Functional common principal components models

Graciela Boente, Daniela Rodriguez, Mariela Sued

Functional common principal components models

Graciela Boente[1] , Daniela Rodriguez[1] and Mariela Sued[1]
[1] *Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires and CONICET, Argentina*

Instituto de Cálculo, Facultad de Ciencias Exactas y Naturales, Ciudad Universitaria, Pabellón 2, Buenos Aires, C1428EHA, Argentina.

Abstract

In this paper, we discuss the extension to the functional setting of the common principal component model that has been widely studied when dealing with multivariate observations. We provide estimators of the common eigenfunctions and study their asymptotic behavior.

Résumé

Dans cet exposé, nous discutons l'extension au cas fonctionnel du modèle de composantes principales communes, qui a été largement étudié lorsqu'on s'intéresse à des observations multivariées. Nous proposons des estimateurs pour les composantes principales communes et nous étudions leur distribution asymptotique.

Some key words: Common principal components, Functional data analysis.

1. Introduction

Functional data analysis is an emerging field in statistics that has received considerable attention during the last decade due to its applications to many other different areas. It provides modern data analytical tools for data that are recoded as a continuous phenomenon over a period of time. Because of the intrinsic nature of these data, they can be viewed as realizations of random functions $X_1(t), \ldots, X_n(t)$ often assumed to be in $L^2([0, 1])$. In this context, principal components analysis offers an effective way for dimension reduction and it has been extended from the traditional multivariate setting to accommodate functional data. In the functional data analysis literature, it is usually referred to as functional principal component analysis (FPCA).

In many situations, we have independent observations $X_{i,1}(t), \cdots, X_{i,n_i}(t)$ from k independent samples of random functions in $L^2[0, 1]$ with mean μ_i and covariance operators $\mathbf{\Gamma}_i$. As it is the case in the finite–dimensional setting, the covariance operators may exhibit some common structure. The common principal components model, introduced by Flury [**?**] for $p-$th dimensional data, allow the covariance matrices to have different eigenvalues but identical eigenvectors. A natural extension to the functional setting of the common principal components model is to assume that the covariance operators $\mathbf{\Gamma}_i$ have common eigenfunctions $\phi_j(t)$ but different eigenvalues λ_{ij}. We will denote this model the functional common principal component (FCPC) model.

The aim of this talk is to provide estimators of the common eigenfunctions and the eigenvalues under a FCPC model and to study their asymptotic behavior. Proofs are given by Boente, Rodriguez and Sued [**?**].

2. Notation and Preliminaries

Let $X_{i,1}(t), \cdots, X_{i,n_i}(t)$, $1 \le i \le k$, be independent observations from k independent samples of smooth random functions in $L^2[\mathcal{I}]$, where $\mathcal{I} = [0, 1]$, with mean μ_i. Denote by γ_i and $\mathbf{\Gamma}_i$ the covariance function and operator, respectively, related to each population. To be more precise, we are assuming that $\{X_{i,1}(t) : t \in \mathcal{I}\}$ are k stochastic processes defined in $(\Omega, \mathcal{A}, P)$ with continuous trajectories, mean μ_i and finite second moment, i.e., $E\left(X_{i,1}(t)\right) = \mu_i(t)$ and $E\left(X_{i,1}^2(t)\right) < \infty$ for $t \in \mathcal{I}$. Each covariance function $\gamma_i(t, s) = \text{COV}(X_{i,1}(s), X_{i,1}(t))$, $s, t \in \mathcal{I}$ has an associated linear operator $\mathbf{\Gamma}_i : L^2[0, 1] \to L^2[0, 1]$ defined as $\left(\mathbf{\Gamma}_i\, u\right)(t) = \int_0^1 \gamma_i(t, s)u(s)ds$, for all $u \in L^2[0, 1]$. As in the case of one population, throughout this paper, we will assume that the covariance functions satisfy $\|\gamma_i\|^2 = \int_0^1 \int_0^1 \gamma_i^2(t, s)dtds < \infty$. Therefore, $\mathbf{\Gamma}_i$ is a self–adjoint continuous linear operator. Moreover, $\mathbf{\Gamma}_i$ is a Hilbert-Schmidt operator. The FCPC model assume that the covariance operators $\mathbf{\Gamma}_i$ have common eigenfunctions $\{\phi_j(t) : j \ge 1\}$, to be estimated, as well as the eigenvalues associated to each covariance operator $\mathbf{\Gamma}_i$, $1 \le i \le k$.

When dealing with one population, estimators of the eigenfunctions and eigenvalues of $\mathbf{\Gamma}$ were defined, in a natural way, through the empirical covariance operator by Dauxois, Pousse and Romain [**?**]. In the present setting, we will give two proposals to estimate the common eigenfunctions under a FCPC model. Both of them are based on estimators $\widehat{\mathbf{\Gamma}}_i$ of the covariance operators $\mathbf{\Gamma}_i$, like $\widehat{\mathbf{\Gamma}}_{i,\mathrm{R}}$, the operator associated to the empirical covariance functions $\widehat{\gamma}_{i,\mathrm{R}}(s,t) = \frac{1}{n_i}\sum_{j=1}^{n_i}\left(X_{i,j}(s) - \overline{X}_i(s)\right)\left(X_{i,j}(t) - \overline{X}_i(t)\right)$.

Assume $n_i = \tau_i N$ with $0 < \tau_i < 1$ fixed numbers such that $\sum_{i=1}^k \tau_i = 1$ and where $N = \sum_{i=1}^k n_i$ denotes the total number of observations in the sample. Define the weighted covariance function as $\gamma = \sum_{i=1}^k \tau_i \gamma_i$ and its related operator as $\mathbf{\Gamma} = \sum_{i=1}^k \tau_i \mathbf{\Gamma}_i$. Therefore, $\widehat{\gamma}_\mathrm{R} = \sum_{i=1}^k \tau_i \widehat{\gamma}_{i,\mathrm{R}}$ and $\widehat{\mathbf{\Gamma}}_\mathrm{R} = \sum_{i=1}^k \tau_i \widehat{\mathbf{\Gamma}}_{i,\mathrm{R}}$ provide estimators of γ and $\mathbf{\Gamma}$, respectively. It is worth noticing that our results do not make use of the explicit expression of the covariance operator estimators, but they only require their consistency and asymptotic normality.

3. The proposals

Let us assume that the FCPC model hold, i.e., $\mathbf{\Gamma}_i$ have common eigenfunctions $\phi_j(t)$ but possible different eigenvalues λ_{ij}, where $\lambda_{ij} = \langle \phi_j, \mathbf{\Gamma}_i \phi_j \rangle$. Moreover, throughout this paper we will assume that

A1. $\lambda_{i1} \geq \lambda_{i2} \geq \cdots \geq \lambda_{ip} \geq \lambda_{ip+1} \cdots$, for $1 \leq i \leq k$

A2. There exists ℓ such that for any $1 \leq j \leq \ell$, there exists $1 \leq i \leq k$ such that $\lambda_{ij} > \lambda_{i\,j+1}$.

The first proposal is based on the fact that under the FCPC model, the common eigenfunctions $\{\phi_j : j \geq 1\}$ are also a basis of eigenfunctions for the operator $\mathbf{\Gamma} = \sum_{i=1}^k \tau_i \mathbf{\Gamma}_i$, with eigenvalues given by $\nu_1 = \sum_{i=1}^k \tau_i \lambda_{i1} \geq \cdots \geq \nu_p = \sum_{i=1}^k \tau_i \lambda_{ip} \geq \nu_{p+1} = \sum_{i=1}^k \tau_i \lambda_{i\,p+1} \cdots$. Note that **A1** and **A2** entail that the first ℓ eigenfunctions will be related to the ℓ largest eigenvalues of the operator $\mathbf{\Gamma}$, having multiplicity one and being strictly positive. A first attempt to estimate the common eigenfunctions consists in considering the eigenfunctions $\widetilde{\phi}_j$ related to the largest eigenvalues $\widehat{\nu}_j$ of a consistent estimator $\widehat{\mathbf{\Gamma}}$ of $\mathbf{\Gamma}$, obtained as $\widehat{\mathbf{\Gamma}} = \sum_{i=1}^k \tau_i \widehat{\mathbf{\Gamma}}_i$ where $\widehat{\mathbf{\Gamma}}_i$ denotes any estimator of the $i-$th covariance operator. An example of such estimators are those associated to the empirical covariance functions $\widehat{\gamma}_{i,\mathrm{R}}$. The eigenvalue estimators can then be defined as $\widehat{\lambda}_{ij} = \langle \widetilde{\phi}_j, \widehat{\mathbf{\Gamma}}_i \widetilde{\phi}_j \rangle$.

The second proposal tries to improve the efficiency of the previous one for gaussian processes. To that purpose, we will have in mind that, in the finite–dimensional case, the maximum likelihood estimators of the common directions for normal data solve a system of equations involving both the eigenvalue and eigenvector estimators (see Flury, [**?**]). Using consistent estimators of the eigenvalues, we generalize the system obtained by Flury to the infinite–dimensional case. Effectively, let $\widehat{\lambda}_{ij}$ be initial estimators of the eigenvalues and $\widehat{\mathbf{\Gamma}}_i$ any consistent estimator of the covariance operator of the $i-$th population. Define for $j \le \ell$ and $m \le \ell$, $\widehat{\mathbf{\Gamma}}_{mj} = \sum_{i=1}^k \tau_i \frac{\widehat{\lambda}_{ij} - \widehat{\lambda}_{im}}{\widehat{\lambda}_{im}\widehat{\lambda}_{ij}} \widehat{\mathbf{\Gamma}}_i$, which will be asymptotically well defined under **A2** if in addition $\lambda_{i\ell} > 0$ for $1 \le i \le k$. The second proposal considers the solution $\widehat{\phi}_j$ of the system of equations

$$\begin{cases} \delta_{mj} = & \langle \widehat{\phi}_m, \widehat{\phi}_j \rangle \\ 0 = & \langle \widehat{\phi}_m, \widehat{\mathbf{\Gamma}}_{mj} \widehat{\phi}_j \rangle \qquad 1 \le j < m \;. \end{cases} \tag{1}$$

4. Asymptotic distribution

It is clear that consistency of each population covariance operator estimator ensures consistency of the pooled one. The results in Section 2.1 of Dauxois, Pousse and Romain [**?**], allow to obtain the asymptotic distribution of the estimators of the common eigenfunctions when considering the first proposal using the sample covariance operators. In particular, we obtain the following result (see, Boente, Rodriguez and Sued, [**?**], for details).

Proposition 4.1. *Let us assume that $\widehat{\mathbf{\Gamma}}_i$ is the empirical operator $\widehat{\mathbf{\Gamma}}_{i,\mathrm{R}}$, that $E(\|X_{i,1}\|^4) < \infty$, for $1 \le i \le k$, and that **A1** and **A2** hold. For each eigenfunction ϕ_j of $\mathbf{\Gamma}$ related to the eigenvalue $\nu_j = \sum_{i=1}^k \tau_i \lambda_{ij}$ with multiplicity one, we have that*

a) $\sqrt{N}(\widetilde{\phi}_j - \phi_j, \phi_j) \xrightarrow{p} 0$

b) For any $j \ne m$ $\sqrt{N}\langle \widetilde{\phi}_j - \phi_j, \phi_m \rangle \;\to\; \mathcal{N}(0, \sigma^2_{mj})$ *with*

$$\sigma^2_{jm} \;=\; \left\{ \sum_{i=1}^k \tau_i (\lambda_{ij} - \lambda_{im}) \right\}^{-2} \sum_{i=1}^k \tau_i \lambda_{im} \lambda_{ij} E[f^2_{im} f^2_{ij}]$$

Moreover, if $X_{i,1}$ are gaussian processes, for all $1 \le i \le k$, we get that

$$\sigma^2_{jm} \;=\; \left\{ \sum_{i=1}^k \tau_i (\lambda_{ij} - \lambda_{im}) \right\}^{-2} \sum_{i=1}^k \tau_i \lambda_{im} \lambda_{ij} \;. \tag{2}$$

The following Theorem provides the asymptotic behavior of the eigenvalue estimators under mild conditions on the eigenfunction estimators. It can be used to derive the asymptotic normality of the eigenvalue estimators when using, either Proposal 1 or Proposal 2 to estimate the eigenfunctions.

Theorem 4.1. *Let $\widehat{\mathbf{\Gamma}}_i$ be an estimator of the covariance operator of the $i-$th population such that $\sqrt{n_i}(\widehat{\mathbf{\Gamma}}_i - \mathbf{\Gamma}_i) \xrightarrow{\mathcal{D}} \mathbf{U}_i$, where $\mathbf{U}_i$ is zero mean gaussian random element with* covariance operator $\boldsymbol{\Upsilon}_i$. Let $\widetilde{\phi}_j$ be consistent estimators of the common eigenfunctions such that $\sqrt{N}\left(\widetilde{\phi}_j-\phi_j\right)=O_p(1)$ and define estimators of λ_{ij} as $\widehat{\lambda}_{ij}=\langle\widetilde{\phi}_j,\widehat{\boldsymbol{\Gamma}}_i\widetilde{\phi}_j\rangle$. For any fixed m, denote $\widehat{\boldsymbol{\Lambda}}_i^{(m)}=\left\{\sqrt{n_i}\left(\widehat{\lambda}_{ij}-\lambda_{ij}\right)\right\}_{1\le j\le m}$. Then,

a) For each $1\le i\le k$, $\sqrt{n_i}\left(\widehat{\lambda}_{ij}-\lambda_{ij}\right)$ has the same asymptotic distribution as $\sqrt{n_i}\left(\langle\phi_j,\widehat{\boldsymbol{\Gamma}}_i\phi_j\rangle-\lambda_{ij}\right)$.

b) For any m fixed, $\widehat{\boldsymbol{\Lambda}}_1^{(m)},\dots,\widehat{\boldsymbol{\Lambda}}_k^{(m)}$ are asymptotically independent.

c) If, in addition, the covariance operator $\boldsymbol{\Upsilon}_i$ of $\mathbf{U}_i$ is given by

$$\boldsymbol{\Upsilon}_i=\sum_{m,r,o,p}s_{im}s_{ir}s_{io}s_{ip}E[f_{im}f_{ir}f_{io}f_{ip}]\,\phi_m\otimes\phi_r\tilde{\otimes}\phi_o\otimes\phi_p-\sum_{m,r}\lambda_{im}\lambda_{ir}\,\phi_m\otimes\phi_m\tilde{\otimes}\phi_r\otimes\phi_r$$

then, $\widehat{\boldsymbol{\Lambda}}_i^{(m)}$ is jointly asymptotically normally distributed with zero mean and covariance matrix $\mathbf{C}^{(i,m)}$ such that $\mathbf{C}_{jj}^{(i,m)}=\lambda_{ij}^2\left[E\left(f_{ij}^4\right)-1\right]$ and $\mathbf{C}_{js}^{(i,m)}=\lambda_{ij}\lambda_{is}\left[E\left(f_{ij}^2f_{is}^2\right)-1\right]$, that is, the asymptotic variance of $\sqrt{n_i}\left(\widehat{\lambda}_{ij}-\lambda_{ij}\right)$ is given by $\lambda_{ij}^2\left[E\left(f_{ij}^4\right)-1\right]$ and the asymptotic correlations are given by

$$\frac{E\left(f_{ij}^2f_{is}^2\right)-1}{\left[E\left(f_{ij}^4\right)-1\right]^{\frac{1}{2}}\left[E\left(f_{is}^4\right)-1\right]^{\frac{1}{2}}}\,.$$

Moreover, in the normal case, we get that the components of $\widehat{\boldsymbol{\Lambda}}_i^{(m)}$ are asymptotically independent with asymptotic variances $2\lambda_{ij}^2$.

Finally, in order to study the asymptotic behavior of the second proposal, let $\boldsymbol{\Gamma}_{mj}=\sum_{i=1}^k\tau_i\left[(\lambda_{ij}-\lambda_{im})/(\lambda_{im}\lambda_{ij})\right]\boldsymbol{\Gamma}_i$ and denote $\phi_j^\star$ any solution of

$$\left\{\begin{array}{rl}\delta_{mj}= & \langle\phi_m^\star,\phi_j^\star\rangle\\ 0= & \langle\phi_m^\star,\boldsymbol{\Gamma}_{mj}\phi_j^\star\rangle\qquad 1\le j<m\,.\end{array}\right.\tag{3}$$

It is easy to see that if the covariance operators satisfy a FCPC model, then ϕ_j satisfies (**??**). Moreover, in Boente, Rodriguez and Sued [**?**] the consistency of the estimators defined through (**??**) is derived under mild conditions. The following result states the asymptotic behavior of the coordinates $\{\langle\widehat{\phi}_j,\phi_s\rangle : s\ge 1\}$ of the common eigenfunctions estimators $\widehat{\phi}_j$ defined through Proposal 2 that will allow to establish an improvement in efficiency for gaussian processes.

Theorem 4.1. *Let $\widehat{\boldsymbol{\Gamma}}_i$ be an estimator of the covariance operator of the $i-th$ satisfying the same assmptions as in Theorem 4.1. Let $\widehat{\lambda}_{ij}$ be consistent estimators of the eigenvalues* *of the $i-$th population λ_{ij} and $\widehat{\phi}_j$ consistent estimators of the common eigenfunctions ϕ_j, solution of (**??**) and denote $\widehat{g}_j = \sqrt{N}\left(\widehat{\phi}_j - \phi_j\right)$. Assume **A1**, **A2** and that $\lambda_{i\ell} > 0$, for all $1 \le i \le k$. If, in addition, for any $j \le \ell$, $m \le \ell$, the following two conditions hold*

i) $\langle \widehat{g}_j, \widehat{\phi}_m - \phi_m \rangle = o_p(1)$

ii) the operators $\mathbf{\Gamma}_i$ have finite rank ℓ, for all $1 \le i \le k$, or $\langle \widehat{g}_j, \mathbf{\Gamma}_i \left(\widehat{\phi}_m - \phi_m\right) \rangle = o_p(1)$.

then, for any $j \le \ell$, $m \le \ell$, $m \ne j$ we have that

a) $\langle \widehat{g}_m, \phi_j \rangle$ *has the same asymptotic distribution as* $-\langle \widehat{g}_j, \phi_m \rangle$.

b) For $j < m$, $\langle \widehat{g}_j, \phi_m \rangle \xrightarrow{\mathcal{D}} \mathcal{N}(0, \theta^2_{jm})$, *where*

$$\theta^2_{jm} = \frac{\sum_{i=1}^{k} \tau_i \frac{(\lambda_{im} - \lambda_{ij})^2}{\lambda_{im}\lambda_{ij}} E\left(f^2_{im} f^2_{ij}\right)}{\left\{\sum_{i=1}^{k} \tau_i \frac{(\lambda_{im} - \lambda_{ij})^2}{\lambda_{im}\lambda_{ij}}\right\}^2}. \tag{4}$$

Remark 4.1. Note that in the gaussian case, we get $E\left(f^2_{im} f^2_{ij}\right) = 1$ and so the asymptotic variance of coordinates of the common eigenfunction estimates, defined through Proposal 2, reduces to

$$\theta^2_{jm} = \left\{\sum_{i=1}^{k} \tau_i \frac{(\lambda_{im} - \lambda_{ij})^2}{\lambda_{im}\lambda_{ij}}\right\}^{-1}$$

On the other hand, the common eigenfunction estimates, defined through Proposal 1, have asymptotic variances σ^2_{jm} given by (**??**). Since $\theta^2_{jm} \le \sigma^2_{jm}$, we obtain that the estimates of Proposal 2 are more efficient that those of Proposal 1 for gaussian processes.

References

[1] Boente, G.; Rodriguez, D. and Sued, M. (2010). Inference under functional proportional and common principal components models. *J. Multivar. Anal.*, **101**, 464-475.

[2] Dauxois, J., Pousse, A. and Romain, Y. (1982). Asymptotic theory for the principal component analysis of a vector random function: Some applications to statistical inference. *J. Multivar. Anal.*, **12**, 136-154.

[3] Flury, B. K. (1984). Common principal components in k groups. *J. Am. Statist. Assoc.* **79,** 892-8.