Asymptotic Properties of Principal Component Projections with Repeated Eigenvalues

Justin Petrovich and Matthew Reimherr
Pennsylvania State University, Statistics Department

Introduction

One of the most challenging features of functional data methods is the assumption of infinite-dimensional data, which generally forces practitioners to employ some method of dimension reduction in order to apply multivariate data tools. 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, also known as the functional principal components (FPCs). When estimating these FPCs, it is typically assumed that the associated eigenvalues must be distinct[1]. However, this assumption turns out to be unnecessarily restrictive; instead, what matters is the uniqueness of the first eigenvalue included and last eigenvalue excluded from the dimension reduction.

Notation

Let $\{X_1, \cdots, X_N\}$ be a sequence of identically distributed square integrable random functions in a real separable Hilbert space $\mathcal{H}$. We define $\langle \cdot, \cdot \rangle$ and $|| \cdot ||$ to be the inner product and 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, and $\hat{C}$ is the sample covariance operator, which is also a Hilbert-Schmidt operator. Let $\{v_j, j \geq 1\}$ be the eigenfunctions of C and $\{\lambda_j, j \geq 1\}$ the corresponding eigenvalues.Define the principal component projection to be

$$P_J = \sum_{j=1}^{J} v_j \otimes v_j$$

and its estimate as

$$\hat{P}_J = \sum_{j=1}^{J} \hat{v}_j \otimes \hat{v}_j.$$

Here we take $x \otimes y$, the tensor product between the functions x and y, as an operator, such that $(x \otimes y)(h) := \langle y, h \rangle x$. 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 are crucial for our proofs because 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 the assumption, requiring only that the J^{th} and $(J+1)^{th}$ eigenvalues be unique.

Main Theorem

Assumption 1: Let $Z_N = \sqrt{N}(\hat{C} - C)$, and assume there is a mean zero Gaussian Hilbert-Schmidt operator, Z, such that $Z_N \xrightarrow{d} Z$.

Theorem 1

Define the quantity

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

If Assumption 1 holds and as $N \to \infty$, J is such that

$$N^{1/2}(\lambda_J - \lambda_{J+1}) \to \infty,$$

then

$$||\sqrt{N}(\hat{P}_J - P_J) - R_{J,N}||_{\mathcal{S}}^2 = \left[\frac{||\hat{C} - C||^2}{(\lambda_J - \lambda_{J+1})^2} + ||\hat{P}_J - P_J||^2 \right] O_p\left((\lambda_J - \lambda_{J+1})^{-2}\right).$$

Corollaries

Assumption 2: As $N \to \infty$, we assume that

$$\frac{J}{N\lambda_J^2(\lambda_J - \lambda_{J+1})^4} \to 0.$$

Corollary 1: If Assumption 1 holds, J is fixed, and $\lambda_J \neq \lambda_{J+1}$, then

$$\sqrt{N}(\hat{P}_J - P_J) \xrightarrow{\mathcal{D}} \mathcal{N}_{\mathcal{S}}(0, \Gamma_J),$$

where Γ_J is a covariance operator acting on, $\mathcal{S}$, the space of Hilbert-Schmidt operators and is given by

$$\Gamma_J = \sum_{j=1}^{J} \sum_{i>J} \sum_{j'=1}^{J} \sum_{i'>J} \frac{\langle \mathrm{E}[Z_N \otimes Z_N], v_j \otimes v_i \otimes v_{j'} \otimes v_{i'} \rangle}{(\lambda_j - \lambda_i)(\lambda_{j'} - \lambda_{i'})} \times (v_j \otimes v_i + v_i \otimes v_j) \otimes (v_{j'} \otimes v_{i'} + v_{i'} \otimes v_{j'}).$$

Corollary 2: Under Assumptions 1 and 1, if we fix x and y in $\mathcal{H}$, then

$$\frac{\langle \sqrt{N}(\hat{P}_J - P_J)(x), y \rangle}{\langle \Gamma_J(x \otimes y), x \otimes y \rangle^{1/2}} \xrightarrow{\mathcal{D}} \mathcal{N}(0, 1).$$

Corollary 3: If Assumptions 1 and 1 hold and there exists a sequence of Gaussian processes, $\{Y_N : N = 1, 2, \dots\}$, such that

$$\|Z_N - Y_N\|_{\mathcal{S}} = O_P(N^{-\delta}),$$

then there exist Gaussian processes $\tilde{R}_{J,N}$ in $\mathcal{S}$ such that

$$\|\sqrt{N}(\hat{P}_J - P_J) - \tilde{R}_{J,N}\|_{\mathcal{S}} = O_P(2(\lambda_J - \lambda_{J+1})^{-2} N^{-2\delta}),$$

where $\tilde{R}_{J,N}$ is mean zero with covariance operator Γ_J.

Some Applications

These results impact all aspects of FDA in which FPCA is used, including but not limited to[2]:

- Functional Linear Regression[3]
- Functional GLMs
- Multilevel Functional Models
- Hypothesis Testing for Functional Linear Models[4]
- Functional Classification
- Functional Cluster analysis[5]

Conclusion

These results shed further light on the role of the eigenvalues in establishing theory for FPCs and their projections. By focusing on estimating the FPC projections, we find estimates under milder assumptions, and use this result to confirm the asymptotic normality of the FPC projections under various circumstances. We acknowledge that the rates provided in Corollary 3 can potentially be improved. Since achieving optimal convergence rates was not our focus, we leave this for future work. Indeed, it would be of great use to establish whether our method of utilizing the FPC projections can lead to minimax rates of convergence, say, in the instance of functional linear regression, or if the convergence rate presented in Theorem 1 can be tightened.

References

[1] Hall P. and Hosseini-Nasab M.
On Properties of Functional Principal Components Analysis.
Journal of the Royal Statistical Society, Series B, 68(1):109–126, 2006.

[2] Shang H. L.
A Survey of Functional Principal Component Analysis.
Advances in Statistical Analysis, 98(2):121–142, 2014.

[3] Reimherr M.
Functional Regression with Repeated Eigenvalues.
Statistics and Probability Letters, 107:62–70, 2015.

[4] A. Staicu D. Kong and A. Maity.
Classical Testing in Functional Linear Models.
Journal of Nonparametric Statistics, 28(4):813–838, 2016.

[5] J. Jacques and C. Preda.
Functional Data Clustering: A Survey.
Advances in Data Analysis and Classification, 8(3):231–255, 2014.