Minimax estimation for FPCA on discretized data

Nassim Bourarach[1] & Franck Picard[2] & Vincent Rivoirard [1] & Angelina Roche [1]

Given p, n two positive integers, we consider p noisy evaluations of n realizations of random functions on a common design $(t_j)_{j=1}^{p} \in [0,1]$:

$$Y_i(t_j) := X_i(t_j) + \varepsilon_{i,j}, \qquad (i,j) \in [\![1,n]\!] \times [\![1,p]\!].$$

where $\varepsilon_{i,j} \overset{i.i.d.}{\sim} \mathcal{N}(0, \sigma^2)$ with $\sigma^2 > 0$. The $\varepsilon_{i,j}$'s are independent of the random functions X_i which are also i.i.d. and defined on $[0,1]$.

We are interested in the estimation of $(\psi_\ell^*, \lambda_\ell^*)$, respectively the ℓ-th eigenfunction and the ℓ-th eigenvalue of the covariance integral operator associated to the covariance kernel $\mathbb{E}[X_1(s)X_1(t)]$ for $(s,t) \in [0,1]^2$. Indeed, by virtue of Mercer's theorem and the Karhunen-Loève decomposition these eigenelements are crucial to FPCA (see Ramsay and Silverman (2005) or Hsing and Eubank (2013) for a more thorough account of FPCA).

Our first contributions are non-asymptotic minimax lower bounds for the estimation of these eigenelements when the covariance kernel is m-Hölder regular (for all $m \in \mathbb{R}_+^*$) and when the spectrum of the covariance integral operator associated with the kernel obey some constraints.
These new bounds generalize and complement the first bound obtained by Belhakem et al. (2021). The class of processes used for the minimax study allows us to analyze the impact of the spectrum of the covariance operator on the estimation rates. Analogous quantities to that of "relative rank" (which can be found in Jirak and Wahl (2022) and Mas and Ruymgaart (2015)) come into play; we also obtain inconsistency results if the constraints are not satisfied.

Then, we present simple estimators of the eigenelements, based on a projection onto a wavelet basis. The obtained estimators are minimax optimal under a few additional assumptions, and attain rates (in n and p) of the form $n^{-1} + p^{-2m}$.

Surprisingly enough, even if the problem is non-parametric in nature, there is actually no need for data smoothing.

References

Belhakem, R., Picard, F., Rivoirard, V. and Roche, A. (2021) Minimax estimation of functional principal components from noisy discretized functional data.

Hsing, T. and Eubank, R. (2013) *Theoretical Foundations of Functional Data Analysis, with an Introduction to Linear Operators.*

Jirak, M. and Wahl, M. (2022) Relative perturbation bounds with applications to empirical covariance operators.

Mas, A. and Ruymgaart, F. (2015) High Dimensional Principal Projections. *Complex Analysis and Operator Theory*, **9**, 35 – 63. URL: https://hal.science/hal-00772880.

Ramsay, J. O. and Silverman, B. W. (2005) *Functional Data Analysis, 2nd Edition.* Springer.

[1]CEREMADE, Université Paris-Dauphine PSL

[2]LBMC, ENS de Lyon