Supplementary material to "A geometric interpretation of the multivariate functional principal components analysis"

Steven Golovkine[*] Edward Gunning[†]

March 21, 2023

Abstract

Your abstract.

1 Univariate case

For now, we will consider the univariate case ($P = 1$), and estimate the principal components. Assume the existence of a continuous covariance function

$$C(s, t) = \mathbb{E}(\{X(s) - \mu(s)\}\{X(t) - \mu(t)\}), \quad s, t \in \mathcal{T}_0.$$

The covariance operator of the process X is given by

$$\Gamma f(\cdot) = \int C(s, \cdot) f(s) ds, \quad f \in \mathcal{L}^2(\mathcal{T}_0).$$

Let $\lambda_1 \geq \lambda_2 \geq \dots$ and $\phi_1, \phi_2, \dots$ be the eigenvalues and eigenfunctions of the covariance operator Γ. The set $\phi = \{\phi_k\}_{k \geq 1}$ forms a complete orthonormal basis of $\mathcal{L}^2(\mathcal{T}_0)$. Note that ϕ contains an infinite number of elements. Using the Karhunen-Loève decomposition, we get

$$X(t) = \mu(t) + \sum_{k=1}^{\infty} \mathfrak{c}_k \phi_k(t), \quad t \in \mathcal{T}_0,$$

where $\mathfrak{c}_k = \langle X - \mu, \phi_k \rangle$ are the projection of the centered curve onto the eigenfunctions. We have that $\mathbb{E}(\mathfrak{c}_k) = 0$, $\mathbb{E}(\mathfrak{c}_k^2) = \lambda_k$ and $\mathbb{E}(\mathfrak{c}_k \mathfrak{c}_r) = 0$ for $k \neq r$.

Assume we observe N realizations $X_1, \dots, X_N$ of the process X. Estimators of the mean and covariance function are given by

$$\widehat{\mu}(t) = \frac{1}{N} \sum_{n=1}^{N} X_i(t) \quad \text{and} \quad \widehat{C}(s, t) = \frac{1}{N} \sum_{i=1}^{N} (X_i(t) - \widehat{\mu}(t)) (X_i(s) - \widehat{\mu}(s)).$$

And thus, the estimator of the covariance operator follows as

$$\widehat{\Gamma}_N f(\cdot) = \int \widehat{C}(s, \cdot) f(s) ds, \quad f \in \mathcal{L}^2(\mathcal{T}_0).$$

****

Assume that there exists a basis of functions $\{\psi_k\}_{1 \leq k \leq K}$ such that the curves can be extended into this basis

$$X(t) = \sum_{k=1}^{K} c_k \psi_k(t), \quad t \in \mathcal{T}_0.$$

The mean function is given by

$$\mu(t) = \sum_{k=1}^{K} \mathbb{E}(c_k) \psi_k(t), \quad t \in \mathcal{T}_0.$$

[*]MACSI, Department of Mathematics and Statistics, University of Limerick, Ireland steven.golovkine@ul.ie
[†]MACSI, Department of Mathematics and Statistics, University of Limerick, Ireland edward.gunning@ul.ie

The covariance function is given by

$$C(s,t) = \sum_{k=1}^{K}\sum_{l=1}^{K} \left(\mathbb{E}(c_k c_l) - \mathbb{E}(c_k)\mathbb{E}(c_l)\right)\psi_k(s)\psi_l(t), \quad s,t \in \mathcal{T}_0.$$

Using N realizations of the process, we have $X(t) = C\Psi(t)$ where

$$X(t) = \begin{pmatrix} X_1(t) \\ \vdots \\ X_N(t) \end{pmatrix}, \quad C = \begin{pmatrix} c_{11} & \cdots & c_{1K} \\ \vdots & \ddots & \vdots \\ c_{N1} & \cdots & c_{NK} \end{pmatrix} \quad \text{and} \quad \Psi(t) = \begin{pmatrix} \psi_1(t) \\ \vdots \\ \psi_K(t) \end{pmatrix}.$$

The estimation of the mean and covariance functions are given by

$$\widehat{\mu}(t) = \frac{1}{N}\mathbb{1}_N^\top C\Psi(t) \quad \text{and} \quad \widehat{C}(s,t) = \frac{1}{N}\Psi(s)^\top C^\top C\Psi(t).$$

We denote by W the matrix of inner products of the functions of the basis Ψ. The entries of W are given by

$$W_{kl} = \langle \psi_k, \psi_l \rangle, \quad 1 \leq k,l \leq K.$$

1.1 By diagonalization of the covariance operator

From [3]. The estimation of the eigenvalues and eigenfunctions of the covariance operator is usually performed by estimating the covariance surface on a fine grid and diagonalize it. Let $H = (t_1, \dots, t_m)$ a grid. In that case, the empirical covariance matrix $\widehat{C}$ is of size $m \times m$. It results to sets of eigenvalues $(\rho_1, \dots, \rho_m)$ and eigenvectors $(u_1, \dots, u_m)$ such that

$$\widehat{C}u = \rho u. \tag{1}$$

Given ϕ an eigenfunction of the covariance operator Γ, let $\widetilde{\phi}$ be the vector of length m with entries $\phi(t)$ for $t \in H$. Then, for each $t \in H$,

$$\widehat{\Gamma}\phi(s) = \int \widehat{C}(s,t)\phi(t)dt \approx \frac{|\mathcal{T}_0|}{|H|}\sum_{k=1}^{m}\widehat{C}(s,t_k)\widetilde{\phi}_k.$$

The equation $\Gamma\phi = \lambda\phi$ has the approximate discrete form

$$\frac{|\mathcal{T}_0|}{|H|}\widehat{C}\widetilde{\phi} = \lambda\widetilde{\phi}. \tag{2}$$

The solutions of the equations (1) and (2) are the same with eigenvalues $\lambda = \frac{|\mathcal{T}_0|}{|H|}\rho$. To approximate the eigenfunction ϕ from $\widetilde{\phi}$, we use interpolation techniques.

****

Using the expansion of the realizations of the process into the basis of functions Ψ. The eigenfunctions of the estimation of the covariance operator can also be expanded in Ψ

$$\widehat{\phi}_k(t) = \sum_{l=1}^{K} b_{kl}\psi_l(t) = \Psi(t)^\top b_k, \quad \text{where} \quad b_k = (b_1, \dots, b_K)^\top.$$

Then,

$$\begin{aligned}
\widehat{\Gamma}_N\widehat{\phi}(t) &= \int \widehat{C}(t,s)\widehat{\phi}(s)ds \\
&= \int \frac{1}{N}\Psi(t)^\top C^\top C\Psi(s)\Psi(s)^\top b ds \\
&= \frac{1}{N}\Psi(t)^\top C^\top CWb \\
&= \lambda\Psi(t)^\top b.
\end{aligned}$$

As this relationship should be true for all t, we get

$$\frac{1}{N}C^\top CWb = \lambda b.$$

1.2 By diagonalization of the inner product matrix

In this section, we use the duality relation between row and column spaces of a data matrix (from [1]). Consider the matrix M with entries

$$M_{ij} = \langle X_i - \widehat{\mu}, X_j - \widehat{\mu} \rangle, \quad i, j = 1, \dots, N.$$

The relationship between all nonzero eigenvalues $\widehat{\lambda}_1 \geq \widehat{\lambda}_2 \geq \cdots$ of the empirical covariance operator $\widehat{\Gamma}_N$ and the eigenvalues $l_1 \geq l_2 \geq \cdots$ of the matrix M is given by

$$\widehat{\lambda}_k = \frac{l_k}{N}, \quad k = 1, 2, \dots.$$

Let $\widehat{\phi}_1, \widehat{\phi}_2, \dots$ be the eigenfunctions of $\widehat{\Gamma}_N$ and $v_1, v_2, \dots$ be the orthonormal eigenvectors of M. These quantities are linked by

$$\widehat{\phi}_k(\cdot) = \frac{1}{\sqrt{l_k}} \sum_{i=1}^{N} v_{ik}(X_i(\cdot) - \widehat{\mu}(\cdot)), \quad k = 1, 2, \dots$$

The empirical scores are given by

$$\widehat{\mathfrak{c}}_{ik} = \sqrt{l_k} v_{ik}, \quad k = 1, 2, \dots$$

****

For the eigenvalues and the eigenfunctions. For $k = 1, 2, \dots,$ we have

$$M v_k = l_k v_k. \tag{3}$$

Let $X = (X_1(\cdot) - \widehat{\mu}(\cdot) \cdots X_N(\cdot) - \widehat{\mu}(\cdot))^\top$. By left multiplying (7) by $X^\top$, we obtain

$$X^\top M v_k = l_k X^\top v_k. \tag{4}$$

Then, we have

$$\begin{aligned}
X^\top M v_k &= \sum_{i=1}^{N} (X_i(\cdot) - \widehat{\mu}(\cdot)) \sum_{j=1}^{N} \langle X_i(\cdot) - \widehat{\mu}(\cdot), X_j(\cdot) - \widehat{\mu}(\cdot) \rangle v_{jk} \\
&= \int_{\mathcal{T}_0} \sum_{i=1}^{N} (X_i(\cdot) - \widehat{\mu}(\cdot)) (X_i(s) - \widehat{\mu}(s)) \sum_{j=1}^{N} (X_j(s) - \widehat{\mu}(s)) v_{jk} ds \\
&= \int_{\mathcal{T}_0} N\widehat{C}(\cdot, s) \sum_{j=1}^{N} (X_j(s) - \widehat{\mu}(s)) v_{jk} ds \\
&= N\widehat{\Gamma}_n \left(\sum_{j=1}^{N} (X_j(\cdot) - \widehat{\mu}(\cdot)) v_{jk} \right) (\cdot)
\end{aligned}$$

and

$$l_k X^\top v_k = l_k \sum_{i=1}^{N} (X_i(\cdot) - \widehat{\mu}(\cdot)) v_{ik}.$$

From (4), we get

$$\widehat{\Gamma}_n \left(\sum_{j=1}^{N} (X_j(\cdot) - \widehat{\mu}(\cdot)) v_{jk} \right) (t) = \frac{l_k}{N} \sum_{i=1}^{N} (X_i(t) - \widehat{\mu}(t)) v_{ik}, \quad \text{for all } t \in \mathcal{T}_0.$$

By identification, we find that

$$\widehat{\lambda}_k = \frac{l_k}{N} \quad \text{and} \quad \widehat{\phi}_k(\cdot) = \sum_{i=1}^{N} v_{ik}(X_i(\cdot) - \widehat{\mu}(\cdot)), \quad k = 1, 2, \dots$$

For $k = 1, 2, \ldots$, we have

$$\begin{aligned}
\left\|\widehat{\phi}_k\right\|^2 &= \langle \sum_{i=1}^N v_{ik}(X_i(\cdot) - \widehat{\mu}(\cdot)), \sum_{j=1}^N v_{jk}(X_j(\cdot) - \widehat{\mu}(\cdot))\rangle \\
&= \sum_{i=1}^N \sum_{j=1}^N v_{ik} v_{jk} \langle (X_i(\cdot) - \widehat{\mu}(\cdot)), (X_j(\cdot) - \widehat{\mu}(\cdot))\rangle \\
&= \sum_{i=1}^N \sum_{j=1}^N v_{ik} v_{jk} \int_{\mathcal{T}_0} (X_i(s) - \widehat{\mu}(s))(X_j(s) - \widehat{\mu}(s)) ds \\
&= \sum_{i=1}^N \sum_{j=1}^N v_{ik} v_{jk} M_{ij} \\
&= \sum_{i=1}^N v_{ik} \sum_{j=1}^N M_{ij} v_{jk} \\
&= \sum_{i=1}^N v_{ik} l_k v_{ik} \\
&= l_k \left\|v_k\right\|^2 \\
&= l_k
\end{aligned}$$

We normalise $\widehat{\phi}_k$ by $1/\sqrt{l_k}$ to have an orthonormal basis.

For the scores. For $k = 1, 2, \ldots$, we have

$$\begin{aligned}
\widehat{\mathfrak{c}}_{ik} &= \langle X_i(\cdot) - \widehat{\mu}(\cdot), \widehat{\phi}_k(\cdot)\rangle \\
&= \langle X_i(\cdot) - \widehat{\mu}(\cdot), \frac{1}{\sqrt{l_k}} \sum_{j=1}^N v_{jk}(X_j(\cdot) - \widehat{\mu}(\cdot))\rangle \\
&= \frac{1}{\sqrt{l_k}} \sum_{j=1}^N v_{jk} \langle X_i(\cdot) - \widehat{\mu}(\cdot), X_j(\cdot) - \widehat{\mu}(\cdot)\rangle \\
&= \frac{1}{\sqrt{l_k}} \sum_{j=1}^N v_{jk} M_{ij} \\
&= \frac{1}{\sqrt{l_k}} (M v_k)_i \\
&= \frac{1}{\sqrt{l_k}} l_k v_{ik} \\
&= \sqrt{l_k} v_{ik}
\end{aligned}$$

****

Concerning the expansion of the data into the basis of function Ψ, we write

$$M = \left(CW^{1/2}\right)\left(CW^{1/2}\right)^T.$$

We also assume that $\widehat{\phi}_1, \widehat{\phi}_2, \ldots$ the eigenfunctions of $\widehat{\Gamma}_N$ have a decomposition into the basis Ψ

$$\widehat{\phi}_k(t) = \sum_{l=1}^K b_{kl} \psi_l(t) = \Psi(t)^\top b_k, \quad \text{where} \quad b_k = (b_{k1}, \ldots, b_{kK})^\top.$$

We have, for all $t \in \mathcal{T}_0$,

$$\begin{aligned}
\widehat{\Gamma}_N \phi_k(t) &= \int_{\mathcal{T}_0} \widehat{C}(t,s)\phi_k(s)ds \\
&= \frac{1}{N}\int_{\mathcal{T}_0} \Psi(t)^\top C^\top C \Psi(s)\Psi(s)^\top b_k ds \\
&= \frac{1}{N}\Psi(t)^\top C^\top C \int_{\mathcal{T}_0} \Psi(s)\Psi(s)^\top ds b_k \\
&= \frac{1}{N}\Psi(t)^\top C^\top C W b_k.
\end{aligned}$$

From the eigenequation, we have that

$$\widehat{\Gamma}_N \phi_k(t) = \lambda_k \phi_k(t) \iff \frac{1}{N}\Psi(t)^\top C^\top C W b_k = \lambda_k \Psi(t)^\top b_k, \quad t \in \mathcal{T}_0.$$

Since this equation must be true for all $t \in \mathcal{T}_0$, this imply the equation

$$C^\top C W b_k = N\lambda_k b_k \tag{5}$$

As the eigenfunctions are assumed to be normalized, $\|\phi_k\|^2 = 1$. And so, $b_k^\top W b_k = 1$. Let $u_k = W^{1/2} b_k$. Then, from Equation (5), we obtain

$$W^{1/2} C^\top C W^{1/2} u_k = N\lambda_k u_k \iff \left(CW^{1/2}\right)^\top \left(CW^{1/2}\right) u_k = N\lambda_k u_k. \tag{6}$$

From the eigendecomposition of the matrix M, we get

$$M v_k = l_k v_k \iff \left(CW^{1/2}\right)\left(CW^{1/2}\right)^T v_k = l_k v_k. \tag{7}$$

The equations (6) and (7) are eigenequations in the classical PCA case, with the duality $X^\top X$ and $XX^\top$. Following [2], we find that, for $1 \leq k \leq K$,

$$\lambda_k = \frac{l_k}{N}, \quad v_k = \frac{1}{\sqrt{l_k}} C W^{1/2} u_k \quad \text{and} \quad u_k = \frac{1}{\sqrt{l_k}} W^{1/2} C^\top v_k.$$

And finally, to get the coefficient of the eigenfunctions, for $1 \leq k \leq K$,

$$b_k = W^{-1/2} u_k = \frac{1}{\sqrt{l_k}} C^\top v_k.$$

Acknowledgment

S. Golovkine was partially supported by Science Foundation Ireland under Grant No. 19/FFP/7002 and co-funded under the European Regional Development Fund.

References

[1] M. Benko, W. Härdle, and A. Kneip. Common functional principal components. *The Annals of Statistics*, 37(1):1–34, Feb. 2009.

[2] J. Pagès. *Multiple Factor Analysis by Example Using R*. CRC Press, Nov. 2014.

[3] J. Ramsay and B. W. Silverman. *Functional Data Analysis*. Springer Science & Business Media, June 2005.