

Data Reduction by means of Principal Differential Analysis: an Application to the Study of the Geometrical Features of the Internal Carotid Artery

Matilde Dalla Rosa, Laura M. Sangalli, Simone Vantini

Abstract In this work we show how in some cases, principal differential analysis (PDA) can be an effective tool - alternative to functional principal component analysis (FPCA) - for performing a dimensional reduction of a functional data set. In particular, we compared the results provided by PDA and FPCA in the dimensional reduction of three synthetic datasets and of the AneuRisk data set. The latter involves the geometric features of 65 internal carotid artery (ICA) of 65 patients under observation within the AneuRisk Project¹. The analyses of the synthetic datasets show how PDA, when effective as a dimensional reduction tool, can provide a representation of functional data more easily interpretable than FPCA and that this representation is not affected, differently by FPCA, by the presence of clusters or strong correlations among the original components. Moreover, in the analysis of the AneuRisk dataset, PDA is able to detect important features of the data that FPCA is not able to detect.

Key words: Functional data analysis, Dimensionality reduction.

Matilde Dalla Rosa

Eni S.p.A. Divisione E&P, San Donato M.se, Italy, e-mail: matilde.dalla.rosa@eni.it

Laura M. Sangalli and Simone Vantini

MOX - Department of Mathematics “Francesco Brioschi”, Politecnico di Milano, Piazza Leonardo da Vinci, 32, 20133, Milano, Italy, e-mail: laura.sangalli@polimi.it, simone.vantini@polimi.it

¹ The project involves MOX (Dip. di Matematica, Politecnico di Milano), LaBS (Dip. di Ingegneria Strutturale, Politecnico di Milano), Istituto Mario Negri (Ranica), Ospedale Niguarda Ca’ Granda (Milano), and Ospedale Maggiore Policlinico (Milano), and is supported by Fondazione Politecnico di Milano and Siemens Medical Solutions Italia.

1 Introduction

The analysis here presented has been developed within the AneuRisk project. This is a joint project involving scientists of different fields that aims at providing evidence of a relation between aneurysmal pathologies, and the geometry and the hemodynamics of brain vessels. The interest in studying cerebral aneurysms (i.e. malformations of cerebral vessels characterized by a bulge of the vessel wall) is due to the fact that the rupture of a cerebral aneurysm is usually a very severe event, taking the patient to death in 50% of the cases and causing various permanent damages in most of the other cases.

Aim of this analysis is to test a conjecture grounded on practical experience of neuroradiologists at Niguarda Ca Granda Hospital (E. Boccardi, personal communication): “cerebral arteries of patients with an aneurysm at or after the terminal bifurcation of the internal carotid artery (ICA) show peculiar geometric features”.

This conjecture has already been tested in the two exploratory works [7, 8]. Aim of the present work is to confirm or deny conclusions that have been drawn in [7, 8]. In particular, we want to explore the same problem from the new perspective that is offered by the use of Principal Differential Analysis (PDA) as a dimensional reduction tool. The basic idea is to reduce data dimensionality, not only taking into account the observed functions, but also their derivatives (i.e. linear combinations of them).

2 PDA as a Dimensional Reduction Tool

PDA is a technique that enables to estimate a differential operator from a functional dataset. This technique has already been used to analyze various types of applications such as the study of free handwriting ([4]), the analysis of the movement of the lips during the release of the syllable Bob ([3]), economic models ([9]), weather ([5]) and chemical models ([2]). In these papers, the main goal of PDA was the estimation of the unknown parameters of a well known underlying differential operator. In our case, the operator is instead a tool for achieving a better understanding of the phenomenon. In fact, the estimated operator becomes a tool for obtaining a convenient representation of the data. It provides a finite-dimensional space onto which the data can be projected and that enables to explore the variability related to linear relations among derivatives.

In order to estimate a linear differential operator L of order m with constant coefficients by means of PDA, the observed functions x_i , $i = 1, \dots, N$, and their derivatives - up to the m order - need to be contained in the Hilbert space $L^2(a, b)$, with (a, b) being the interval where the data are observed. Given this assumption, a natural procedure to estimate the operator L is to choose the estimate $\hat{L}$ such that it represents, at best, the linear relation between the data x_i and their derivatives $D^j x_i$, $j = 1, \dots, m$. More formally, it means finding, among all linear differential operators L of the form

$$Lx_i = D^m x_i - \beta_{m-1} D^{m-1} x_i - \dots - \beta_1 D x_i - \beta_0 x_i, \quad (1)$$

the one (i.e. $\hat{L}$) that minimizes the sum of squared differential residuals:

$$RSS(L) = \sum_{i=1}^N \|Lx_i\|_{L^2(a,b)}^2. \quad (2)$$

The minimization problem (2) is solved for the value $\hat{\beta} = -\mathbf{R}^{-1}\mathbf{s}$ of the parameter vector $\beta = (\beta_0, \beta_1, \dots, \beta_{m-1})'$, with $(\mathbf{R})_{j_1 j_2} = \sum_{i=1}^N \langle D^{j_1} x_i, D^{j_2} x_i \rangle_{L^2(a,b)}$ and $(\mathbf{s})_{j_1} = \sum_{i=1}^N \langle D^{j_1} x_i, D^m x_i \rangle_{L^2(a,b)}$ with $j_1, j_2, \dots \in \{0, 1, \dots, m-1\}$.

The estimation of the linear operator L and the Partitioning Principle of Hilbert spaces (e.g. [6]), jointly enable the orthogonal decomposition of the functional Hilbert space which the data belong to, into two different components $\ker(\hat{L})$ and $\ker(\hat{L})^\perp$, where $\ker(\hat{L})$ is the m -dimensional linear space of all functions $\hat{x}$ satisfying the linear differential relation $\hat{L}\hat{x} = 0$ and $\ker(\hat{L})^\perp$ is its orthogonal counter part. This means that for $i = 1, \dots, N$, the function x_i can be univocally decomposed in the sum of two orthogonal components (with respect to the usual $L^2(a,b)$ inner product) $\hat{x}_i$ and $\hat{e}_i$ such that $\hat{L}\hat{x}_i = 0$ and $\hat{L}\hat{e}_i = \hat{L}x_i$.

The term $\hat{x}_i$ is named the structural component of x_i , i.e. the m -dimensional component of the function x_i that can completely be explained by means of the linear differential relation $\hat{L}\hat{x}_i = 0$, and $\hat{e}_i$ is named the residual component of x_i , i.e. the ∞ -dimensional component of the function x_i that makes x_i not satisfy $\hat{L}x_i = 0$.

The partitioning principle together with the well known result (e.g. [6]) that $\ker(\hat{L}) = \text{span}\{e^{\hat{\lambda}_j s}\}_{j=1, \dots, m}$, with $\hat{\lambda}_j$ being the m different complex roots of the characteristic polynomial $\lambda^m - \hat{\beta}_{m-1}\lambda^{m-1} - \dots - \hat{\beta}_1\lambda - \hat{\beta}_0$, makes the computation of $\hat{x}_i$ and $\hat{e}_i$ straightforward. Indeed, due to the orthogonality of $\hat{x}_i$ and $\hat{e}_i$, the structural component $\hat{x}_i$ turns out to be the solution of the following minimization problem:

$$\min_{x \in \ker(\hat{L})} \|x - x_i\|_{L^2(a,b)}^2 \quad \text{for } i = 1, \dots, N, \quad (3)$$

that because of the finite dimensionality of $\ker(\hat{L})$ reduces to $\hat{x}_i(s) = \sum_{j=1}^m \hat{c}_{ij} e^{\hat{\lambda}_j s}$ with $(\hat{c}_{i1}, \hat{c}_{i2}, \dots, \hat{c}_{im})' = -\bar{\mathbf{R}}^{-1} \bar{\mathbf{s}}_i$, where $(\bar{\mathbf{R}})_{j_1 j_2} = \langle e^{\hat{\lambda}_{j_1} s}, e^{\hat{\lambda}_{j_2} s} \rangle_{L^2(a,b)}$ and $(\bar{\mathbf{s}}_i)_{j_1} = \langle e^{\hat{\lambda}_{j_1} s}, x_i(s) \rangle_{L^2(a,b)}$. Finally, $\hat{e}_i = x_i - \hat{x}_i$.

A useful tool to measure the effectiveness of the obtained dimensional reduction is the quantity RSQ introduced in [5]. Under the assumption of constant coefficients, RSQ is a ratio between the structural variability and the overall variability. Indeed $RSQ = 1$ when there is only structural variability (i.e. $x_i = \hat{x}_i$ for $i = 1, 2, \dots, N$) and $RSQ = 0$ when there is no structural variability (i.e. $x_i = \hat{e}_i$ for $i = 1, 2, \dots, N$).

A couple of theoretical remarks about the dimensional reduction achieved by means of the projection of functional data onto $\ker(\hat{L})$ need to be mentioned.

The first remark pertains to the link between PDA and functional regression. Note that, while in the minimization problem (3), N actual functional regression

analyses are performed (the only unusual feature with respect to a traditional functional regression is that the regressors are not known but estimated in a previous stage of the analysis), in the minimization problem (2), the link is instead just formal and not pertaining to the modeling. Indeed, in (2), the functional “regressors” $D^j x_i$, $j = 0, \dots, m-1$, are not deterministic and the random functional “error” Lx_i is not independent from the functional “regressors”.

The second remark pertains instead to the link between PDA and functional principal component analysis (FPCA). Differently from FPCA, where the dimensional reduction is driven just by the punctual values of the functional data along the domain (a, b) , the dimensional reduction obtained by PDA is driven by the values of linear combinations of the first m derivatives along the domain (a, b) . By definition, FPCA is expected to provide an effective dimensional reduction in any situation where most of the functional variability is expressed within a finite dimensional subspace; at the same time a simple analytical expression of the principal components is often missing and the interpretation of these components is often non-trivial. On the contrary, PDA is expected to provide an effective dimensional reduction only in those situations where most of the functional variability is expressed within some particular finite dimensional subspaces, i.e. the ones generated by some functions of the form $e^{(\alpha + \omega i)s}$ with α and $\omega \in \mathbb{R}$; despite of its reduced applicability, PDA always provides very clear results that are easily interpretable in terms of constant, exponential, sinusoidal, or damped-sinusoidal functions.

Since the effectiveness of PDA is related to particular finite dimensional subspaces, it is straightforward that the dimensional reduction provided by FPCA is always more effective or at least comparable with the one provided by PDA. It is thus clear that in all situations where FPCA provides an effective dimensional reduction while PDA does not, the use of the latter is not suggested. It is more interesting to determine if PDA can be a useful tool to have a better insight of the functional variability when both PDA and FPCA provide an effective dimensional reduction. To unravel this doubt, in Section 3, we compare the results of PDA and FPCA in the analysis of three very simple synthetic datasets that can be effectively reduced by both techniques (i.e. datasets made of sinusoids of fixed frequency with random amplitude, horizontal and vertical shifts). In particular, we will show that the presence of strong correlations, or of clusters, strongly depletes the interpretability of the principal components making them unable to reveal the simple nature of the phenomenon while it is not affecting the results of PDA.

3 Comparing FPCA and PDA of three synthetic datasets

In this section we perform the dimensional reduction of three synthetic datasets (Figure 1), made of $N = 200$ functions generated according the following model:

$$y_i(s) = a_i + b_i \sin\left(\frac{2\pi}{10}s\right) + c_i \cos\left(\frac{2\pi}{10}s\right), \quad (4)$$

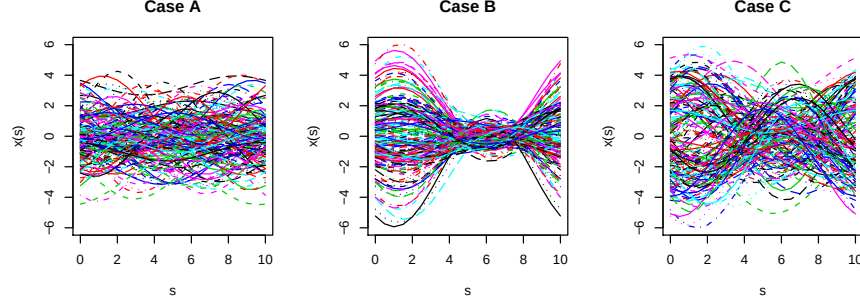


Fig. 1 Synthetic datasets.

with $s \in [0, 10]$, $i = 1, \dots, 200$, and the random coefficients $(a_i \ b_i \ c_i)' \sim iid \mathcal{N}_3(\mu, \Sigma)$:

Case A: $\mu = \mathbf{0}$ and $\Sigma = \mathbf{I}$;

Case B: $\mu = \mathbf{0}$ and $\Sigma = \begin{pmatrix} 1 & 0.9 & 0.9 \\ 0.9 & 1 & 0.9 \\ 0.9 & 0.9 & 1 \end{pmatrix}$;

Case C: $\mu = \mathbf{1}$ for $i = 1, \dots, 100$, $\mu = -\mathbf{1}$ for $i = 101, \dots, 200$, and $\Sigma = \mathbf{I}$.

Note that $\{1, \sin(\frac{2\pi}{10}s), \cos(\frac{2\pi}{10}s)\}$ is an orthogonal basis for the three-dimensional space which all functions belong to, and that this space coincides with $ker(D^3 + (\frac{2\pi}{10})^2 D)$. It is thus obvious that both the projection on the first three sample principal components and the projection on the kernel of a third-order operator estimated as in (2), are expected to be effective dimensional reduction tools, and indeed they are. In Figure 2 the basis proposed by FPCA is compared with the basis detected by PDA in the three cases.

In case A, the non-correlation of the original components makes the basis detected by FPCA almost identical to the basis detected by PDA with the first component nearly constant and the remaining two components associated to periodic oscillations.

In case B, due to the strong correlation among original components, the basis detected by FPCA drifts away from the basis detected by PDA that instead remains unchanged. The constant and sinusoidal terms are indeed confounded in a unique component explaining more than 90% of the total variability. If the model behind the data had not been known, and FPCA had been used to perform dimensional reduction, probably just the first principal component would have been taken in to account and interpreted as describing the amplitude variability within a semi-period heuristically identified, and the simple description of functions as a sum of a constant term and a sinusoidal term would have probably remained unrevealed.

In case C, a strong correlation among original components is artificially induced by the presence of two well separated clusters of data. As expected, the first principal component detects the direction connecting the two clusters. This direction is not associated to just one of the original components and thus, similarly to the former case, along this principal component all three original components are put together

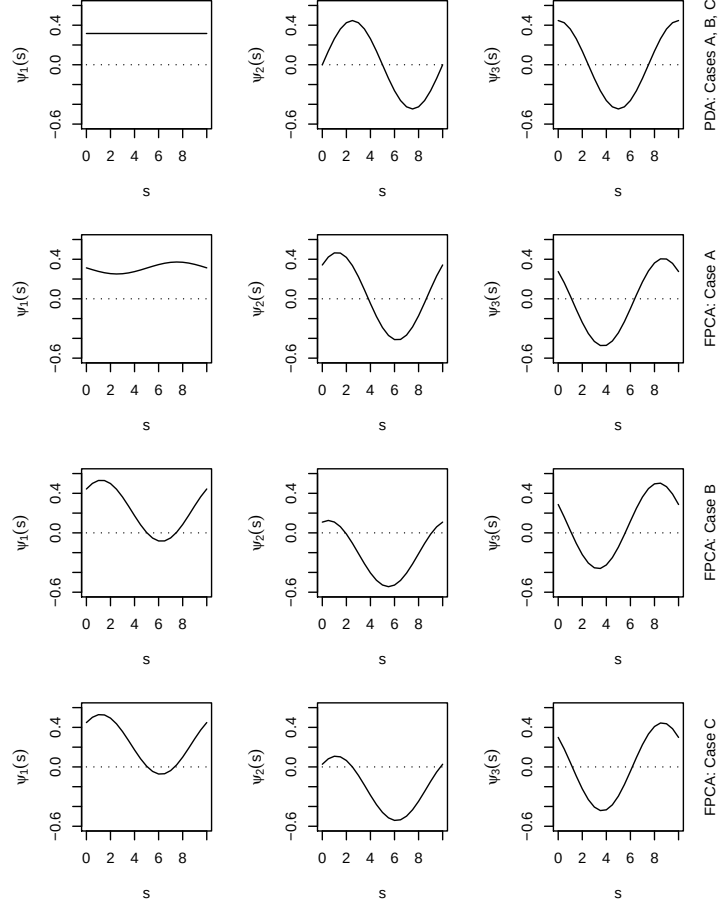


Fig. 2 First row: basis functions ψ_1 , ψ_2 , ψ_3 detected by PDA in case A, B, and C (results are graphically undistinguishable). Second, third, and fourth rows: basis functions (principal components) detected by FPCA in case A, B, and C respectively.

masking once again the simple description of functions as as sum of a constant term and a sinusoidal term.

Note that PDA provides explicit estimates of the frequency of the sinusoidal component while in FPCA the latter can be estimated just heuristically by comparing subsequent maxima. This is not a big issue in this ideal case where no error component is added, it could become an issue in favor of PDA in real applications where a 100% effective dimensional reduction is hardly achievable. On the other hand, differently from FPCA, PDA necessitates the estimate of the first m derivatives of each function and not just of the function itself. Once again this is not an issue in these synthetic cases where derivatives are known exactly, but in real applications an efficient estimation of derivatives could be tough (as shown in [7]).

On the whole, these easy examples show how PDA, when able to effectively reduce functional data dimensionality, can be able to provide an alternative and easily interpretable representation of functional data in terms of constant, exponential, sinusoidal, or damped-sinusoidal functions.

4 FPCA of Internal Carotid Artery Radius

In this section we briefly sketch out the study of the radius of the inner carotid artery performed by means of FPCA in [8] to enable the comparison with the new PDA-based analysis.

The AneuRisk dataset is based on three-dimensional angiographies of 65 patients hospitalized at the Neuroradiology Department of the Niguarda Ca'Granda Hospital in Milan. According to the location of the aneurysm and to the conjecture introduced in Section 1, the 65 patients can be divided into two groups:

Lower group: thirty-two patients that are healthy or affected by aneurysm before the terminal bifurcation of the ICA;

Upper group: thirty-three patients that are affected by aneurysm at the ICA terminal bifurcation or after it.

The raw dataset consists of 65 sets of measures of the ICA radius (one set for each patient). For each patient, the ICA radius is sampled along a very dense discrete grid, differing from patient to patient, that is indexed by means of an approximated curvilinear abscissa associated to the vessel centerline ([1]). Such data are clearly functional in nature, although they are sampled in a discrete way. An accurate functional representation of the radii and of their derivatives, is obtained by means of free-knot regression splines of order five. This method provides in fact accurate estimates of derivatives, as it is shown in [7], where free-knot splines are generalized to the case of multidimensional curves. Moreover, in order to enable meaningful comparisons across patients, the 65 ICA radius functions are then registered by means of the alignment procedure described in [8].

In [8] the first two principal components (explaining the 66% and 13% of the total variance, respectively) result amenable of a biological interpretation. The first one is intended as an average size of the carotid (distinguishing between narrow and wide ICAs), while the second one as a tapering factor (distinguishing between cylindric and conic ICAs).

From the analysis of the associated scores, it appears that the patients belonging to the Upper group present more tapered and wider carotids, with a lower within-group variability in the size of the radius than the Lower group.

Notice that radii are available on different abscissa intervals, since they depend on the point where the angiography was centered. Thus both the FPCA of [8] and the PDA of the present work are carried out only on the common portion of the 65 abscissa intervals, where the records are present for all patients. Such portion is highlighted in Figure 3.

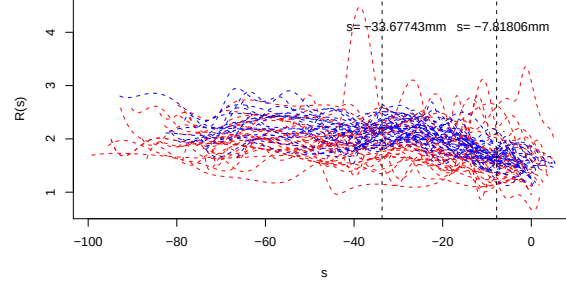


Fig. 3 The 65 ICA radius functions after registration. Red curves are relative to the Lower group, while blue curves refer to the Upper group. The common interval of definition is highlighted by means of vertical dashed lines.

5 PDA of Internal Carotid Artery Radius

In this section we present the PDA of the 65 internal carotid artery radius functions introduced in Section 4. As already stated, the focus is on the possible relations between the patient's ICA radius and his membership to the Lower or Upper group.

The first two problems that need to be jointly solved are (i) finding the order m of the linear operator $\hat{L}$, compromising a satisfactory goodness of fit with an easy interpretability of the results and (ii) finding out if one operator alone (i.e. $\hat{L}^{Tot}$) is able to explain the variability within the two groups or if two different operators are needed (i.e. $\hat{L}^{Low}$ and $\hat{L}^{Up}$).

In the following table, the RSQ achieved by performing a PDA with $m = 2$ and $m = 3$ for the Lower, the Upper, and the two groups together are reported:

RSQ	$\hat{L}^{Low}$	$\hat{L}^{Up}$	$\hat{L}^{Tot}$
$m = 2$	< 1%	< 1%	< 1%
$m = 3$	36%	29%	33%

It is evident from the previous table that linear differential operators of order two are not able to explain any significant portion of the functional variability of any of the two groups. On the other hand, linear differential operators of order three are able to explain nearly one third of the overall variability in both groups. This means that nearly one third of the functional variability lies within a three-dimensional space, i.e. the kernel of the corresponding operator. We choose $m = 3$ to be the suitable order of the linear differential operator and we decide not to move to a higher order to avoid the use of possibly inaccurate estimates of high order derivatives.

Focusing on third order operators, the minor difference between the RSQ associated to $\hat{L}^{Low}$ and $\hat{L}^{Up}$ leaves space for the possibility that $\hat{L}^{Tot}$ can jointly explain the structural variability of both groups. Since no inferential tool is available to test this hypothesis, we heuristically proceed by comparing the roots of two characteristic

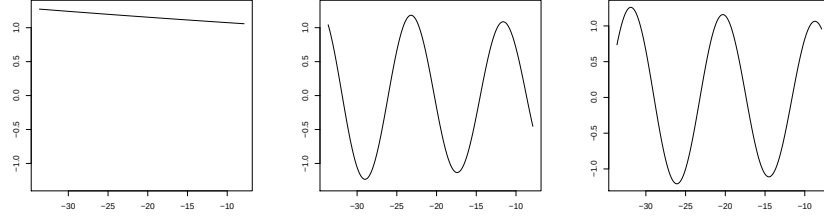


Fig. 4 Basis functions $\{\hat{\psi}_1, \hat{\psi}_2, \hat{\psi}_3\}$ of $\ker(\hat{L}^{Tot})$ with $m = 3$.

polynomials of $\hat{L}^{Low}$ and $\hat{L}^{Up}$. Both polynomials present one real root and two complex conjugate roots. The similarity of the three roots supports our final choice of using $\ker(\hat{L}^{Tot})$ as the three-dimensional space onto which project all 65 ICA radius functions.

Let us now characterize $\ker(\hat{L}^{Tot})$. Using the procedure presented in Section 2 the third order operator is estimated as follows:

$$\hat{L}^{Tot}x = D^3x + 0.0217D^2x + 0.2940D^1x + 0.0021x, \quad (5)$$

with $\hat{\lambda}_1 = -0.0072$, $\hat{\lambda}_2 = -0.0072 + 0.5400i$, and $\hat{\lambda}_3 = -0.0072 - 0.5400i$ being the three roots of the associated characteristic polynomial and $\{\hat{\psi}_1, \hat{\psi}_2, \hat{\psi}_3\}$ being a consequent practical real-valued basis:

$$\hat{\psi}_1(s) = e^{-0.0072s}, \quad \hat{\psi}_2(s) = e^{-0.0072s} \cos(0.5400s), \quad \hat{\psi}_3(s) = e^{-0.0072s} \sin(0.5400s).$$

Also a constrained PDA with constraints $\beta_2 = \beta_0 = 0$ has been performed achieving conclusions similar to one presented in Section 6. Because of the limited space, this alternative analysis is not reported here.

Due to the very small values of the real parts of the three roots, the first basis function $\hat{\psi}_1$ defines a slightly decreasing function and the remaining basis functions $\hat{\psi}_2$ and $\hat{\psi}_3$ jointly define a nearly sinusoidal function of period 11.6 mm with arbitrary phase.

In the left plot of Figure 5 the 65 ICA radius functions R_i are reported (red color for the Lower group and blue color for the Upper one). In the central plot of Figure 5 instead the 65 structural components $\hat{R}_i$ of the previous functions, computed as shown in Section 2, are reported. For comparison in the right plot of the same figure, the result achieved using a second order operator is reported.

Most of the 65 ICA radius functions R_i (left plot of Figure 5) present a very similar behavior with the exception of five curves presenting very “unusual” oscillations. It is surprising how well the third order operator $\hat{L}^{Tot}$ is able to jointly describe both the 60 “usual” and the five “unusual” ICA radius functions. In [8], the dimensional reduction achieved by FPCA completely smooths those unusual oscillations letting them be explained by high order principal components. This discrepancy points out

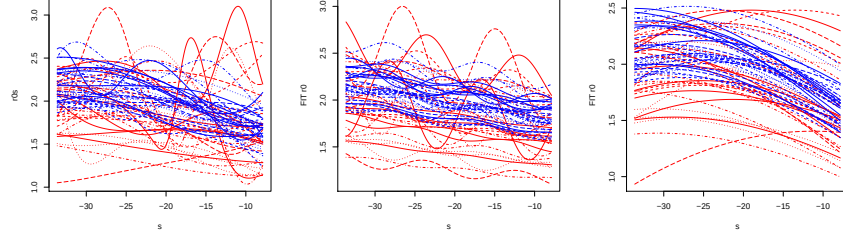


Fig. 5 The original radius functions R_i (left), their projections onto $\ker(\hat{L}^{Tot})$ for $m = 3$ (center) and $m = 2$ (right). Blue lines indicate Upper group patients while red ones Lower group patients.

the fact that, even if the values of these five ICA radius functions are unusual, the linear relations among the derivatives are not unusual but similar to the ones presented by the other 60 functions.

The last part of the analysis is within a more traditional functional data analysis framework, indeed we now want to analyze the structural variability pointed out by PDA. In particular, we want to perform a FPCA restricted to $\ker(\hat{L}^{Tot})$, in order to identify within this three-dimensional space a new origin and a new set of basis functions that are statistically optimal (i.e. with zero-mean and uncorrelated sample components).

The structural variability is strongly dominated by its first functional principal component. Indeed, the first, the second, and the third principal components of the structural variability account for 92%, 5%, and 3% of the structural total variance, respectively.

The first component is strongly associated with ψ_1 and accounts for the average amplitude of the radius along the ICA, thus being a scale factor among patients; the latter two are essentially rotations of ψ_2 and ψ_3 and thus jointly describe sinusoidal oscillations of period 11.6 mm. In Figure 6, the boxplots of the scores for the Upper and the Lower groups for the first, second, and third principal components within $\ker(\hat{L}^{Tot})$ are reported. It is evident from these that the first component discriminates between the two groups in terms of both location and dispersion. Upper group patients seem to have wider carotids than Lower group patients and a lower within-group variability. Moreover some outliers appear along the second and third principal components.

In order to statistically quantify such differences between the two groups we perform three Wilcoxon tests for comparing the median values of the scores of the two groups. The corresponding p -values are 0.7%, 11.0%, and 2.0% for the first, the second, and the third principal component, respectively, pointing out a significant difference in the median values of the first principal component scores for the two groups.

Moreover, not surprisingly, for the Lower group the scores associated to the second and third principal components appear to be very far from the normal distribution (p -values of the Shapiro-Wilk test both much less than 0.1%), supporting the

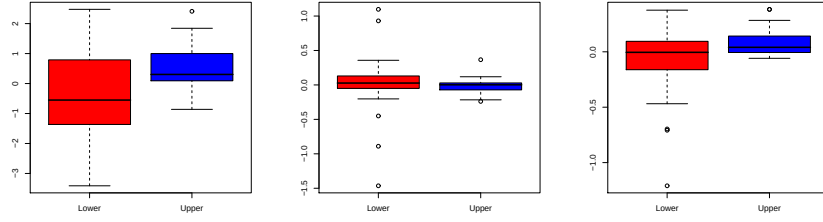


Fig. 6 Boxplots of the scores the first, second, and third principal components within $\ker(\hat{L}^{Tot})$ for $m = 3$ (left, center, right respectively) for the Upper and the Lower groups.

idea that these two principal components are accounting for the structural variability of the “unusual” radius functions. On the other hand, there is no evidence of the absence of normality for the scores associated to the first principal component, neither for the Lower group nor for the Upper group (p -values of the Shapiro-Wilk test 45.3% and 41.9% respectively).

Under the assumption of normality of the scores associated to the first principal component, we perform an F-test for testing the equality of the variances of the groups. The associated p -value is much less than 0.1%, strongly supporting the hypothesis of two different variances for the two groups.

Putting together the results of the previous tests with the interpretations of the functional principal components within $\ker(\hat{L}^{Tot})$ we can draw some interesting conclusions.

6 Conclusions

Nearly one third of the functional variability of the 65 ICA radius functions of the AneuRisk dataset can be described by the third order linear differential equation $\hat{L}^{Tot}x = 0$ with $\hat{L}^{Tot}$ as reported in (5). This variability is essentially due to a term defining the average radius of each ICA and a second one defining the amplitude and the phase of sinusoidal oscillations of the radius with period 11.6 mm around the average value.

Most of the structural variability (93%) is associated to the former component (the average radius of each ICA). Along this component the Lower group and the Upper group present different behaviors in terms of both location and dispersion: in particular the Upper group patients (the ones that are affected by aneurysm at the ICA terminal bifurcation or after it) present wider ICA than the Lower group patients (the ones that are healthy or affected by aneurysm before the terminal bifurcation of the ICA). Moreover, the variance of the average radius of the Upper group results significantly lower than the one of the Lower group, making the former group a very well defined group in terms of average radius.

Some patients, all belonging to the Lower group, present some anomalously large oscillations of the ICA radius, that should be further investigated from a medical point of view. On the other hand, except for these few cases, oscillations of the radius represent a minor mode of variability and no significant differences appear evident between the two groups from this perspective.

The conclusions about the average radius of the ICA are in complete agreement with the results drawn in [8]. Conclusions about the oscillations of the ICA radius are instead specific of the present analysis, indeed nothing similar is pointed out by the FPCA performed in [8]. This is not surprising since these oscillations, even if large, do not occur for similar values of the abscissa, and thus FPCA - that, differently from PDA, focuses just on punctual values of the functions - is not able to recognize them as a unique variability feature. On the other hand, contrary to [8], the present analysis does not point out anything regarding the tapering of the terminal part of the ICA.

These conclusions together with the analysis of the synthetic datasets show how in some cases, PDA can be a useful tool, alternative to FPCA, for performing a dimensional reduction of a functional dataset. In particular, when effective as a dimensional reduction tool, PDA *(i)* can provide a representation of functional data more easily interpretable than FPCA, *(ii)* this representation is not affected, differently by FPCA, by the presence of clusters or strong correlations among the original components, and *(iii)* it can detect important features of the data that FPCA is not able to detect.

References

1. Antiga L., Piccinelli M., Botti L., Ene-Iordache B., Remuzzi A., Steinman D.A. (2008): An image-based modeling framework for patient-specific computational hemodynamics, *Medical and Biological Engineering and Computing*, **46**, 1097-1120.
2. Poyton A.A., Varziri M.S., McAuley K.B., McLellan P.-J., Ramsay J.O. (2006): Parameter estimation in continuous-time dynamic models using principal differential analysis. *Computer and chemical engineering* **30**, 698-708.
3. Ramsay J.O., Munhall K.G., Gracco V.L., Ostry D.J. (1996): Functional data analysis of lip motion. *Journal of the Acoustical Society of America* **99**, 3718-3727.
4. Ramsay J.O. (2000): Functional components of variation in handwriting. *Journal of the American Statistical Association* **95**, 9-16.
5. Ramsay J.O., Silverman B.W. (2005): *Functional data analysis II Ed.*. Springer-Verlag, NY.
6. Rudin W. (1991): *Functional analysis, II Ed.*. McGraw-Hill, NY.
7. Sangalli L.M., Secchi P., Vantini S., Veneziani A. (2009): Efficient estimation of three-dimensional curves and their derivatives by free-knot regression splines, applied to the analysis of inner carotid artery centerlines. *Journal of the Royal Statistical Society*, vol **58**, Part 3, 285-306.
8. Sangalli L.M., Vantini S., Secchi P., Veneziani A. (2009): A Case Study in Exploratory Functional Data Analysis: Geometrical Features of the Internal Carotid Artery. *Journal of the American Statistical Association*, Vol. **104**, n. 485, 37-48.
9. Wang S., Jank W., Shmueli G., Smith P. (2008): Modeling price dynamics in eBay auctions using differential equations. *Journal of the American Statistical Association*, Vol. **103**, n. 483, 1100-1118.