THE NIPALS ALGORITHM FOR MISSING FUNCTIONAL DATA

CRISTIAN PREDA, GILBERT SAPORTA
and MOHAMED HEDI BEN HADJ MBAREK

Time-average approximation and principal component analysis of the stochastic process underlying the functional data are the main tools for adapting NIPALS algorithm to estimate missing data in the functional context. The influence of the amount of missing data in the estimation of linear regression models is studied using the PLS method. A simulation study illustrates our methodology.

AMS 2000 Subject Classification: 62G20, 62G08, 62G35, 62E20.

Key words: functional data, missing data, principal components, partial least squares, functional regression models.

1. **INTRODUCTION**

Statistical methods for data representing functions or curves have received much attention in the recent years. The interest for such data, known in literature as functional data, is due mainly to the difficulty to deal with infinite dimensional spaces in the context of classical multivariate methods. Examples of functional data can be found in several application domains such as medicine, economics, chemometrics and many others (Ramsay and Silverman [11]). A well accepted model for functional data is to consider it as paths of a stochastic process $X = \{X_t,\ t \in [0,T]\}$ taking values into a Hilbert space of functions on some interval $[0,T]$. For example, a second order stochastic process $X = \{X_t,\ t \in [0,T]\}$ L_2-continuous with sample paths in L_2([0,T]) can be used as model for describing the behavior of some quantitative parameter associated to a process observed on a time interval of length T.

Let us suppose that for each statistical unit of the learning sample, ω, we observe the associated curve X_ω and a single real response Y_ω. For a new unit ω' for which $X_{\omega'}$ is known, we are interested in predicting $Y_{\omega'}$ from $X_{\omega'}$. The linear functional regression model is the simplest approach to be considered and an important number of research papers in the functional data field are devoted to the estimation of the model

$$Y = \int_0^T X_t \beta(t)\mathrm{d}t + \varepsilon. \tag{1}$$

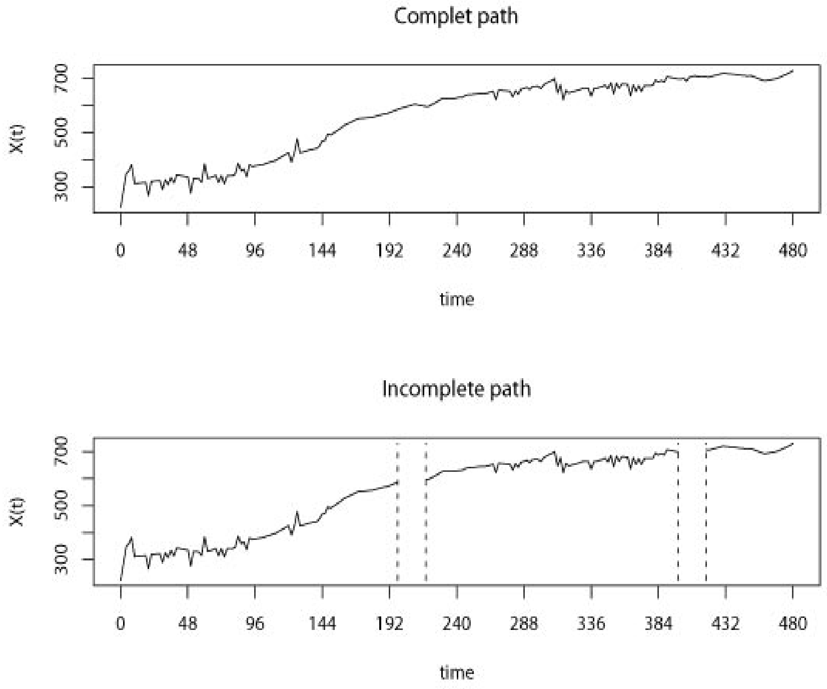
It is well known that the direct estimation of the regression coefficient function using the least square criterion yields to an ill posed problem. Solutions based on elements derived from the principal component analysis of X have been proposed by Aguilera *et al.* [1] and Cardot *et al.* [2]. These techniques are known in the literature as principal component regression (PCR). However, the choice of principal components is not an easy task, since one has to choose between robustness of the model (the most explanatories pc's) and his performances (the pc's the most correlated with the response). As an alternative to functional PCR, Preda and Saporta [9] have extended Partial Least Squares (PLS) regression to the case of a functional predictor.

But how to estimate such regression models when the functional predictor is subject to missing observation ?

If missing data is quite a common concept in finite multivariate analysis [see Little and Rubin (1987)] that is not the case for functional data. In practice, a curve is generally observed in a finite number of time points $0 = t_0 < t_1 < \cdots < t_k = T$ and thus, with missing information. However, the true form of the curve can be approximated from the points $\{(t_i, X_{t_i)},\ {}_{i=1,\ldots,k}\}$ using interpolation or smoothing procedures (Aguilera *et al.* [1]). We consider that a curve has missing data when one or several continuous part of the curve is missing, i.e., observation was not possible. This situation occurs, for example, for instruments recording curves (spectrometers, oscilloscopes) that are out of service for some short time intervals. Figure 1 provides an example of curve with missing data in two intervals of time.

The aim of this paper is to provide a methodology to estimate missing data for functional random variables, and thus, to estimate functional regression models in the context of missing data.

Our approach for dealing with missing data in the functional framework is to consider that the underlying process generating missing data is a jump stochastic process $M = \{M_t,\ t \in [0, T]\}$ with two states, $\{0, 1\}$, corresponding to the presence or absence of information. For estimating the linear model (1) in presence of missing data, we propose to use the PLS approach after time-average approximation (Preda [8]) and imputation of missing data by the NIPALS algorithm (Tenenhaus [13]). The paper is organized as follows. In Section 2 we introduce the principal component analysis for functional data and the functional linear model. The time-averge approximation and PCR and PLS regression approaches are presented. The process generating missing data as well as the methodology for applying the NIPALS algorithm for



Complet path

Incomplete path

Fig. 1. Complete and incomplete paths.

estimation of missing data is presented in Section 3. Section 4 is devoted to a simulation study.

2. PRINCIPAL COMPONENT ANALYSIS AND LINEAR REGRESSION FOR FUNCTIONAL DATA

Let us consider the functional data as sample paths of a stochastic process $X = \{X_t,\ t \in [0,T]\}$ with continuous time and Y be a random real variable defined on the same probability space as X. We assume that Y is centered and X is of second order, L_2 continuous and centered for each $t \in [0,T]$.

Also known as Karhunen-Loève expansion, the principal component analysis (PCA) of the stochastic process X consists in representing X_t as

$$X_t = \sum_{i\geq 1} u_i(t)\xi_i, \quad \forall t \in [0,T],$$

where the set $\{u_i\}_{i\geq 1}$ (the principal factors) forms an orthonormal system of deterministic functions of $L_2([0,T])$ and $\{\xi_i\}_{i\geq 1}$ (principal components) are uncorrelated zero-mean random variables. The principal factors $\{u_i\}_{i\geq 1}$ are solution of the eigenvalue equation

$$\int_0^T C(t,s)u_i(s)\mathrm{d}s = \lambda_i u_i(t),$$

where $C(t,s) = \mathrm{cov}(X_t, X_s)$, $\forall t, s \in [0,T]$. Therefore, the principal components $\{\xi_i\}_{i\geq 1}$ defined as $\xi_i = \int_0^T u_i(t)X_t\mathrm{d}t$ are eigenvectors of the Escoufier operator $\mathbf{W}^X$ (Escoufier [14]) defined by

$$\mathbf{W}^X Z = \int_0^T E(X_t Z)X_t \mathrm{d}t, \quad Z \in L_2(\Omega).$$

Notice that the process X and the set of its principal components, $\{\xi_k\}_{k\geq 1}$, span the same linear space.

2.1. Linear regression on principal components (PCR)

Under the least squares criterion, the estimation of the coefficient function β of linear regression model (1) is in general a distribution rather than a function of $L_2([0,T])$ (Saporta [12]). This difficulty appears also in practice because one has generally more predictors than the number of observations. Regression on principal components (PCR) of X (Aguilera *et al.* [1]) and PLS approach (Preda and Saporta [9]) provide efficient solutions to this problem.

As in the classical setting, the process $\{X_t\}_{t\in[0,T]}$ and the set of its principal components, $\{\xi_k\}_{k\geq 1}$, span the same linear space. Thus, the linear regression of Y on X is equivalent to the regression on $\{\xi_k\}_{k\geq 1}$ and we have $\widehat{Y} = \sum\limits_{k\geq 1} \frac{E(Y\xi_k)}{\lambda_k}\xi_k$.

In practice one has to choose an approximation of order q, $q \geq 1$:

$$\widehat{Y}_{PCR(q)} = \sum_{k=1}^{q} \frac{E(Y\xi_k)}{\lambda_k}\xi_k = \int_0^T \widehat{\beta}_{PCR(q)}(t)X_t\mathrm{d}t,$$

where

$$\widehat{\beta}_{PCR(q)} = \sum_{k=1}^{q} \frac{E(Y\xi_k)}{\lambda_k} f_k(t)$$

is the estimator of the coefficient regression function β obtained with the first q principal components.

Using the first q principal components raises a problem since they are computed independently of the response. Principal components with a great power of explanation yield generally stable models but could be uncorrelated with the response, whereas the principal components highly correlated with the response could be less explanatory for X. Moreover, for functional data, the number of principal components could be infinite. Thus, the choice of principal components is a trade-off between stability of the linear model and its predictive power (see also Escabias *et al.* [3]). A solution to this problem is the PLS approach.

2.2. **Partial least squares regression (PLS)**

The basic idea of PLS approach is to construct a set of uncorrelated random variables $\{T_i,\ i \geq 1\}$ (PLS components) in the linear space spanned by X, taking into account the correlation between Y and X. Replacing the least squares criterion with that of maximal covariance between X and Y,

$$\max_{w \in L_2([0,T])} \operatorname{cov}^2\left(Y, \int_0^T X_t w(t)\mathrm{d}t\right), \tag{2}$$

the PLS regression offers a good alternative to PCR (Preda and Saporta [9]). The first PLS component is given by $T_1 = \int_0^T X_t w(t)\mathrm{d}t$ and further PLS components are obtained by maximizing the covariance criterion between the residuals of both Y and X_t with the previous components.

The PLS approximation is given by

$$\widehat{Y}_{PLS(k)} = \sum_{i=1}^{k} c_i T_i = \int_0^T X_t \beta_{PLS(k)}(t)\mathrm{d}t. \tag{3}$$

As in the finite multivariate setting (de Jong [5]), in the functional context PLS fits closer than PCR, i.e., $R^2(Y, \widehat{Y}_{PCR(k)}) \leq R^2(Y, \widehat{Y}_{PLS(k)})$, where R is the linear correlation coefficient.

2.3. **Time average approximation**

Generally, the principal components analysis of X is realized by approximating the principal factors into a finite dimensional space of functions. One of the approximations which is convenient in presence of missing data is the time-average approximation developed in Preda [8]. This approximation, easy to put in practice, consists into approximate X by a stochastic process whose the sample paths are constant piecewise functions. If $\Delta = \{0 = t_0 < t_1 < \cdots < t_p = T\}$ is a discretization of $[0, T]$ then the time-average approximation of X is given by X^{Δ} defined as

$$X_t^{\Delta} = m_i = \frac{1}{t_i - t_{i-1}} \int_{t_{i-1}}^{t_i} X_t \mathrm{d}t, \quad \forall t \in [t_{i-1}, t_i],\ i = 1, \ldots, p. \tag{4}$$

Properties of this approximation with respect to the accuracy of the approximations provided by the elements derived from principal components analysis are presented in Preda [8]. Let observe that the principal component analysis in this case is equivalent with the principal component analysis of the set of variables $\{m_i,\ i = 1, \dots, p\}$ using as metric $\operatorname{diag}(t_1 - t_0, \dots, t_p - t_{p-1})$. The principal factors u_k of the process X are approximated by constant piecewise functions u_k^{Δ} obtained from the principal factors of the set $\{m_i,\ i = 1, \dots, p\}$ and so are for the principal components, ξ_k^{Δ}. The functional PCR regression of Y on X is then approximated by the PCR of Y on the set $\{\xi_i^{\Delta},\ i = 1, \dots, k,\ k \leq p\}$. In the same way, the PLS regression of Y on X is approximated by the PLS regression of Y on the set of variables $\{\sqrt{t_i - t_{i-1}} \times m_i,\ i = 1, \dots, p\}$ (Preda and Saporta [9]).

3. MISSING DATA FOR FUNCTIONAL DATA AND NIPALS ALGORITHM

At our knowledge, there are no works dealing with missing data for functional variables. We can observe that when the situation occurs, it is often question of the end of the curve and thus imputation of missing data is synonym of time series prediction.

3.1. Missing data model

In our approach we consider that the missing information could occur in any continuous time interval of $[0, T]$. Thus, a curve can miss information of a set of intervals $[a_1, b_1], \dots, [a_m, b_m]$. In general, the number m of such intervals is random as well as their length. One possible model for missing data in this context is to consider an underlying jump continuous time process M_t with two states, 0 and 1, defined by

$$M_t = \begin{cases} 0 & \text{if } X_t \text{ is observed at time } t, \\ 1 & \text{otherwise.} \end{cases} \tag{5}$$

Thus, for each observation $X(\omega)$ of X one has associated one observation $M(\omega)$ of M. A curve $M(\omega)$ that corresponds to the "0" constant function means that the curve $X(\omega)$ is completely observed. In the multivariate finite case, it is usually to speak about the ratio of the missing data in the whole dataset. In the functional context, we can extend this notion to the ratio of the sum of the length of the intervals $[a_i, b_i]$ within $[0, T]$ for all available curves. However, if this ratio has some interpretation when the missing data is "completely at random" (see Little and Rubin, 1987), it is difficult to justify this measure in the case of functional data.

Inspired by the reliability theory of repairable systems, we propose as measure for quantify the missing information the mean time of missing observation ($MTMO$) defined by

$$MTMO = \frac{1}{T}\int_0^T U(t)\mathrm{d}t, \tag{6}$$

where $U(t)$ is the probability that the process X is not observable at the instant t, i.e., $U(t) = P(M_t = 1)$. Obviously, the simplest model for M is a two state markovian process with exponential times for each state. Considering that $M_0 = 0$ and the rate parameters describing the system are λ (for state 0) and μ (for state 1) then one can show (Iosifescu *et al.* [6]) that

$$MTMO = \frac{\lambda}{\lambda+\mu} - \frac{\lambda}{(\lambda+\mu)^2 T} \times (1 - \mathrm{e}^{-\frac{\lambda+\mu}{T}}). \tag{7}$$

For example, for $T = 1$, $\lambda = 1$ and $\mu = 100$ we have $MTMO = 0.009802$ that means the process is unobservable about of 1% of time.

3.2. **Estimation of missing data by the NIPALS algorithm**

The NIPALS (Nonlinear Itérative Partial Least Squares) algorithm is a Jacobi-like iterative method used to estimate the elements of the principal component analysis of a finite dimensional random vector. It is interesting to note that, due to the duality between the principal factors and the principal components, this algorithm can be adapted for datasets with missing data (Tenenhaus [13]). In this context, NIPALS provides not only an estimation of principal factors and components, but also, by the mean of the data reconstitution formula, an imputation method for missing data.

Let us introduce the NIPALS algorithm in the multivariate finite dimensional case. Let $X = (X_1, X_2, \dots, X_p)'$ be a random vector of dimension p, $p \geq 1$, such that $E(X_i) = 0$, $\forall i \in 1, \dots, p$. The expansion of the vector X in terms of principal components and principal factors is a well-known result in multivariate data analysis (Escoufier [14])

$$X = \sum_{h=1}^{q} \xi_h u_h, \tag{8}$$

where $q = \dim L_2(X)$ and $\{\xi_h\}_{h=1,\dots,q}$, respectively $\{u_h\}_{h=1,\dots,q}$, are the principal components (random variables), respectively the principal factors (vectors in $\mathbf{R}^p$) of the principal component analysis of X.

If only the first r components are used in (8), $r < q$, one obtains the approximation of order r of X by

$$\widehat{X}^{(r)} = \sum_{h=1}^{r} \xi_h u_h,$$

and, for each $i \in 1, \dots, p$,

$$\widehat{X}_i^{(r)} = \sum_{h=1}^{r} \xi_h u_h(i).$$

The main idea of the NIPALS algorithm consists in the fact that for each $h = 1, \dots, q$, $u_h(i)$ represents the slope coefficient in the linear regression of the variable X_i on the component ξ_h. In the same way, if ω is an element of Ω, $\forall h = 1, \dots, q$, $\xi_h(\omega)$ represents the slope coefficient in the linear regression of the "variable" $(X_1(\omega), X_2(\omega), \dots, X_p(\omega))$ on the "variable" u_h (considered as elements of $\mathbf{R}^p$).

The input data of the NIPALS algorithm are N independent realizations of the random vector X, $N \geq 1$, as a $N \times p$ matrix with entries $x(i,j)$, $i = 1, \dots, N$, $j = 1, \dots, p$. We suppose that each column of the matrix is centered. The output is represented by N independent realizations of the q principal components and an estimate for the q principal factors.

The NIPALS algorithm

```
1. $X_0 = X$;
2. for $h = 1, 2, \dots q$,
    2.1. $\xi_h = X_{h-1}(\,\cdot\,, 1)$(the first column of $X_{h-1}$);
    2.2. repeat until convergence of $u_h$,
        2.2.1. for $i = 1, 2, \dots, p$,
            $$u_h(i) = \frac{\sum_{j: x(j,i),\, \xi_h(j) \text{ exist}} x_{h-1}(j,i)\xi_h(j)}{\sum_{j: x(j,i),\, \xi_h(j) \text{ exist}} \xi_h^2(j)};$$
        2.2.2. normalize $u_h$;
        2.2.3. for $i = 1, 2, \dots, N$,
            $$\xi_h(i) = \frac{\sum_{j: x(i,j) \text{ exist}} x_{h-1}(i,j)u_h(j)}{\sum_{j: x(i,j) \text{ exist}} u_h^2(j)};$$
    2.3. $X_h = X_{h-1} - \xi_h u'_h$.
```

If there are no missing data, the NIPALS algorithm is equivalent to the SVD Jacobi algorithm for which the convergence is well known (see for example, Golub and Van Loan [4]). In presence of missing data at random, the quality of NIPALS algorithm depends of several parameters, the most important being the distribution of X, the degree of linear dependence between the X_i's and the size of N with respect to p (see for details Preda and Duhamel [10]).

If $\{\widehat{\xi}_h\}_{h=1,\dots,q}$ and $\{\widehat{u}_h\}_{h=1,\dots,q}$ are the approximations of $\{\xi_h\}_{h=1,\dots,q}$ and $\{u_h\}_{h=1,\dots,q}$ provided by NIPALS, then $x(i,j)$ can be approximated by

$$\widehat{x}(i,j) = \sum_{h=1}^{q} \widehat{\xi}_h(i)\widehat{u}_h(j). \tag{9}$$

The explicit formula (9) defines also the approximation for missing data by the NIPALS algorithm.

Notice that if the PCA is carried out with a particular metric $M = TT'$, then NIPALS is applied to the matrix XT'.

NIPALS for functional data. Let X be a stochastic process.

Obviously, if the trajectories of X are piecewise constant functions with the same set of discontinuity points, i.e., $(X_t)_{t\in[0,T]}$ is defined by the sets $\{t_i\}_{i=0,\dots,p}$, $0 = t_0 < t_1 < t_2 < \cdots < t_{p-1} < t_p = T$ and $\{c_i\}_{i=0,\dots,p-1}$, $c_i \in L_2(\Omega)$, $\forall i = 1,\dots,p$, such that

$$X_t = c_i, \quad \forall t \in [t_i, t_{i+1}),\ \forall i = 0,\dots,p-1,$$

then the PCA of X is equivalent to the PCA of $\{c_i\}_{i=0,\dots,p-1}$ with the metric $M = \operatorname{diag}(t_1 - t_0, \dots, t_p - t_{p-1})$. If missing data occur, they are associated to the variables $\{c_i\}_{i=0,\dots,p-1}$ and the NIPALS algorithm can be applied for the estimation.

Let now consider that the sample paths of the stochastic process X are not piecewise constant but completely known except some missing data intervals. For each curve $\omega \in \Omega$ let $[a_1(\omega), b_1(\omega)], \dots, [a_{m_\omega}(\omega), b_{m_\omega}(\omega)]$ be the set of intervals corresponding to missing data (eventually empty).

In order be apply the NIPALS algorithm, we define $\Delta = 0 = t_0 < t_1 < \cdots < t_p = T$ as an equidistant discretization of $[0,T]$ with $\delta = t_{i+1} - t_i$ such that the length of each missing data interval $[a_i(\omega), b_i(\omega)]$, $i = 1,\dots,m_\omega$, $\omega \in \Omega$, is multiple of δ. Then, using this discretization, the time average approximation defined by (4) yields a set of random variables $m_i,\ i = 1,\dots,p$, $p = \frac{T}{\delta}$, which has missing data if there exists i, $i = 1,\dots,p$, such that $i \times \delta$ belongs to a missing data interval. The application of the NIPALS algorithm to $m_i,\ i = 1,\dots,p$, provides approximations for the principal component analysis of X. For each missing time interval, data is estimated by the time-average approximation given by (9). More details on the choice of δ can be found in Preda [8].

4. SIMULATION STUDY

In this study missing data are estimated by the NIPALS algorithm after time average approximation of sample paths. The quality of this estimation procedure is measured when linear models with functional predictor are fitted.

Let Y be a real random variable defined by the linear regression model

$$Y = \int_0^1 X_t \beta(t) \mathrm{d}t + \varepsilon, \tag{10}$$

where X is the standard Brownian motion on $[0, 1]$, $\beta(t) = 3t^3$, $t \in [0, 1]$ and ε is the error term such that $V(\varepsilon) = 0.1$. Notice that $V(Y) = 0.5$ and $R^2 = 0.8$. We generate $n = 100$ curves representing the sample paths of X observed on a discretization Δ_1 of the interval $[0, 1]$ in 1000 equidistant intervals. The Simpson quadrature method provides the values of Y for each curve.

The estimation of the linear model (10) is realized using the PLS approach after time-average approximation with a discretization Δ_2 of 100 equidistant intervals. The PLS regression coefficient function, $\widehat{\beta}_{PLS}$, is obtained with three PLS components (cross-validation) and is represented in Figure 2.

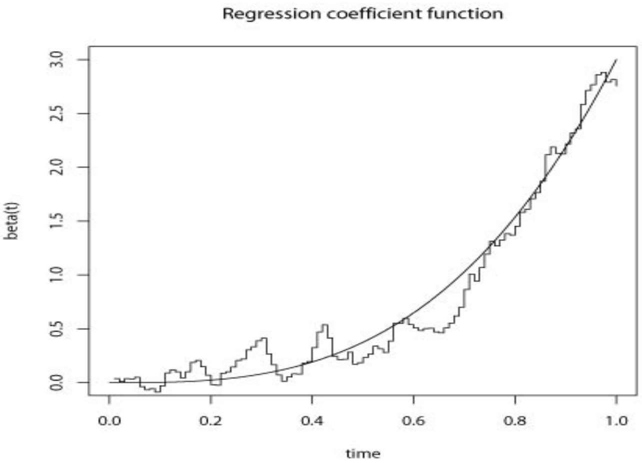


Fig. 2. Regression coefficient function, $\beta(t) = 3t^3$, and the PLS estimate (step function) with complete data.

Missing data are now simulated for several values of $MTMO$. We use for simulation a two state Markovian jump process with exponential times for the two states as in (7). The exponential distribution is simulated with a precision of 1/1000, thus the change points belong to Δ_1. The size of the discretization for the time average approximation is the *gcd* of the missing interval lengths as multiple of 1/1000. In Table 1 are presented the performances (measured by R^2) of the PLS regression after the estimation of missing data by the NIPALS algorithm for several values of $MTMO$ (λ and μ). For each value of $MTMO$, the value of R^2 is averaged over 50 independent samples.

Table 1. Performances of the PLS regression with missing data

Parameters	MTMO	R^2
complete data	0	0.7645
$\lambda = 1,\ \mu = 100$	0.00980	0.7263
$\lambda = 1,\ \mu = 50$	0.01922	0.7288
$\lambda = 1,\ \mu = 20$	0.04535	0.7144
$\lambda = 2,\ \mu = 20$	0.08677	0.6625
$\lambda = 2,\ \mu = 10$	0.15277	0.6218
$\lambda = 2,\ \mu = 5$	0.24493	0.4872

One can observe on this example that small amounts of missing data ($MTMO < 5\%$) have not significant influence on the quality fit of the model (less than 1% of R^2), whereas for large amounts the quality of fit degenerates rapidly (up to 37% of R^2 for $MTMO = 24.4\%$).

5. CONCLUSION

Based on the estimation of simple linear regression models with missing data, the NIPALS algorithm provides estimations for the principal component analysis of X. The expansion formula of X in terms of principal factors and components allows estimation of missing data. A simulation study using as model for missing data a two-state Markovian process shows the influence of the amount of missing data on the quality fit of a linear regression model with functional predictor.

REFERENCES

[1] A.M. Aguilera, F. Ocana and M.J. Valderrama, *An approximated principal component prediction model for continuous-time stochastic process.* Appl. Stoch. Models Data Anal. **13** (1997), 61–72.

[2] H. Cardot, F. Ferraty and P. Sarda, *Functional linear model.* Statist. Probab. Lett. **45** (1999), 11–22.

[3] M. Escabias, A.M. Aguilera and M.J. Valderrama, *Principal component estimation of functional logistic regression: discussion of two different approaches.* J. Nonparametric Statist. **16** (2004), 365–384.

[4] G.H. Golub and C.F. Van Loan, *Matrix Computation*, Third Edition. The John Hopkins University Press, 1996.

[5] S. de Jong, *PLS fits closer than PCR.* Chemometrics **7** (1993), 551–557.

[6] M. Iosifescu, N. Limnios and G. Oprisan, *Modèles stochastiques.* Hermes Science, 2007.

[7] R.J.A. Little and D.B. Rubin, *Statistical Analysis with Missing Data.* Wiley, New York, 1987.

[8] C. Preda, *Approximation par moyennage de l'analyse en composantes principales d'un processus stochastique.* C.R. Math. Acad. Sci. Paris **330** (2000), 605–610.

[9] C. Preda and G. Saporta, *PLS regression on a stochastic process.* Comput. Statist. Data Anal. **48** (2005), 149–158.

[10] C. Preda and A. Duhamel, *Tools for statistical analysis with missing data. An application to a large medical dataset.* Studies in Health Technology and Informatics **116** (2005), 181–186.

[11] J.O. Ramsay and B.W. Silverman, *Applied Functional Data Analysis: Methods and Case Studies.* Springer Series in Statistics. Springer, New York, 2002.

[12] G. Saporta, *Méthodes exploratoires d'analyse de données temporelles.* Cahiers B.U.R.O., pp. 37–38. Univ. Pierre et Marie Curie, Paris, 1981.

[13] M. Tenenhaus, *La régression PLS. Théorie et pratique.* Editions Technip, 1998.

[14] Y. Escoufier, *Echantillonnage dans une population de variables aléatoires réelles.* Publ. Inst. Statist. Univ. Paris **19** (1970), 1–47.

Received 19 April 2010

Université Lille I
Laboratoire P. Painlevé, UMR 8524 CNRS
Bât. M2, Cité Scientifique
F-59655 Villeneuve d'Ascq Cedex, France
cristian.preda@polytech-lille.fr

CNAM, Chaire de Statistique Appliquée & CEDRIC
292, Rue Saint-Martin
75141 Paris Cedex 03, France
gilbert.saporta@cnam.fr

and

Institut Supérieur de Gestion de Sousse
Tunisie
benmbarekmhedi@yahoo.fr