Penalized PCA approaches for B-spline expansions of smooth functional data

- Ana M. Aguilera; M. Carmen Aguilera-Morillo

- Penalized PCA approaches for B-spline expansions of smooth functional data

Penalized PCA approaches for B-spline expansions of smooth functional data

A.M. Aguilera [a,*], M.C. Aguilera-Morillo [b]

[a] Facultad de Ciencias, Campus de Fuentenueva s/n, 18071 Granada, Spain
[b] Facultad de Farmacia, Campus de Cartuja s/n, 18071 Granada, Spain

ABSTRACT

Functional principal component analysis (FPCA) is a dimension reduction technique that explains the dependence structure of a functional data set in terms of uncorrelated variables. In many applications the data are a set of smooth functions observed with error. In these cases the principal components are difficult to interpret because the estimated weight functions have a lot of variability and lack of smoothness. The most common way to solve this problem is based on penalizing the roughness of a function by its integrated squared d-order derivative. Two alternative forms of penalized FPCA based on B-spline basis expansions of sample curves and a simpler discrete penalty that measures the roughness of a function by summing squared d-order differences between adjacent B-spline coefficients (P-spline penalty) are proposed in this paper. The main difference between both smoothed FPCA approaches is that the first uses the P-spline penalty in the least squares approximation of the sample curves in terms of a B-spline basis meanwhile the second introduces the P-spline penalty in the orthonormality constraint of the algorithm that computes the principal components. Leave-one-out cross-validation is adapted to select the smoothing parameter for these two smoothed FPCA approaches. A simulation study and an application with chemometric functional data are developed to test the performance of the proposed smoothed approaches and to compare the results with non penalized FPCA and regularized FPCA.

© 2013 Elsevier Inc. All rights reserved.

1. Introduction

Many application areas have functional data that come from the observation of a random function at a discrete set of sampling points. In the majority of cases the sample curves are functions of time but in many others the argument is a different magnitude. In spectroscopy, for example, the NIR spectrum is a functional variable whose observations are measured as functions of wavelengths. The potential of functional data analysis methodologies for the chemometric analysis of spectroscopic data was shown in [1]. A wide variety of applications with functional data in different fields were collected and analyzed by [2].

When analyzing a functional data set it is usual to have a large number of regularly spaced observations for each sample curve. Because of this a reduction dimension technique is necessary for explaining the main features of a set of sample curves in terms of a small set of uncorrelated variables. This problem was solved by generalizing principal component analysis to the case of a continuous-time stochastic process [3]. Asymptotic properties of the estimators of FPCA were deeply studied in the general context of functional variables [4]. Nonparametric methods were developed to perform FPCA for the case of a small number of irregularly spaced observations of each sample curve [5,6]. As in the multivariate case, the interpretation

of the principal component scores and loadings is a useful tool for discovering the relationships among the variables associated to a functional data set. To avoid misinterpretation of PCA, a new type of plots, named Structural and Variance Information plots, were recently introduced by [7].

FPCA is a flexible tool in functional data analysis that is successfully used to solve important problems as the estimation of the functional parameter in different functional regression models [8–14]. An alternative methodology for solving this estimation problem in the functional linear model is the functional version of partial least squares (PLS) regression. A estimation procedure based on basis expansions of sample curves was introduced by [15]. A Bayesian approach to FPCA based on a generative model for noisy and sparse observations of curves was developed in [16].

One usual form of estimating FPCA from discrete observations of the sample curves is based on basis expansion approximation. This way, FPCA of a set of curves is reduced to multivariate PCA of a transformation of the matrix of basis coefficients [17]. B-spline basis are appropriate to approximate smooth curves. Cubic spline interpolation with B-spline basis can be considered for approximating smooth sample curves observed without error [18]. On the other hand, least squares approximation with B-spline basis is appropriate for reconstructing the true functional form of noisy smooth curves. This type of approximation was performed to forecast lupus flares from time evolution of stress level [19]. The problem is that regression splines do not control the degree of smoothness and, consequently, the principal component curves show substantial variability and their interpretation is difficult. This problem must be solved by introducing some kind of smoothing in the estimation of principal component curves.

There are different ways of introducing smoothing in the estimation of FPCA. On the one hand, the data can be smooth first and then an unpenalized FPCA is carried out. A spline smoothing that penalizes the integrated squared second derivative of each sample path was considered by [20,21]. This approach was applied for smoothing and reconstructing a magnetic resonance imaging (FMRI) functional data in [22]. On the other hand, the smoothing can be introduced within the FPCA algorithm. Two different approaches for smoothing functional principal components analysis were proposed by [23] and [24]. Both approaches use a continuous penalty that measures the roughness of the principal component curves by their integrated squared d-order derivative but they differ in the way they incorporate the penalty. The Rice-Silverman approach introduces the roughness penalty in the definition of the sample variance of the principal component weight functions. The Silverman approach is known as regularized FPCA (RFPCA) and introduces the penalty in the orthonormality constraint between principal components. This FPCA approach was extended to the case of multivariate functional data sets by using Gaussian basis functions instead of B-splines [23]. An application of regularized FPCA with B-splines basis in actuarial science was performed to estimate the risk of occurrence of a claim in terms of the driver's age and others significative variables [24]. A third way of penalizing FPCA is based on smoothing not the data or the components, but the covariance operator, whose eigenfunctions are the principal component functions [6]. Penalized rank one approximation was recently proposed as an alternative approach to the estimation of FPCA [25].

Penalized spline regression [26] is an increasingly popular smoothing approach that was used to estimate the functional sample mean and to develop an iterative P-spline algorithm for estimating FPCA in [27]. The P-spline penalty measures the roughness of a function by summing squared d-order differences between adjacent basic coefficients. In this paper, two different versions of smoothed FPCA based on penalized splines (P-splines) with B-splines basis are introduced and compared. The first approach carries out an unpenalized FPCA on the P-spline smoothing of the sample curves. The second approximates the sample curves by unpenalized least squares and then incorporate the P-spline penalty in the orthonormality constraint within the FPCA algorithm. The accuracy of the estimates provided by both P-spline smoothed approaches is tested with simulated and real data, and the results compared with non penalized FPCA and regularized FPCA. In order to get an optimum estimation of the smoothing parameter, leave-one-out cross-validation is adapted to this context.

2. Functional principal component analysis

Let $\{x_i(t) : t \in T, i = 1, \ldots, n\}$ be a sample of functions that are the sample information related to a functional variable X. It will be supposed that they are observations of a second order stochastic process $X = \{X(t) : t \in T\}$, continuous in quadratic mean whose sample functions belong to the Hilbert space $L^2(T)$ of square integrable functions with the usual inner product $\langle f, g \rangle = \int_T f(t)g(t)dt, \forall f, g \in L^2(T)$. Multivariate PCA was extended to the functional case to reduce the infinite dimension of a functional predictor and to explain its dependence structure by a reduced set of uncorrelated variables [3]. In order to compute the functional principal components, let us assume without loss of generality that the observed curves are centered so that the sample mean $n^{-1}\sum_{i=1}^n x_i(t)$ is zero.

The principal components are obtained as uncorrelated generalized linear combinations with maximum variance (Var). In general, the j-th principal component scores are given by

$$\xi_{ij} = \int_T x_i(t)f_j(t)dt, \quad i = 1, \ldots, n, \tag{1}$$

where the weight function or loading f_j is obtained by maximizing the variance solving

$$\begin{cases} Max_f \ Var\left[\int_T x_i(t)f(t)dt\right], \\ s.t. \|f\|^2 = 1 \quad \text{and} \quad \int f_\ell(t)f(t)dt = 0, \ \ell = 1, \ldots, j-1. \end{cases}$$

It can be shown that the weight functions are the eigenfunctions of the sample covariance operator C. That is, the solutions to the eigenequation

$$C(f_j)(t) = \int c(t,s)f_j(s)ds = \lambda_j f_j(t), \tag{2}$$

where $c(t,s)$ is the sample covariance function and $\lambda_j = Var[\xi_j]$. The sample curves are then expressed in terms of the functional principal components by

$$x_i(t) = \sum_{i=1}^{n-1} \xi_{ij} f_j(t).$$

This principal component decomposition can be truncated providing the best linear approximation of the sample curves in the least squares sense $x_i^q(t) = \sum_{i=1}^{q} \xi_{ij} f_j(t)$, whose explained variance is given by $\sum_{i=1}^{q} \lambda_i$.

The problem inherent to many applications is that interpreting the components is not always straightforward. This problem is usually solved by a rotation of the principal component curves that simplifies the factor structure and therefore makes the interpretation easier. There are two main types of rotation: orthogonal when the resulting factors are also orthogonal to each other, and oblique when the new factors are not required to be orthogonal to each other. Oblique rotations relax the orthogonality constraint in order to simplify the interpretation. They are used more rarely than their orthogonal counterparts but, recently, new techniques are developed based on oblique rotations. An example is independent component analysis that was originally created in the domain of signal processing and neural networks, and derives, directly from the data, an oblique solution that maximizes statistical independence [28].

The most popular orthogonal rotation method is indubitably Varimax. This rotation criterion has been applied to functional PCA in two different ways, one based on Varimax rotation of the matrix of basis coefficients of the principal component curves, and the other based on Varimax rotation of the matrix of values of the principal component curves in a grid of equally spaced time points (see [29] for a detailed development of this functional Varimax rotation). The extension of oblique rotations to functional PCA is out of the scope of this paper but could be developed by following the same idea that in the Varimax case.

2.1. Basis expansion estimation

In order to compute the principal component weights it is necessary to solve the second order integral Eq. (2). It is a difficult problem that is further complicated in practice because the sample curves are usually observed at a finite set of sampling points that can be different for the sample individuals. This means that in real data applications the sample information is given by the vectors $\{x_i = (x_{i0}, \ldots, x_{im_i})', i = 1, \ldots, n\}$, with x_{ik} being the observed value for the sample path, $x_i(t)$ at the sampling point $t_{ik}, k = 0, 1, \ldots, m_i$.

The functional form of sample paths must be reconstructed from the discrete observations by using several different methods that must be chosen depending on how the functional data was observed and the main characteristics of sample curves.

One usual solution is to assume that sample paths belong to a finite-dimension space spanned by a basis $\{\phi_1(t), \ldots, \phi_p(t)\}$, so that they are expressed as

$$x_i(t) = \sum_{j=1}^{p} a_{ij} \phi_j(t), \quad i = 1, \ldots, n.$$

The main objective of this paper is to solve the problem of estimating the functional principal components from a sample of smooth curves observed with error so that the sample data are given by

$$x_{ik} = x_i(t_{ik}) + \epsilon_{ik} \quad k = 0, 1, \ldots, m_i, \ i = 1, \ldots, n.$$

This means that least squares approximation with B-splines basis [30] are an appropriate choice to approximate the basis coefficients.

A B-spline basis of degree q generates the space of the splines of the same degree, defined as curves consisting of piecewise polynomials of degree q that join up smoothly at a set of definition knots with continuity in their derivatives up to order $q - 1$. The dimension of the B-spline basis of degree q equals the order of the polynomials plus the number of interior breakpoints (see [30] for a detailed explanation). The spline functions of degree q are smooth and well-behaved functions that provide design flexibility so that by increasing the degree q, we can progressively switch from the simplest piecewise constant ($q = 0$) and piecewise linear ($q = 1$) representations to the other extreme, which corresponds to a bandlimited model ($n \to \infty$). Formulas and computational algorithms for optimal smoothing curves with B-splines basis for given set of discrete, not necessarily equally spaced, data were studied in detail in [31,32]. Basis of splines were recently used to evaluate paper manufactured using Eucalyptus globulus by means of multivariate adaptive regression splines [33].

Other useful basis systems are Fourier functions for periodic data, piecewise constant functions for counting processes or wavelets bases for curves with strong local behavior. A successful application of FPCA with Fourier basis expansions was developed in [34] where different functional principal component regression models were developed to forecast cypress

pollen concentration from daily evolution of temperatures. An alternative solution could be to use non-parametric procedures for estimating curves [35].

Let us suppose that the sample paths are expressed in terms of basis functions, so that $x = A\phi$, where $A = (a_{ij})$ the coefficient matrix, $\phi = (\phi_1, \ldots, \phi_p)'$ and $x = (x_1, \ldots, x_n)'$. Then, the principal component weight function f_j admits the basis expansion

$$f_j(t) = \sum_{k=1}^{p} b_{jk}\phi_k(t) = \phi(t)'b_j,$$

with $b_j = (b_{j1}, \ldots, b_{jp})'$. In this case,

$$Var[\xi] = Var\left[\int_T x_i(t)f(t)dt\right] = n^{-1}\sum_{i=1}^{n}\left[\int_T \left(\sum_{j=1}^{p}a_{ij}\phi_j(t)\right)\left(\sum_{k=1}^{p}b_k\phi_k(t)\right)\right] = n^{-1}\sum_{i=1}^{n}b'\Psi a_i'a_i\Psi b = b'\Psi V\Psi b,$$

where $V = n^{-1}A'A, A = (a_{ij})_{n \times p}$ and $\Psi = (\Psi_{ij})_{p \times p} = \int_T \phi_i(t)\phi_j(t)dt$. Therefore, computing the j-th principal weight function is reduced to solve the maximization problem

$$\begin{cases} Max_b \ b'\Psi V\Psi b, \\ s.t. b'\Psi b = 1 \quad \text{and} \quad b_l\Psi b = 0, \ \ell = 1, \ldots, j-1. \end{cases}$$

This means that FPCA is equivalent to the multivariate PCA of matrix $A\Psi^{\frac{1}{2}}$, with $\Psi^{\frac{1}{2}}$ being the squared root of the matrix of inner products between basis functions [17].

Then, the vector b_j of basis coefficients of the jth principal weight function is given by $b_j = \Psi^{-\frac{1}{2}}u_j$, where the vectors u_j are computed as the solutions to the eigenvalue problem $n^{-1}\Psi^{\frac{1}{2}}A'A\Psi^{\frac{1}{2}}u_j = \lambda_j u_j$, where $n^{-1}\Psi^{\frac{1}{2}}A'A\Psi^{\frac{1}{2}}$ is the sample covariance matrix of $A\Psi^{\frac{1}{2}}$.

The principal components curves estimated by this unpenalized FPCA approach with B-spline basis can present a lot of variability and have difficult interpretation. To solve this problem two new ways of introducing smoothing in FPCA are proposed in this work. The first one consists of FPCA of the P-spline smoothing of sample curves. The second one is a P-spline version of the smoothed FPCA carried out in [36].

3. Functional PCA of P-splines

The first version of smoothed FPCA consists of functional PCA of the P-spline smoothing of the original data.

An appropriate method to approximate the basis coefficients of smooth sample curves observed with error is ordinary least squares approximation in terms of B-spline basis. The curves fitted by this way are known as regression splines and their main problem is that they don't control the degree of smoothness. A penalized least squares approach that measures the roughness of a function by means of its integrated dth derivative was first introduced in [37] for solving this problem with $d = 2$. Smoothing splines are computed by using this continuous penalization in terms of B-spline basis expansions. A simpler discrete penalty approach is based on defining the roughness of a function by summing squared d-order differences between adjacent B-splines [26]. This discrete penalty has two main advantages. First, it is mathematically simpler than smoothing splines because is a good discrete approximation of the integrated square of the d-th derivative. Second, the number of knots is not so determinant as in regression splines and can be easily compute by using Ruppert's law of thumb [38]. In general, the knots of a P-spline must be equally spaced and its number sufficiently large to fit the data and not so large that computation time is unnecessarily big [39].

P-spline smoothing of each sample curve is obtained by minimizing the following penalized least squares error in terms of B-spline basis functions

$$PMSE_d(a_i|x_i, \lambda) = (x_i - \Phi_i a_i)'(x_i - \Phi_i a_i) + \lambda a_i'P_d a_i,$$

where $\Phi_i = (\phi_j(t_{ik}))_{m_i \times p}, a_i = (a_{i1}, \ldots, a_{ip})'$ and P_d is the matrix of penalties given by

$$P_d = (\triangle^d)'\triangle^d,$$

with $\triangle^d$ being the matrix representation of the d-order difference operator.

These smoothers are known as penalized splines (P-splines) and their B-spline basis coefficients are given by

$$\hat{a}_i = (\Phi_i'\Phi_i + \lambda P_d)^{-1}\Phi_i'x_i. \tag{3}$$

The most used penalty order is $d = 2$ although a different order can be used depending on the variability of the curve and the quantity of noise in the observed data. A penalty of order 2 is equivalent to

$$a_i'P_2 a_i = (a_{i1} - 2a_{i2} + a_{i3})^2 + \cdots + (a_{ip-2} - 2a_{ip-1} + a_{ip})^2,$$

where $P_2 = (\triangle^2)' \triangle^2$ with $\triangle^2$ matrix given by

$$\triangle^2 = \begin{bmatrix} 1 & -2 & 1 & 0 & \cdots \\ 0 & 1 & -2 & 1 & \cdots \\ 0 & 0 & 1 & -2 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

As we can see in the above development there are some important choices related to the P-spline fitting: the smoothing parameter, the order of the penalty, the degree of the B-spline basis and the number of knots. The choice of the smoothing parameter is discussed later in this section. The simplest and most usual choice for the other three parameters that should work well in most applications is use a quadratic penalty, cubic splines and one knot for every four or five observations up to a maximum of about 40 knots [38].

The cubic spline functions (piece-wise polynomial curves that has continuous derivatives to order two) tend to be the most popular in applications with smooth functions because of their minimum curvature property (they have less tendency to oscillate). In fact, cubic spline functions have proven to be a good compromise between accuracy and complexity. Higher order splines allow more strongly curved surfaces to be modeled but also requires more calculations. Usually, degree 3 or 4 is sufficient for B-splines.

Once the P-spline approximation of sample curves have been computed FPCA is performed on the approximated curves instead of the original sample paths. This way FPCA on P-splines is equivalent to multivariate PCA of the matrix $A\Psi^{\frac{1}{2}}$ where A is the matrix whose rows are the estimated coefficients of the P-spline smoothing of the sample curves in terms of a basis of B-splines.

3.1. Selection of the smoothing parameter

As in other smoothing methods, the smoothing parameter of P-splines controls the smoothness of the fitted curve. P-splines penalize distant coefficients so that the larger is the parameter the smoother is the fitted curve. Classical methods for smoothing parameter selection are leave-one-out cross-validation (CV), generalized cross-validation [40], the Akaike information criterion (AIC) [41] and the Bayesian information criterion (BIC) [42]. A nonparametric procedure for choosing the P-spline parameters has been performed by [43], where mixed model methods based on restricted maximum likelihood (REML) estimation were applied for smoothing parameter selection. The P-spline smoothing of the FPCA (SFPCA) introduced in the next section is not associated to the estimation of a regression model and because of this a direct implementation of the GCV, AIC, BIC and REML criteria is not possible. To make the results comparable with those of the SFPCA approach, CV criterion is used in this paper to choose the optimum smoothing parameter.

In order to select the same smoothing parameter for the n fitted P-splines, a CV method based on minimizing the mean of the cross-validation errors over all P-splines is applied in this paper. This CV criterion consists of selecting the smoothing parameter λ that minimizes the expression

$$CVMSE(\lambda) = \frac{1}{n} \sum_{i=1}^{n} \sqrt{\sum_{k=0}^{m_i} \left(x_{ik} - \hat{x}_{ik}^{(-k)} \right)^2 / (m_i + 1)},$$

where $\hat{x}_{ik}^{(-k)}$ are the values of the i-th sample path estimated at time t_{ik} avoiding the k-th observation knot in the iterative estimation process.

4. P-Spline smoothed functional PCA

In order to control the roughness of the weight functions f_i, the principal components can be computed by maximizing a penalized sample variance that introduces the roughness penalty into the orthonormality constraint. A continuous penalty based on the integrated squared d-order derivative was first proposed by [36]. Following this idea, a penalized sample variance based on a discrete P-spline roughness penalty is introduced in this paper.

Let us consider a B-spline expansion of the weight functions given by $f(t) = \sum_{k=1}^{p} b_k \phi_k(t) = \phi(t)'b$, where $\phi(t)$ is a B-spline basis. The discrete P-spline roughness penalty function is defined by $PEN_d(f) = b'P_d b$, with P_d defined as in Section 3 and $b = (b_1, \ldots, b_p)$ being the vector of coefficients of the weight function in terms of the B-spline basis.

The principal components are now computed as generalized linear combinations of the functional variable that maximize the penalized sample variance defined by

$$PVar\left[\int_T x_i(t)f(t)dt \right] = \frac{Var[\int_T x_i(t)f(t)dt]}{\|f\|^2 + \lambda PEN_d(f)} = \frac{b'\Psi V \Psi b}{b'\Psi b + \lambda b'P_d b} = \frac{b'\Psi V \Psi b}{b'(\Psi + \lambda P_d)b},$$

whit λ being the smoothing parameter that controls the roughness of the weight function.

The jth principal component is now defined as in Eq. (1) and the factor loading coefficients of the factor loading f_j are obtained by solving

$$\begin{cases} Max_b \ PVar\left[\int_T x_i(t)f(t)dt\right] = \frac{b'\Psi V\Psi b}{b'(\Psi + \lambda P_d)b}, \\ s.t. b'\Psi b = 1 \ \text{and} \ b'\Psi b_l + b'P_d b_l = 0, \quad \ell = 1, \ldots, j-1. \end{cases}$$

Let us observe that the first constraint is the usual requirement $\|f\|^2 = 1$ and the second is a modified form of orthogonality that takes into account the roughness of the weight function.

This variance maximization problem can be converted into a eigenvalue problem

$$\Psi V\Psi b = \beta(\Psi + \lambda P_d)b.$$

By applying a factorization (SVD or Choleski) of the form $LL' = \Psi + \lambda P_d$ with L being a lower triangular matrix, we have $L^{-1}\Psi V\Psi L^{-1'}(L'b) = \beta(L'b)$, so that the weight coefficients are computed by solving the eigenvalue problem

$$L^{-1}\Psi V\Psi L^{-1'}u = \beta u,$$

with $L'b = u$. This way P-spline smoothed PCA is reduced to multivariate PCA of the matrix whose rows are the transformed vectors of coefficients $L^{-1}\Psi a_i$. That is multivariate PCA of matrix $A\Psi L^{-1'}$. Finally, the basis coefficients of the principal components curves are given by $b_j = L^{-1'}u_j$ renormalized so that $b_j'\Psi b_j = 1$. Let us observe that in this case the vectors a_i of basis coefficients of the sample curves in terms of the B-spline basis are first estimated by non penalized least squares approximation that provides $\hat{a}_i = (\Phi_i'\Phi_i)^{-1}\Phi_i'x_i$ with $\Phi_i = (\phi_j(t_{ik}))_{m_i \times p}, i = 1, \ldots, n$.

4.1. Selection of the smoothing parameter

As in previous section, selecting a suitable smoothing parameter is very important to control the smoothness of the weight function associated to each principal component. In this paper, leave one out cross-validation (CV) has been adapted by considering the discrete roughness penalty based on P-splines. It consists of selecting the value of λ that minimizes

$$CV(\lambda) = \frac{1}{p}\sum_{q=1}^{p}CV_q(\lambda),$$

where

$$CV_q(\lambda) = \frac{1}{n}\sum_{i=1}^{n}\|x_i - x_i^{q(-i)}\|^2,$$

with $x_i^{q(-i)} = \sum_{l=1}^{q}\xi_{il}^{(-i)}f_l^{(-i)}$ being the reconstruction of the sample curve x_i in terms of the first q principal components estimated from the sample of size $n-1$ that includes all sample curves except x_i.

In terms of basis expansions these quadratic distances are given by

$$\|x_i - x_i^{q(-i)}\|^2 = \int_T \left[x_i(t) - x_i^{q(-i)}(t)\right]^2 dt = \int_T \left[\sum_{j=1}^{p}a_{ij}\phi_j(t) - \sum_{l=1}^{q}\xi_{il}^{(-i)}\sum_{j=1}^{p}b_{lj}^{(-i)}\phi_j(t)\right]^2 dt = \int_T \left[\sum_{j=1}^{p}d_{ij}\phi_j(t)\right]^2 dt = d_i'\Psi d_i,$$

where $d_i = (d_{i1}, \ldots, d_{ip})'$ and $d_{ij} = a_{ij} - \sum_{l=1}^{q}\xi_{il}^{(-i)}b_{lj}^{(-i)}$.

5. Simulation study

The ability of the two smoothed versions of FPCA proposed in this paper (FPCA of P-splines and P-spline smoothed FPCA) to approximate the true form of the factor loadings functions from a set of smooth curves observed with noise is tested on simulated data. The results are compared with the ones provided by unpenalized FPCA in terms of B-splines and regularized FPCA (RFPCA) based on penalizing the roughness of a the principal component curves by its integrated squared 2-order derivative (see [36] for a detailed study).

Let $\{X_t : t \in [0, T]\}$ be the zero mean gaussian process with covariance function $C(t, s) = P\exp(-\alpha|t - s|)$ known as Ornstein–Uhlenbeck process. It can be shown [44] that the total variance of this process is T and the principal component weights functions are given by the solutions of the integral eigenequation

$$\int_0^T P\exp(-\alpha|t - s|)f(s)ds = \lambda f(t),$$

whose eigenvalues are

$$\lambda_i = \frac{2P\alpha}{\alpha^2 + b_i^2},$$

Table 1
The first 14 solutions b_i of the Eq. (4).

b_1	b_2	b_3	b_4	b_5	b_6	b_7
0.2164	0.8443	1.6020	2.3772	3.1574	3.9397	4.7230
b_8	b_9	b_{10}	b_{11}	b_{12}	b_{13}	b_{14}
5.5069	6.2911	7.0757	7.8603	8.6452	9.4301	10.2151

with b_i being the positive solutions of

$$\tan\left(b_i\frac{T}{2}\right) = \frac{\alpha}{b_i} \ (i\ odd),$$
$$\tan\left(b_i\frac{T}{2}\right) = \frac{-b_i}{\alpha} \ (i\ even). \tag{4}$$

The first fourteen b_i (solutions of Eq. 4) can be found in Table 1. The eigenfunctions normalized in $[0, T]$ are

$$f_i(t) = \frac{\cos\left(b_i\left(t - \frac{T}{2}\right)\right)}{\left[\frac{T}{2}\left(1 + \frac{\sin(b_i T)}{b_i T}\right)\right]^{\frac{1}{2}}} \ (i\ odd)$$

$$f_i(t) = \frac{\sin\left(b_i\left(t - \frac{T}{2}\right)\right)}{\left[\frac{T}{2}\left(1 - \frac{\sin(b_i T)}{b_i T}\right)\right]^{\frac{1}{2}}} \ (i\ even). \tag{5}$$

The principal components of the Ornstein–Uhlenbeck process are uncorrelated random variables with normal distribution so that the principal component decomposition of this process is given by

$$X(t) = \sum_{i=1}^{\infty} \lambda_i f_i(t) \xi_i, \tag{6}$$

where ξ_i are random variables with distribution $N(0, 1)$. Eq. (6) truncated in the 14th term provides a smoothed version of the Orsntein-Ulenbeck process that explains a 99.4% of its total variance. In order to get noisy observations, a random error $\varepsilon(t)$ with distribution $N(0, \sigma^2)$ was added so that the simulated process was

$$Y(t) = X^{14}(t) + \varepsilon(t). \tag{7}$$

The variance of the errors σ^2 was chosen to control the value of $R^2 = \left[Var\left(X^{14}\right)\right]/[Var(Y)]$ close to 0.8. The simulation was made for $T = 4, P = 1$ and $\alpha = 0.1$.

In order to test the performance of the four different FPCA approaches compared in this study, 350 samples of 100 sample curves of the contaminated process $Y(t)$ given by Eq. (7) were simulated at 41 equally spaced knots in the observed domain $[0, 4]$.

The first step is to reconstruct the true functional form of the original sample paths $X^{14}(t)$ from the noisy discrete observations $Y(t_k)$ with $t_k = k/10, k = 0, 1, \ldots, 40$. Following the methodology proposed in this paper, regression splines and P-splines will be computed in terms of B-spline basis. In Fig. 1 it can be seen an example where an original smooth curve (solid line), its contaminated curve (dotted line), the cubic B-spline (dashed line) and the quartic B-spline (dashed-dotted line) approximations are superposed for the non penalized splines (left) and the penalized splines (right) with different number of basis knots $(15, 25$ and $30)$. It can be seen that the reconstructions of the curves provided by cubic and quartic B-splines are very similar for penalized splines independently of the number of knots. In the non penalized case, the approximation provided by the quartic splines is worse when the number of knots increases by loosing control in the extremes. Because of this B-splines of degree 3 was chosen in this simulation study.

Our main goal is to get an accurate approximation of the true eigenfunctions and the original sample curves. The approximations of the first, second and third eigenfunctions of the Ornstein–Uhlenbeck process by using the four considered FPCA approaches with different number of basis knots $(15, 25$ and $30)$ are displayed in Fig. 2 for one of the 350 simulations. The true eigenfunctions given by the solutions of Eq. (5) are superposed with their approximations in the same plot. To show the ability of the smoothed FPCA approaches to approximate the original process $X^{14}(t)$, the reconstructions of two different sample paths with the first three PCs approximated by the four considered approaches are displayed in Fig. 3.

In order to draw general conclusions, the box plots of the MSEs of approximation of the eigenfunctions estimated by using FPCA, FPCA of P-splines, P-spline SFPCA and regularized FPCA with $15, 25$ and 30 basis knots for 350 simulations of the Ornstein–Uhlenbeck process were plotted in Fig. 4. On the other hand, the box plots of the MSEs of the reconstructions of all sample curves with the first three PCs and with all PCs estimated by using the four FPCA approaches for the 350 simulations are displayed in Figs. 5 and 6, respectively.

From the results of this simulation study it can be concluded that the penalized smoothing approaches (FPCA of P-splines, P-spline smoothed FPCA and regularized FPCA) provide approximations of the eigenfunctions and sample curves much more
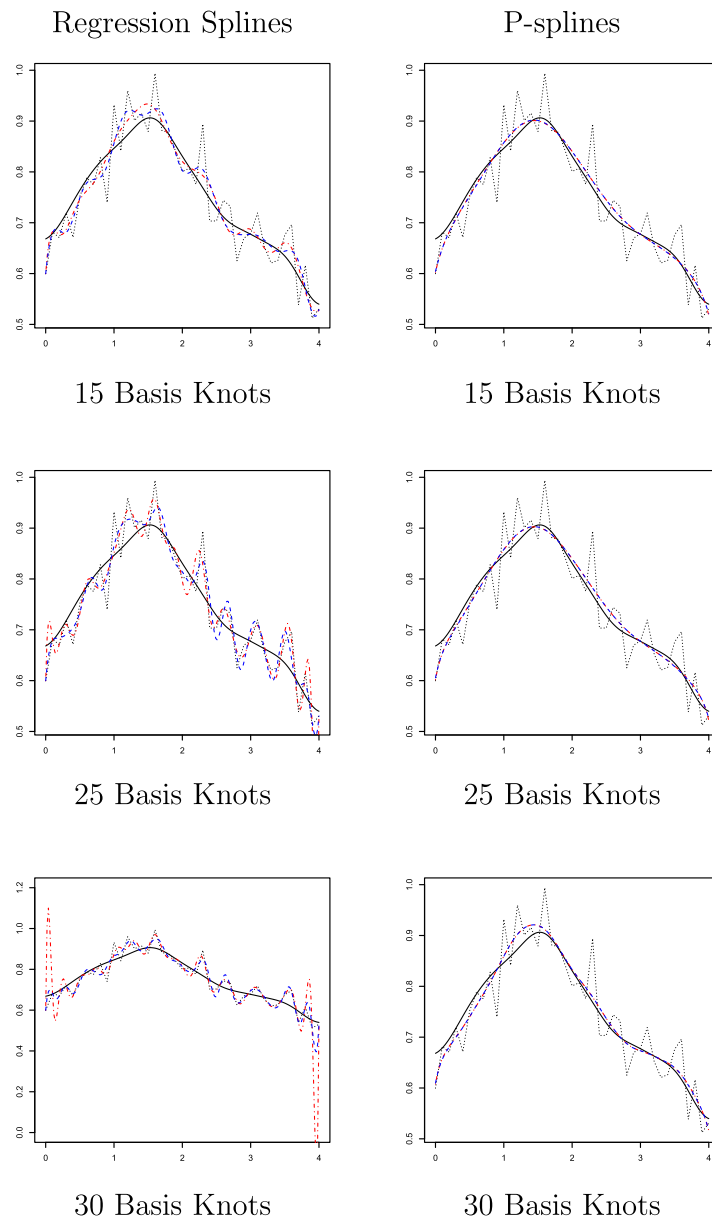
Regression Splines   P-splines

Fig. 1. Sample path of the Ornstein–Uhlenbeck simulated process approximated by regression splines (left column) and P-splines (right column) using B-splines basis of degree 3 (dashed line) and degree 4 (dashed-dotted line). The original sample paths are displayed in solid line and the noisy sample paths in dotted line.

accurate and smoother than the non penalized FPCA. This is because the regression splines used to estimate the FPCA does not control the degree of smoothness and the roughness and variability of the approximated sample paths increase dramatically with the number of knots. The results provided by the three penalized approaches are quite similar for any number of knots but they get their best behavior when a high number of knots is considered. The accuracy of the approximations provided by the two approaches based on penalizing the roughness of the principal component curves (SFPCA and RFPCA) are very similar and the best approach is FPCA of P-splines because it gives the lowest estimation errors and is computationally less expensive.

6. Real data application

Finally, we test the performance of the proposed smoothed FPCAs using the Diesel data set downloaded from the website http://software.eigenvector.com/Data/SWRI/index.html. This data set was measured by the Southwest Research Institute and consists of a training set of 133 samples and a validation set of 112 samples of the NIR transmission spectra. The NIR spectra of a Diesel sample is measured as a function of wavelengths ranging approximately from 750 to 1550 nm and was used to predict the cetane number of Diesel samples by using functional linear models [1]. In this paper, unpenalized
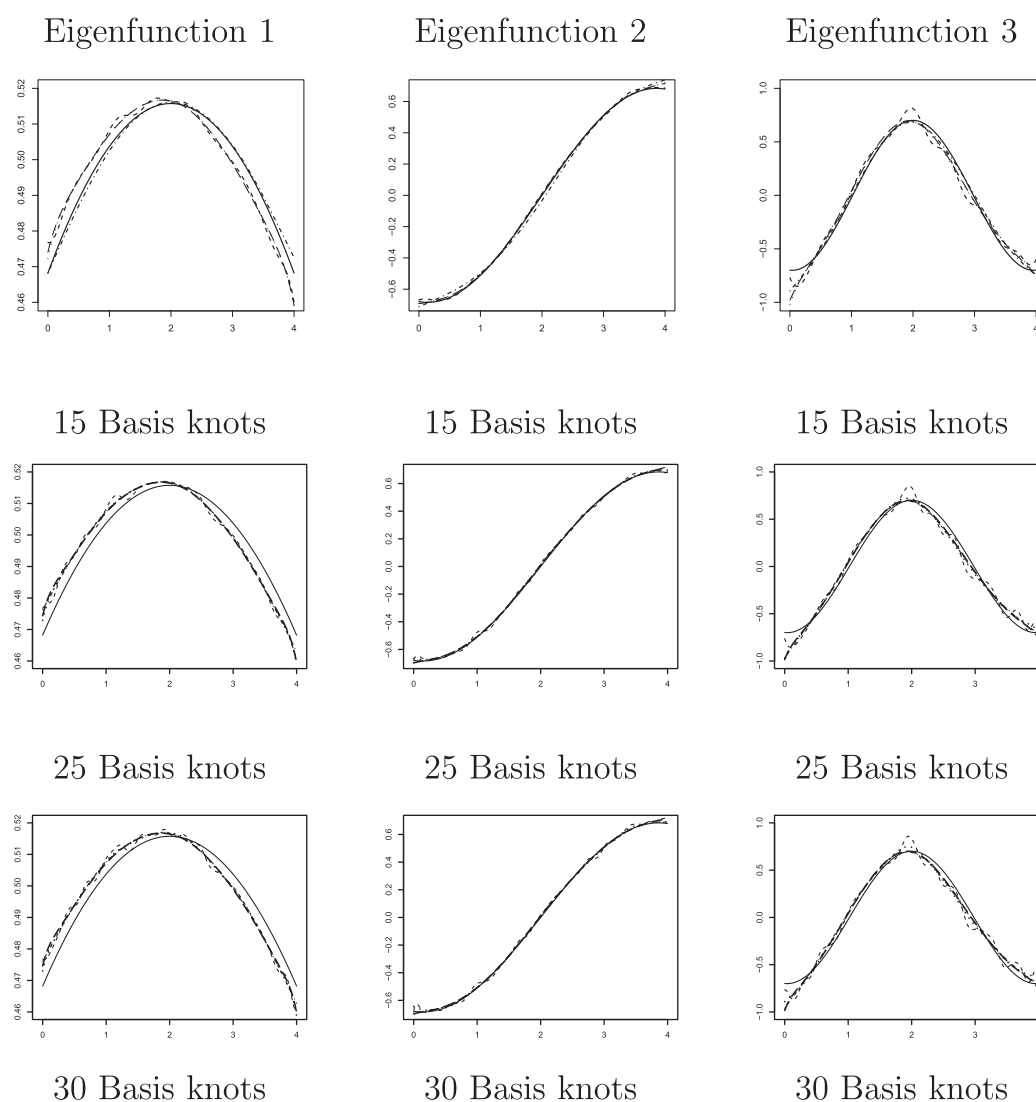
Fig. 2. First, second and third eigenfunctions of the Ornstein–Uhlenbeck process fitted by FPCA (short dashed line), FPCA via P-splines (long dashed line), SFPCA (dotted line) and RFPCA (dashed-dotted line) for 15 basis knots (λ of P-splines 2.77, λ of SFPCA 0.21 and λ of RFPCA 0.0038), 25 knots (λ of P-splines 13.92, λ of SFPCA 0.31 and λ of RFPCA 0.005) and 30 knots (λ of P-splines 24.37, λ of SFPCA 0.15 and λ of RFPCA 0.005). The true eigenfunctions are represented by black and solid line.

FPCA, FPCA of P-splines and P-spline SFPCA were performed to reduce the dimension of Diesel data and to obtain smooth reconstructions of the NIR transmission spectra. In order to test the good behavior of the smoothed FPCA based on P-spline penalties, cubic B-splines with 30 definition knots were used. The smoothing parameters selected by leave-one-out cross-validation are $\lambda = 0.05$ and $\lambda = 0,041$ for FPCA of P-splines and P-spline SFPCA, respectively.

The estimated first and second eigenfunctions are displayed in Fig. 7. It can be seen that smoothed FPCA approaches provide smoother eigenfunctions than unpenalized FPCA, especially for the second eigenfunction. The smoothest eigenfunctions are provided by FPCA of P-splines. On the other hand, it can be seen in Table 2 that the first two PCs estimated by FPCA of P-splines explain a bigger proportion of explained variance than the other two FPCA approaches. In Fig. 8 two sample paths were reconstructed with the first and second PCs estimated by each one of the FPCA approaches. All sample paths of the test sample were reconstructed with the first and second PCs estimated by using the three FPCA approaches and plotted in Fig. 9. It is clearly observed that FPCA of P-splines provides the smoothest approximation of sample curves.

7. Computational considerations

The numerical results presented in the two previous sections were obtained by using the free software R for statistical computing and graphics. The computational cost of each of the four different FPCA approaches compared in the simulation study was measured in based to the CPU times spent in their executions.
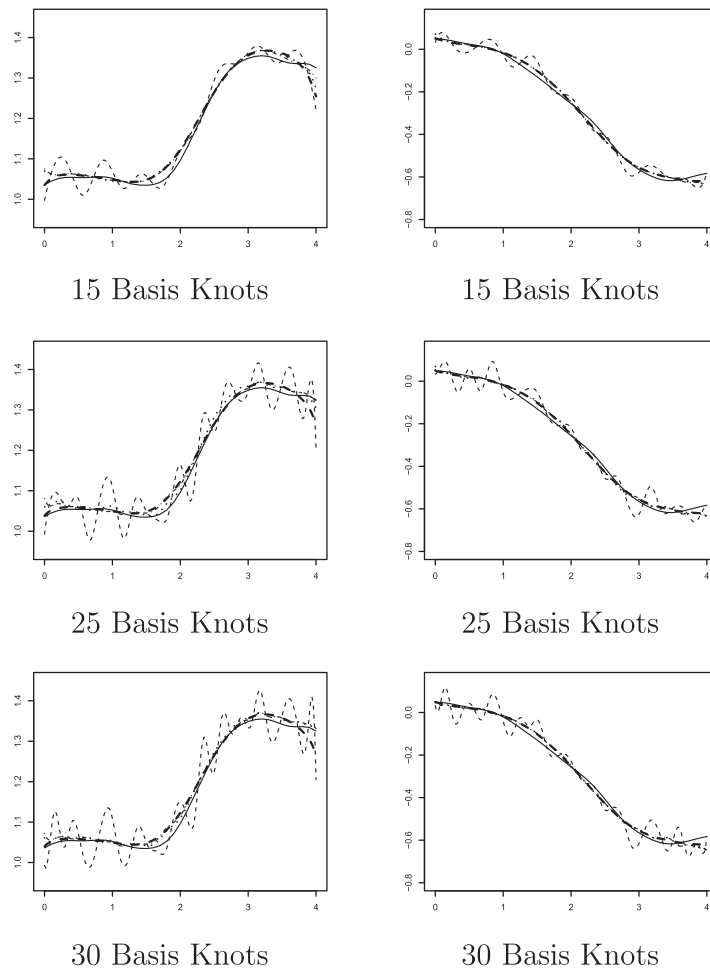
Fig. 3. Sample path (a) (left column) and sample path (b) (right column) of the Ornstein–Uhlenbeck simulated process reconstructed with the first three PCs estimated by FPCA (short dashed line), FPCA of P-splines (long dashed line), P-spline SFPCA (dotted line) and RFPCA (dashed-dotted line) for 15 basis knots (λ of P-splines 2.46(a) and 2.31(b), λ of SFPCA 0.26(a) and 0.05(b) and λ of RFPCA 0.005(a) and 0.004(b)), 25 knots (λ of P-splines 13.29(a) and 12.66(b), λ of SFPCA 0.10(a) and 0.41(b) and λ of RFPCA 0.005(a) and 0.0035(b)) and 30 knots (λ of P-splines 23.48(a) and 22.59(b), λ of SFPCA 0.92(a) and 1.84(b) and λ of RFPCA 0.0045(a) and 0.005(b)). The true original sample paths are displayed in black and solid line.

7.1. Description of R code

In this subsection, the R code generated to carry out the different FPCA approaches studied in this paper is explained. The four compared methods were implemented by using available functions of different R libraries and new code developed by the authors according to the theory described in previous sections.

1. Non penalized FPCA: This FPCA is equivalent to the multivariate PCA of the matrix $A\Psi^{\frac{1}{2}}$, with A being the matrix with columns the B-spline basis coefficients of the sample paths and $\Psi^{\frac{1}{2}}$ being the squared root of the matrix of inner products between basis functions. The R code use the following R functions:
   - The Cubic B-splines basis is obtained with the function *create.bspline.basis ()* of the *fda* package.
   - A is obtained by means of the value *coef* provided by the function *data2fd ()* of the *fda* package.
   - The matrix of inner products between basis functions Ψ is obtained with the function *inprod ()* of the *fda* package.
   - The squared root of the matrix of inner products between basis functions is computed by the singular value decomposition with the function *mroot (, method="svd")* of the *mgcv* package.
   - Finally, multivariate PCA is computed with the function *princom ()* of the *stats* package.
2. FPCA of P-splines: The only difference between FPCA of P-splines and non penalized FPCA is the matrix A. In this case, the basis coefficients of the sample curves are estimated by considering a discrete penalty based on P-splines in the least squared criterion (see Section 3 for more details). The discrete penalty matrix of order 2 (matrix $P_2 = \left(\triangle^2\right)'\triangle^2$ introduced in Section 3) is computed with the following code:

$$\triangle^2 = diff(diff(diag(nnodos + 2))),$$

where "nnodos" is the number of basis knots. Then, the matrix of P-splines basis coefficients A is computed by using Eq. 3.
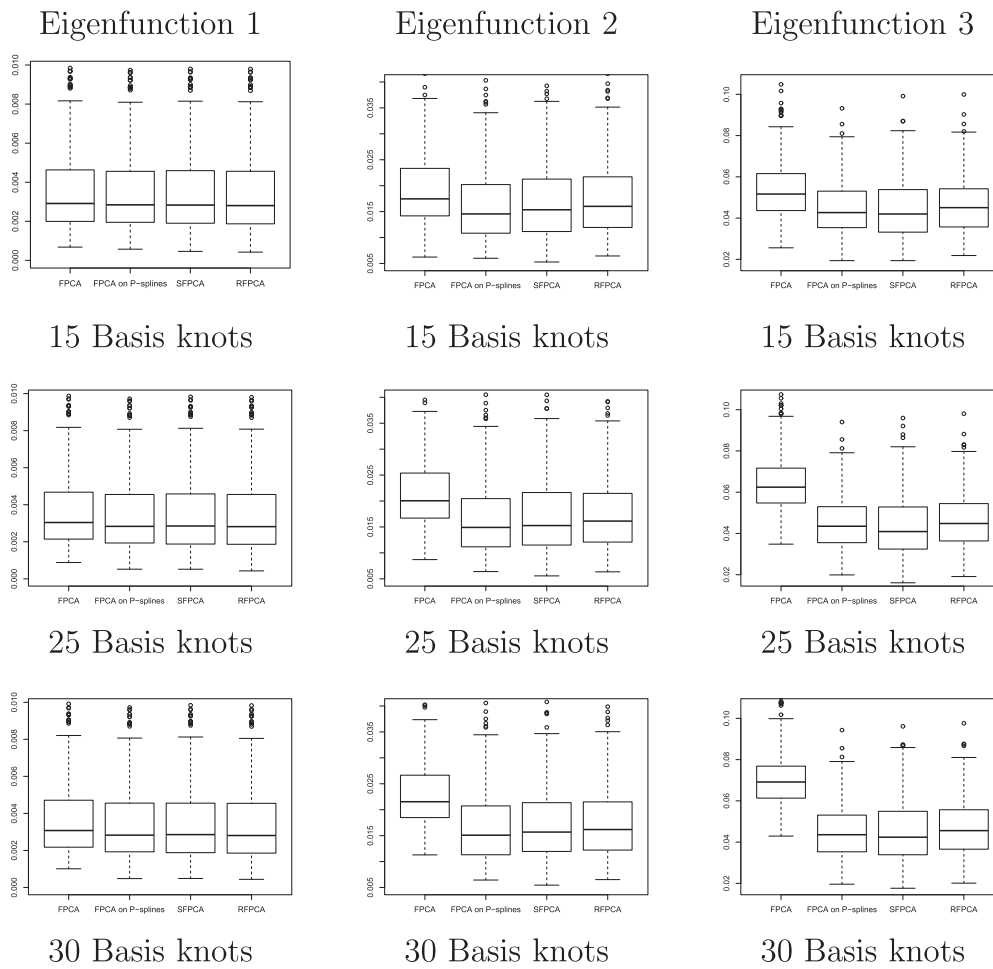
Fig. 4. Box plots of the MSEs of the eigenfunction 1 (left column), eigenfunction 2 (central column) and eigenfunction 3 (right column) estimated by using FPCA, FPCA of P-splines, P-spline SFPCA and RFPCA with 15, 25 and 30 basis knots on 350 simulations of the Ornstein–Uhlenbeck process.
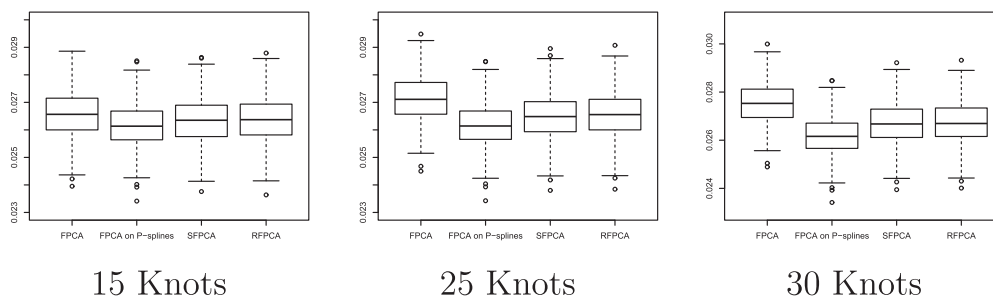

Fig. 5. Box plots of the MSEs of the reconstructions of sample curves with the first three PCs estimated by FPCA, FPCA of P-splines, P-spline SFPCA and RFPCA with 15, 25 and 30 basis knots for the 350 simulations of the Ornstein–Uhlenbeck process.

3. P-spline SFPCA: This smoothed FPCA approach turns into a multivariate PCA of matrix $A\Psi L^{-1'}$ (see Section 4 for more details). In order to compute it the following routines were carried out:
   - A is obtained by means of the value "coef" provides by the function *data2fd ()* of the *fda* package.
   - The matrix of inner products between basis functions Ψ is computed with the function *inprod ()* of the *fda* package.
   - L is a lower triangular matrix obtained by applying a Choleski factorization of the form $LL' = \Psi + \lambda P_2$, with P_2 the discrete penalty matrix previously computed for the FPCA of P-splines. The Choleski factorization was computed by means of the function *chol ($\Psi + \lambda P_2$, pivot = TRUE)* of the *base* package.
   - Finally, multivariate PCA is computed with the function *princom ()* of the *stats* package.
4. Regularized FPCA: The main difference between the computational algorithm of regularized FPCA and P-spline SFPCA is the penalty matrix that is based on the integrated squared 2-order derivative of B-spline functions. The continuous penalty matrix of order 2 is obtained by means of the function *bsplinepen (basis, Lfdobj = 2)* of the *fda* package.
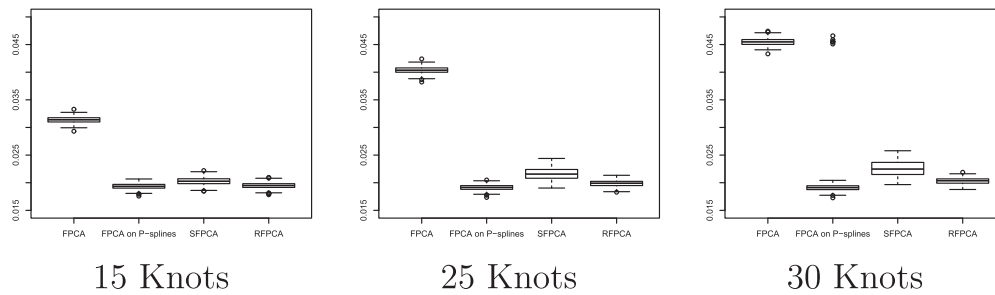
Fig. 6. Box plots of the MSEs of the reconstructions of sample curves with all PCs estimated by FPCA, FPCA of P-splines, P-spline SFPCA and RFPCA with 15, 25 and 30 basis knots for the 350 simulations of the Ornstein–Uhlenbeck process.
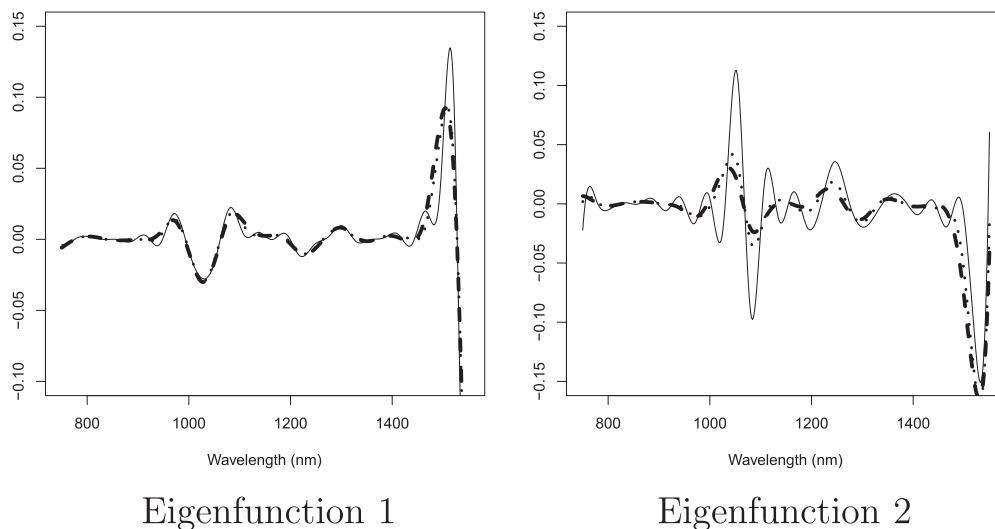

Fig. 7. First (left) and second (right) eigenfunctions estimated by FPCA (solid line), FPCA of P-splines (dashed line) and P-spline SFPCA (dotted line) for Diesel data set.

Table 2
Variances (Var) and percentages of variances (%) explained by the three considered approaches: FPCA, FPCA of P-splines and P-spline SFPCA for Diesel data set.

PC	FPCA		FPCA of P-splines		P-spline SFPCA	
	Var.	%	Var.	%	Var.	%
1	7.07×10^{-4}	85.23	4.44×10^{-4}	87.39	5.73×10^{-4}	86.65
2	6.70×10^{-5}	8.08	4.75×10^{-5}	9.35	5.40×10^{-5}	8.17

7.2. Computational cost

To select the smoothing parameter λ, a cross-validation criterion (leave-one-out) is considered in this paper. Three smoothing parameters are needed for each simulation, one for computing FPCA of P-splines, other for the P-spline SFPCA and another for regularized FPCA. From the results of the simulation study developed before, it has been concluded that there are not big differences in the approximation errors provided by the three smoothing approaches (FPCA of P-splines, P-spline SFPCA and RFPCA). However, FPCA of P-splines is preferable because its computational cost is lower and the approximation errors are slightly smaller. When we talk about computational cost, we make reference to the CPU time that a computational algorithm consumes during its execution.

In order to compare the computational cost of the cross-validation methods used to select λ with FPCA of P-splines, P-spline SFPCA and RFPCA for the 350 simulations of the Ornstein–Uhlenbeck process, the corresponding CPU times were measured by using the function *system.time* of R which returns CPU charged for the execution of user instructions of the calling algorithm. The means of the CPU times required by the different methods are shown in Table 3. It is obvious that CPU times spent by CV method with P-spline SFPCA (discrete penalty) and RFPCA (continuous penlty) are much larger than with FPCA of P-splines.
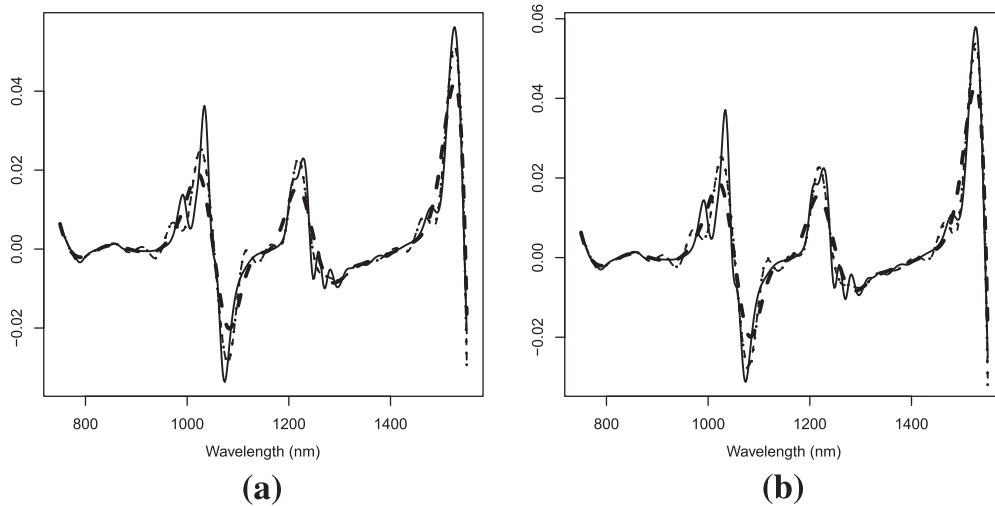
Fig. 8. Noisy sample paths (a) and (b) (solid line) reconstructed by the first and second PCs estimated by FPCA (short dashed line), FPCA of P-splines (long dashed line) and P-spline SFPCA (dotted line).
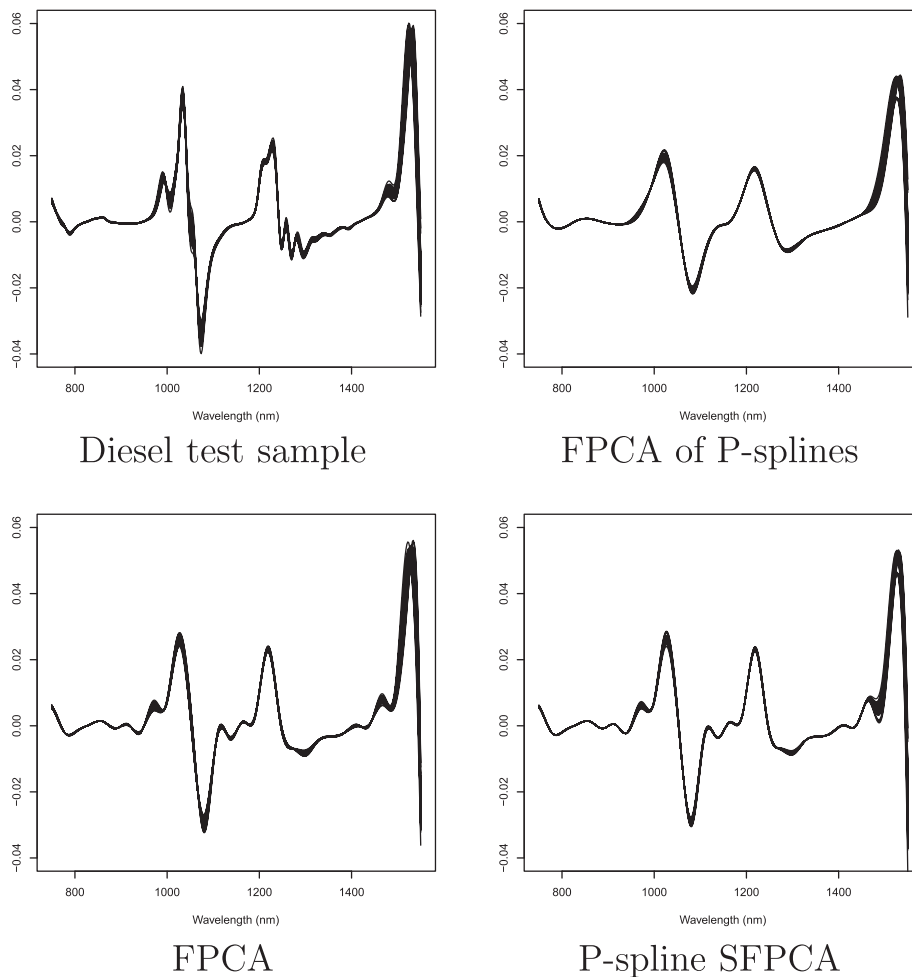

Fig. 9. Original noisy test sample of NIR spectra (top left) reconstructed with the first and second PCs estimated by FPCA (bottom left), FPCA of P-splines (top right), and P-spline SFPCA (bottom right).

All simulations have been analyzed in the same experimental conditions (computers with the same characteristics). Specifying, the simulations were run on a cluster of 30 blade servers each one with two Intel XEON *E*5420 processors running at 2.5 GHz and with 16 GB of RAM memory. Each processor has four cores and the experiments are carried out in virtualized

Table 3
Mean of the CPU times that the computational algorithms (FPCA of Psplines, P-splines SFPCA and RFPCA) consume during their execution.

FPCA of P-splines	P-spline SFPCA	RFPCA
104.79	13171.46	11056.90

Windows XP machines, each one with one virtualized processor and 1 GB of RAM memory. Until 25 Windows XP systems have been simultaneously used to carried our the cross-validation procedures.

8. Conclusions

Two smoothed FPCA approaches based on P-spline penalties have been proposed in this paper to control the degree of smoothness of the principal components weight functions estimated from smooth sample curves observed with error. Both approaches are based on B-spline expansion of sample curves and a P-spline penalty that measures the roughness of a function by summing squared d-order differences between adjacent B-spline coefficients. The first smoothed FPCA approach (called FPCA of P-splines) introduces the P-spline penalty in the least squares approximation of the sample curves with B-spline functions (P-splines) and then carries out a unpenalized FPCA on the approximated curves. The second approach approximates the sample curves by unpenalized least squares (regression splines) and then performed a penalized FPCA estimation based on maximizing a penalized sample variance that introduces the P-spline penalty in the orthonormality constraint between principal components.

A simulation study was performed to test the ability of the proposed smoothed FPCA approaches to provide an accurate and smooth estimation of the principal component curves. The results were compared with the estimations provided by non penalized FPCA of the least squares approximation of sample curves with B-spline basis and regularized FPCA based on penalizing the roughness of the principal component curves by its integrated squared 2-order derivative. From this simulation study it can be concluded that the penalized approaches give much more accurate estimations than non penalized FPCA. This is because FPCA looses control of the smoothness when the dimension of the B-spline base increases. On the other hand, the smoothed FPCA approaches are quite insensitive to the choice of knots so that a relatively large number of equally spaced basis knots is a good election for the definition of the B-spline basis. The advantage of the smoothed FPCA approaches based on P-spline penalties respect to the ones based on penalizing the integrated squared d-order derivatives is that they are mathematically simpler because the difference matrix is easier to compute than the matrix of integrals of products of d-order derivatives between B-spline basis functions (see [45] for a detailed study on the calculation of integrals involving B-splines). Finally, it can be concluded that FPCA of P-splines is preferable to P-spline SFPCA and regularized FPCA because its computational cost is lower and the approximation errors are slightly smaller.

Acknowledgements

References

[1] W. Saeys, B. De Ketelaere, P. Darius, Potential applications of functional data analysis in chemometrics, Journal of chemometrics 22 (2008) 335–344.
[2] J.O. Ramsay, B.W. Silverman, Applied Functional Data Analysis: Methods and Case Studies, Springer-Verlag, 2002.
[3] J.C. Deville, Méthodes statistiques et numériques de l'analyse harmonique, Annales de l'INSEE 15 (1974) 3–101.
[4] J. Dauxois, A. Pousse, Y. Romain, Asymptotic theory for the principal component analysis of a vector random function: some applications to statistical inference, Journal of Multivariate Analysis 12 (1982) 136–156.
[5] G.M. James, T.J. Hastie, C.A. Sugar, Principal component models for sparse functional data, Biometrika 87 (2000) 587–602.
[6] F. Yao, H.-G. Müller, J.-L. Wang, Functional data analysis for sparse longitudinal data, Journal of American Statistical Association 100 (2005) 577–590.
[7] J. Camacho, J. Picó, A. Ferrer, Data understanding with pca: Structural and variance information plots, Chemometrics and Intelligent Laboratory Systems 100 (2010) 48–56.
[8] H. Cardot, F. Ferraty, P. Sarda, Functional linear model, Statistics and Probability Letters 45 (1999) 11–22.
[9] A.M. Aguilera, O.F.A., M.J. Valderrama, Forecasting with unequally spaced data by a functional principal component approach, Test 8 (1999) pp. 233–254.
[10] A. Cuevas, M. Febrero, R. Fraiman, Linear functional regression: the case of fixed design and functional response, Canadian Journal of Statistics 30 (2002) 285–300.
[11] M. Escabias, A.M. Aguilera, M.J. Valderrama, Principal component estimation of functional logistic regression: discussion of two different approaches, Journal of Nonparametric Statistics 16 (2004) 365–384.
[12] H.G. Müller, U. Stadtmüller, Generalized functional linear models, Annals of Statistics 33 (2005) 774–805.
[13] F. Yao, H.G. Muller, J.L. Wang, Functional linear regression analysis for longitudinal data, Annals of Statistics 33 (2005) 2873–2903.
[14] T.T. Cai, P. Hall, Prediction in functional linear regression, Annals of Statistics 34 (2006) 2159–2179.
[15] A.M. Aguilera, M. Escabias, C. Preda, G. Saporta, Using basis expansion for estimating functional pls regression: applications with chemometric data, Chemometrics and Intelligent Laboratory Systems 104 (2010) 289–305.
[16] A. van der Linde, Variational bayesian functional PCA, Computational Statistics and Data Analysis 53 (2008) 517–533.
[17] F.A. Ocaña, A.M. Aguilera, M. Escabias, Computational considerations in functional principal component analysis, Computational Statistics 22 (2007) 449–465.

[18] A.M. Aguilera, R. Gutiérrez, M.J. Valderrama, Approximation of estimators in the PCA of a stochastic proces using B-splines, Communications in Statistics Simulation and Computation 25 (1996) 671–690.
[19] A.M. Aguilera, M. Escabias, M.J. Valderrama, Discussion of different logistic models with functional data: application to systemic lupus erythematosus, Computational Statistics and Data Analysis 53 (2008) 151–163.
[20] P. Besse, J.O. Ramsay, Principal component analysis of sample functions, Psychometrika 51 (1986) 285–311.
[21] P. Besse, H. Cardot, F. Ferraty, Simultaneous nonparametric regression of unbalanced longitudinal data, Computational Statistics and Data Analysis 24 (1997) 255–270.
[22] R. Viviani, G. Gron, M. Spitzer, Functional principal component analysis of fMRI Data, Human Brain Mapping 24 (2005) 109–129.
[23] M. Kayano, S. Konishi, Functional principal component analysis via regularized gaussian basis expansions and its application to unbalanced data, Journal of Statistical Planning and Inference 139 (2009) 2388–2398.
[24] M.M. Segovia-Gonzalez, F.M. Guerrero, P. Herranz, Explaining functional principal component analysis to actuarial science with an example on vehicle insurance, Insurance: Mathematics and Economics 45 (2009) 278–285.
[25] J.Z. Huang, H. Shen, A. Buja, Functional principal components analysis via penalized rank one approximation, Electronic Journal of Statistics 2 (2008) 678–695.
[26] P.H.C. Eilers, B.D. Marx, Flexible smoothing with b-splines and penalties, Statistical Science 11 (1996) 89–121.
[27] F. Yao, T. Lee, Penalized spline models for functional principal component analysis, Journal of the Royal Statistical Society: Series B 68 (2006) 3–25.
[28] A. Hyvärinen, J. Karhunen, E. Oja, Independent Component Analysis, Wiley, 2001.
[29] J.O. Ramsay, B.W. Silverman, Functional Data Analysis, Springer-Verlag, 2005.
[30] C. De Boor, A Practical Guide to Splines (revised edition), Springer, 2001.
[31] H. Kano, H. Nakata, C. Martin, Optimal curve fitting and smoothing using normalized uniform B-splines: a tool for studying complex systems, Applied Mathematics and Computation 169 (2005) 96–128.
[32] H. Kano, H. Fujioka, C.F. Martin, Optimal smoothing and interpolating splines with constraints, Applied Mathematics and Computation 218 (2011) 1831–1844.
[33] P.J. García Nieto, J. Martínez Torres, F.J. de Cos Juez, F. Sánchez Lasheras, Using multivariate adaptive regression splines and multilayer perceptron networks to evaluate paper manufactured using eucalyptus globulus, Applied Mathematics and Computation 219 (2012) 755–763.
[34] M.J. Valderrama, F.A. Ocaña, A.M. Aguilera, F.M. Ocaña-Peinado, Forecasting pollen concentration by a two–step functional model, Biometrics 66 (2010) 578–585.
[35] F. Ferraty, P. Vieu, Nonparametric Functional Data Analysis, Springer-Verlag, 2006.
[36] B.W. Silverman, Smoothed functional principal component analysis by choice of norm, Annal Statistics 24 (1996) 1–24.
[37] F. O'Sullivan, A stastical perspective on ill-posed inverse problems, Statistical Science 1 (1986) 505–527.
[38] D. Ruppert, Selecting the number of knots for penalized splines, Journal of Computational and Graphical Statistics 11 (2002) 735–757.
[39] P.H.C. Eilers, B.D. Marx, Splines, knots, and penalties, Wiley Interdisciplinary Reviews Computational Statistics 2 (2010) 637–653.
[40] P.A. Craven, G.B. Wahba, Smoothing noisy data with spline functions: Estimating the correct degree of smoothing by the method of generalized cross-validation, Numerische Mathematik 31 (1978) 377–403.
[41] H. Akaike, A new look at the statistical model identification, IEEE Transactions on Automatic Control 19 (1974) 716–723.
[42] G.E. Schwarz, Estimating the dimension of a model, Annal Statistics 6 (1978).
[43] I. Currie, M. Durban, Flexible smoothing with p-splines: a unified approach, Statistical Modelling 2 (2002) 333–349.
[44] H.L. Van Trees, Detection, estimation and modulation theory: Part I, Wiley, 1968.
[45] M. Bhatti, P. Bracken, The calculation of integrals involving b-splines by means of recursion relations, Applied Mathematics and Computation 172 (2006) 91–100.