estimator, spline approximation is applied in

* Corresponding author.
E-mail addresses: j.klepsch@tum.de (J. Klepsch), cklu@tum.de (C. Klüppelberg).

http://dx.doi.org/10.1016/j.jmva.2017.01.005

Besse and Cardot [3], and Kargin and Onatski [17] propose a predictive factor method where the principal components are replaced by directions which may be more relevant for prediction.

When moving away from the autoregressive process, results on prediction of functional time series become sparse. An interesting theory for the prediction of general functional linear processes is developed in Bosq [6]. Necessary and sufficient conditions are derived for the best linear predictor to take the form $\phi_n(X_1, \ldots, X_n)$ with ϕ_n linear and bounded. However, due to the infinite-dimensionality of function spaces, boundedness of ϕ_n cannot be guaranteed. Consequently, most results, though interesting from a theoretical point of view, are not suitable for application.

More practical results are given for example in Antoniadis et al. [1], where prediction is performed non-parametrically with a functional kernel regression technique, or in Aue et al. [2], Hyndman and Shang [16] and Klepsch et al. [18], where the dimensionality of the prediction problem is reduced via functional principal component analysis. In a multivariate setting, the Innovations Algorithm proposed in Brockwell and Davis [8] gives a robust prediction method for linear processes. However, as often in functional data analysis, the non-invertibility of covariance operators prevents an ad-hoc generalization of the Innovations Algorithm to functional linear processes.

We suggest a computationally feasible linear prediction method extending the Innovations Algorithm to the functional setting. For a functional linear process $(X_n)_{n\in\mathbb{Z}}$ with values in a Hilbert space H and with innovation process $(\varepsilon_n)_{n\in\mathbb{Z}}$ our goal is the construction of a linear predictor $\widehat{X}_{n+1}$ based on $X_1, \ldots, X_n$ such that $\widehat{X}_{n+1}$ is both computationally tractable and *consistent*. In other words, we want to find a bounded linear mapping ϕ_n with $\widehat{X}_{n+1} = \phi_n(X_1, \ldots, X_n)$ such that the statistical prediction error converges to 0 for increasing sample size; i.e.,

$$\lim_{n\to\infty} \mathrm{E}\|X_{n+1} - \widehat{X}_{n+1}\|^2 = \mathrm{E}\|\varepsilon_0\|^2. \tag{1.1}$$

To achieve convergence in (1.1) we work with finite-dimensional projections of the functional process, similarly as in Aue et al. [2] and Klepsch et al. [18]. We start with a representation of the functional linear model in terms of an arbitrary orthonormal basis of the Hilbert space. We then focus on a representation of the model based on only finitely many basis functions. An intuitive choice for the orthonormal basis consists of the eigenfunctions of the covariance operator of the process. Taking the eigenfunctions corresponding to the D largest eigenvalues results in a truncated Karhunen–Loéve representation, and guarantees to capture most of the variance of the process (see [2]). Other applications may call for a different choice.

Though the idea of finite-dimensional projections is not new, our approach differs significantly from existing ones. Previous approaches consider the innovations of the projected process as the projection of the innovation of the original functional process. Though this may be sufficient in practice, it is in general not theoretically accurate.

The Wold decomposition enables us to work with the exact dynamics of the projected process, which then allows us to derive precise asymptotic results. The task set for this paper is of a purely predictive nature: we assume knowing the dependence structure and do not perform model selection or covariance estimation of the functional process. This will be the topic of a subsequent paper.

The truncated process $(X_{D,n})_{n\in\mathbb{Z}}$ based on D basis functions is called subprocess. We show that every subprocess of a stationary (and invertible) functional process is again stationary (and invertible). We then use an isometric isomorphy to a D-dimensional vector process to compute the best linear predictor of $(X_{D,n})_{n\in\mathbb{Z}}$ by the Multivariate Innovations Algorithm (see [8]).

As a special example we investigate the functional moving average process of finite order. We prove that every subprocess is again a functional moving average process of same order or less. Moreover, for this process the Innovations Algorithm simplifies. Invertibility is a natural assumption in the context of prediction (see Brockwell and Davis [8], Section 5.5, and Nsiri and Roy [21]), and we require it when proving limit results. The theoretical results on the structure of $(X_{D,n})_{n\in\mathbb{Z}}$ enable us to quantify the prediction error in (1.1). As expected, it can be decomposed in two terms, one due to the dimension reduction, and the other due to the statistical prediction error of the D-dimensional model. However, the goal of consistency as in (1.1) is not satisfied, as the error due to dimension reduction does not depend on the sample size.

Finally, in order to satisfy (1.1), we propose a modified version of the Innovations Algorithm. The idea is to increase D together with the sample size. Hence the iterations of our modified Innovations Algorithm are based on increasing subspaces. Here we focus on the eigenfunctions of the covariance operator of X as orthonormal basis of the function space.

Our main result states that the prediction error is a combination of two tail sums, one involving operators of the inverse representation of the process, and the other the eigenvalues of the covariance operator. We obtain a computationally tractable functional linear predictor for stationary invertible functional linear processes. As the sample size tends to infinity the predictor satisfies (1.1) with a rate depending on the eigenvalues of the covariance operator and of the spectral density operator.

Our paper is organized as follows. After summarizing prerequisites of functional time series in Section 2, we recall in Section 3 the framework of prediction in infinite-dimensional Hilbert spaces, mostly based on the work of Bosq (see [4–6]). Here we also clarify the difficulties of linear prediction in infinite-dimensional function spaces. In Section 4 we propose an Innovations Algorithm based on a finite-dimensional subprocess of X. The predictor proposed in Section 4, though quite general, does not satisfy (1.1). Hence, in Section 5 we project the process on a finite-dimensional subspace spanned by the eigenfunctions of the covariance operator of X, and formulate the prediction problem in such a way that the dimension of the subprocess increases with the sample size. A modification of the Innovations Algorithm then yields a predictor which