Statistics and Probability Letters 130 (2017) 42–48

Contents lists available at ScienceDirect

Statistics and Probability Letters

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

Asymptotic properties of principal component projections with repeated eigenvalues

Justin Petrovich *, Matthew Reimherr

Pennsylvania State University, United States

ARTICLE INFO

Article history:
Received 15 February 2017
Received in revised form 6 July 2017
Accepted 12 July 2017
Available online 27 July 2017

Keywords:
Functional data analysis
Repeated eigenvalues
Principal components
Projection

ABSTRACT

In FPCA methods, it is common to assume that the eigenvalues are distinct in order to facilitate theoretical proofs. We relax this assumption, provide a stochastic expansion for the estimated functional principal component projections, and establish their asymptotic normality.

1. Introduction

Functional data analysis (FDA) is an increasingly popular realm of statistics that is receiving much attention and research. With a multitude of data applications as diverse as finance, genomics, and medicine, it is clear why FDA has drawn so much attention from the statistics and, more broadly, the scientific community. One of the most challenging features of functional data methods is the assumption of infinite-dimensional data or parameters, which generally forces practitioners of FDA to employ some method of dimension reduction so that multivariate data tools can be applied. One of the most important dimension reduction techniques is functional principal component analysis (FPCA).

In FPCA, the data are represented in terms of the eigenfunctions of the covariance operator, the functional principal components (FPCs). When estimating these FPCs, it is typically assumed that the associated eigenvalues are distinct. However, recent work Reimherr (2015) has shown that in the context of functional regression, by estimating the entire projection of FPCs, rather than estimating eigenfunctions one at a time, one only needs distinctness of the last eigenvalue included and the first eigenvalue excluded from the projection.

Readers unfamiliar with FDA are recommended to read the definitive works of Ramsay and Silverman (2005) and Bosq (1991), though many other useful overviews of the field exist (Bosq, 2000; Kokoszka and Reimherr, 2017; Hsing and Eubank, 2015; Cuevas, 2014; Goia and Vieu, 2016). For literature more specific to FPCA, we refer readers first to the accessible surveys provided by Shang (2014) and Tran (2008), while the early mathematical framework for the asymptotics of FPCA was established by Dauxois et al. (1982). Another major contribution was (Hall and Hosseini-Nasab, 2006), which provided explicit stochastic expansions of estimated eigenvalues and eigenfunctions, but relied on the distinctness of the eigenvalues. Other relevant works which expand the theory for or use of FPCA include (Kokoszka and Reimherr, 2013; Hörmann and Kidziński, 2015a; Jirak, 2016; Hall et al., 2006; Mas and Ruymgaart, 2015). However, each of these works relies on the assumption of distinct eigenvalues.

* Corresponding author.
E-mail address: jpetrovich@psu.edu (J. Petrovich).

http://dx.doi.org/10.1016/j.spl.2017.07.004
0167-7152/

Particularly relevant to our work is that of Hörmann and Kidziński (2015b) and Reimherr (2015), each of which relaxes the assumption of distinct eigenvalues in working with FPCA. In Hörmann and Kidziński (2015b) it is shown that when using functional regression for prediction, one does not need distinct eigenvalues to guarantee consistency of the prediction. However, in order to establish consistent estimation, they still lean on this assumption. This is taken one step further in Reimherr (2015), which shows that by estimating the entire principal component projection simultaneously rather than each eigenvalue/function one at a time, what is necessary to establish consistent estimators in FPCA-based regression is the spread between the last eigenvalue included and the first eigenvalue excluded, rather than the distinctness of all the eigenvalues. Operating from a similar perspective, we establish a stochastic expansion of the FPC projections, which in turn can be used to prove several forms of asymptotic normality of the projections.

2. Background

Let X_i for $i = 1, \ldots, N$ be a sequence of N identically distributed, square integrable random functions in a real separable Hilbert space $\mathcal{H}$. We define $\langle \cdot, \cdot \rangle$ to be the inner product and $\langle \cdot, \cdot \rangle = \| \cdot \|^2$ to be the norm on $\mathcal{H}$. The covariance operator of X_i is defined as $C(\cdot) = [\langle X_i - EX_i, \cdot \rangle (X_i - EX_i)] \in \mathcal{S}$, where $\mathcal{S}$ represents the space of Hilbert–Schmidt operators. As a Hilbert–Schmidt operator, for any orthonormal basis $\{e_j, j \geq 1\}$ of $\mathcal{H}$, the covariance operator C satisfies $\|C\|_{\mathcal{S}}^2 = \sum_{j=1}^{\infty} \|C(e_j)\|^2 < \infty$. Here, $\| \cdot \|_{\mathcal{S}}^2$ represents the Hilbert–Schmidt norm, and is invariant with respect to the choice of basis. We define $\hat{C}$ to be the sample covariance operator, which is also a Hilbert–Schmidt operator.

We now introduce some terminology and notation used in FPCA. Let $\{v_j, j \geq 1\}$ be the eigenfunctions of C, which form an orthonormal basis in $\mathcal{H}$, and $\{\lambda_j, j \geq 1\}$ the corresponding eigenvalues, which are both nonincreasing and nonnegative. By the Spectral Theorem, we can express $C(x) = \sum_{j=1}^{\infty} \lambda_j \langle x, v_j \rangle v_j$. Similarly, $\{\hat{v}_j, j \geq 1\}$ and $\{\hat{\lambda}_j, j \geq 1\}$ are the eigenfunctions and eigenvalues of $\hat{C}$. These eigenfunctions form the basis by which we expand our random function X_i; we project X_i onto the subspace spanned by $\{v_j, j \leq J\}$ for some (small) finite number of eigenfunctions, J. This is done using the Karhunen–Loéve expansion as follows:

$$X_i - EX_i = \sum_{j=1}^{\infty} \xi_j v_j,$$

where the $\xi_j = \langle X_i - EX_i, v_j \rangle$ are called the scores. FPCA is often used in FDA, in regression for instance, where the scores are used to estimate the functional regression parameters.

Define the principal component projection to be $P_J = \sum_{j=1}^{J} v_j \otimes v_j$, and its estimate to be $\hat{P}_J = \sum_{j=1}^{J} \hat{v}_j \otimes \hat{v}_j$. Note that $x \otimes y$ represents the tensor product between the functions x and y,which we view as an operator $(x \otimes y)(h) := \langle y, h \rangle x$. Thus the principal component projection can be applied to an object x to get $P_J(x) = \sum_{j=1}^{J} \langle v_j, x \rangle v_j$. These projections will be a focal point of our proofs as their use allows for the simultaneous estimation of all the eigenfunctions, v_j. Since estimating each eigenfunction individually requires uniqueness among the eigenvalues, switching to the estimation of projections is the critical step that allows us to weaken this assumption, requiring only that the Jth and $(J+1)$th eigenvalues be unique.

3. Theoretical results

To establish our results, we assume the sample covariance operator, $\hat{C}$, satisfies the following.

Assumption 3.1. Let $Z_N = \sqrt{N}(\hat{C} - C)$ and assume that $\|Z_N\| = O_p(1)$.

Assumption 3.2. Assume there is a mean zero Gaussian Hilbert–Schmidt operator, Z, such that $Z_N \xrightarrow{d} Z$ as $N \to \infty$.

Assumptions 3.1 and 3.2 offer different degrees of control of $\hat{C}$. The former, which ensures that $\|\hat{C} - C\|$ is of order $N^{-1/2}$, is all that is required for our stochastic expansion. Meanwhile, Assumption 3.2 is a slightly stronger version of 3.1, guaranteeing asymptotic normality of Z_N which will be necessary for the asymptotic results we provide in our corollaries. Such assumptions are often satisfied in practice; for example, given independent and identically distributed (i.i.d.) data, or given a weakly dependent sequence—where weakly dependent can be defined, for example, via the concept of $L^4 - m$-approximability as in Kokoszka and Reimherr (2013)—we can guarantee the asymptotic normality of $\hat{C}$. Central limit theorems in Hilbert spaces have also been established under a variety of other conditions (Merlevéde, 2003; Merlevéde et al., 1997; Tone, 2011; Chen and White, 1998; Lavrentyev and Nazarov, 2016). With these assumptions in mind, we proceed to the main result.

Theorem 3.1. *Let Assumption 3.1 hold and define the quantity*

$$R_{J,N} = \sum_{j=1}^{J} \sum_{i>J} \frac{\langle Z_N, v_j \otimes v_i \rangle}{(\lambda_j - \lambda_i)} (v_j \otimes v_i + v_i \otimes v_j).$$