

Functional Data Analysis - Shape Analysis

James Thomson

Simon Fraser University

February 13, 2019

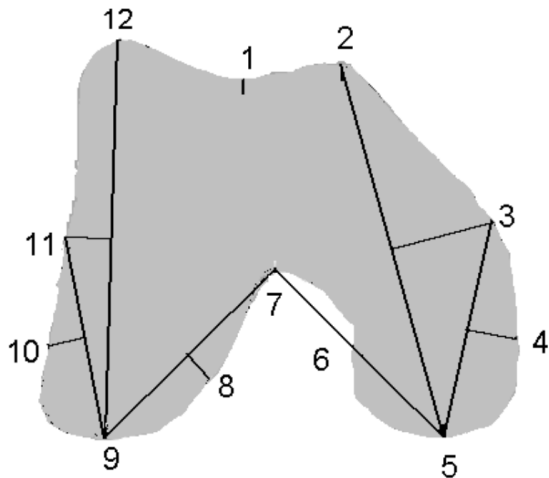
1. Introduction
2. Methods
 - 2.1 Landmark Description
 - 2.2 Shape Analysis with FDA
 - 2.3 Results
3. ICA in Functional PCA Classification
 - 3.1 Methodology
 - 3.2 Application
4. Final Thoughts

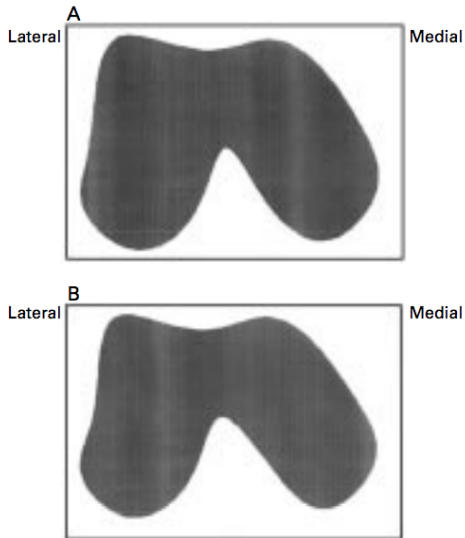
Introduction

Functional Data Analysis in Shape Analysis
Irene Epifanio, Noelia Ventura-Campos
Computational Statistics & Data Analysis, 2011

The data comes from a database of bone shapes from a study analyzed in Ramsey and Silverman (2002, Chapter 4). Each of the 68 bones have 12 $(x_i(t), y_i(t))$ points mapped along their outline. The goal is to classify between eburnated and non-eburnated femora (arthritic and non-arthritic). The data has been rotated ahead of time so that the shapes are all oriented approximately the same way, with left femora reflected to produce 'right' images.







Purpose To classify shapes based on a vector of observation

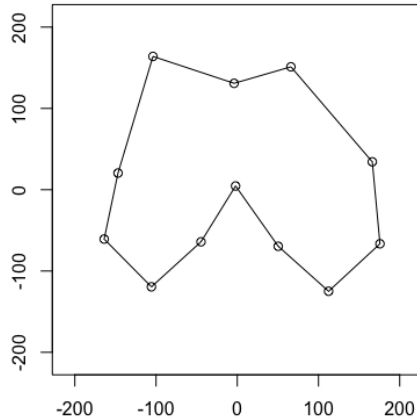
Methods

- ▶ Landmark description
- ▶ Functional Analysis

ICA An alternative method of functional classification

Example Code using the bone data

Methods



- $$\sqrt{\sum_{j=1}^m \|\mathbf{x}_j - \bar{\mathbf{x}}\|^2}$$

- ◀ ◻ ▶ ◀ ◻ ▶ ◀ ≡ ▶ ◀ ≡ ▶ ≡ ↺ 🔍 ↻

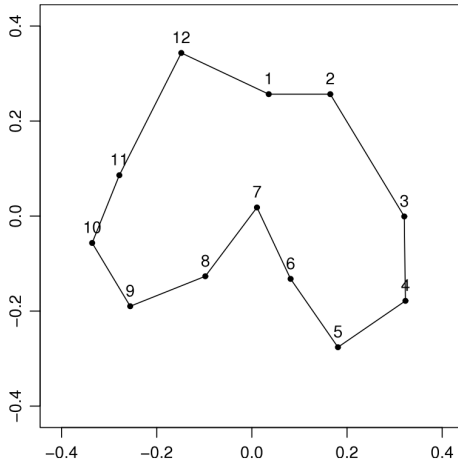


Figure: Mean shape of the 68 bones after centring and scaling

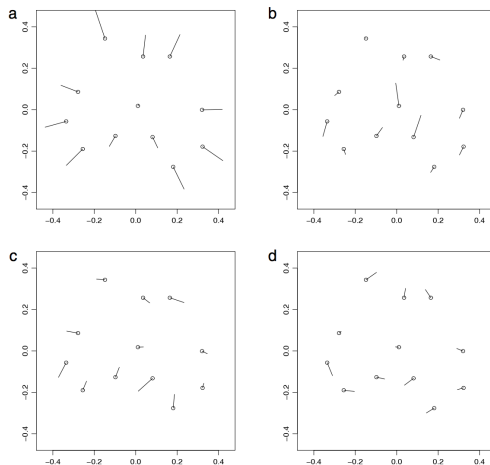


Figure: a) The 1st PC, b) the 2nd PC, c) the 3rd PC, d) the 4th PC

- ▶ Used PCA to reduce the dimensionality of the dataset, and to identify which PCs were significantly different between the classes.
- ▶ There is no guarantee that the separation between groups will be in the direction of the high-variance PC's, since they are uncorrelated we can assess them independently.
- ▶ Therefore we compute all PCs, and choose only those whose difference is significant at level $\alpha = 0.05$ from a t-test
- ▶ The difference is only significant on component 3 out of the 24 components.
- ▶ Using Leave-One-Out-Cross-Validation (LOOCV) with Linear Discriminant Analysis (LDA) from the scores of this component, get 2 false positives, and 14 false negatives
- ▶ If we take the next most significant component and use the same procedure we end up with 13 errors, but the errors are flipped in the other direction

```
> # Without the training data
> one_component = cbind(procg$scores[,3], ebl)
> total_one = as.data.frame(one_component)
>
> one_c = lda(ebl~., total_one, CV = FALSE)
> predict_one = predict(one_c, total_one)$class
> confusionMatrix(ebu, predict_one)$table
```

	Reference	
Prediction	0	1
0	50	2
1	14	2


```
> two_component = cbind(procg$scores[,c(3,7)], ebl)
> total_two = as.data.frame(two_component)
>
> two_c = lda(ebl~., total_two, CV = FALSE)
> predict_two = predict(two_c, total_two)$class
> confusionMatrix(predict_two, ebu)$table
```

	Reference	
Prediction	0	1
0	50	11
1	2	5

- ▶ We define the variance-covariance structure of the data x_i as

$$v(s, t) = N^{-1} \sum_i^K x_i(s)x_i(t)$$

- ▶ In PCA we need to solve the eigenequation such that

$$\mathbf{V}\xi = \rho\xi$$

- ▶ In functional PCA the eigenvectors ξ are replaced with eigenfunctions $\xi(t)$ and the eigenequation with

$$\int v(s, t)\xi(t)dt = \rho\xi(s)$$

- ▶ In the bivariate setting, x_1, \dots, x_K and y_1, \dots, y_K have variance covariance structures v_{XX} and v_{YY} , and cross-covariance v_{XY}
- ▶ A single principle component is defined by $\xi = (\xi^X, \xi^Y)$
- ▶ The principle component score s_i for the i -th bivariate function $(x_i(t), y_i(t))$ is the inner product between bivariate functions for the i -th x and y .
- ▶ A straightforward definition of an inner product between two weight functions

$$\langle \xi_1, \xi_2 \rangle = \int \xi_1^X \xi_2^X + \int \xi_1^Y \xi_2^Y$$

- ▶ So our PC score can be written

$$s_i = \int x_i \xi_x + \int y_i \xi_y$$

- ▶ The solution to the eigenequation system $\mathbf{V}\xi = \rho\xi$ has to satisfy both

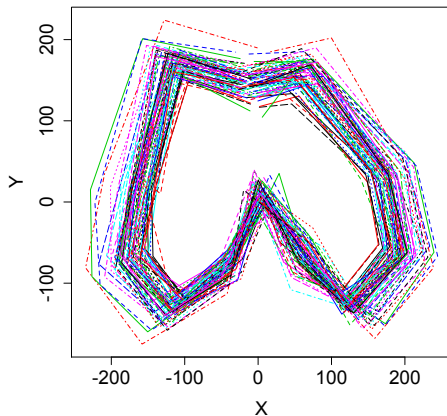
$$\int x_{XX}(s, t)\xi_X(t)dt + \int v_{XY}(s, t)\xi_Y(t)dt = \rho\xi_X(s)$$

and

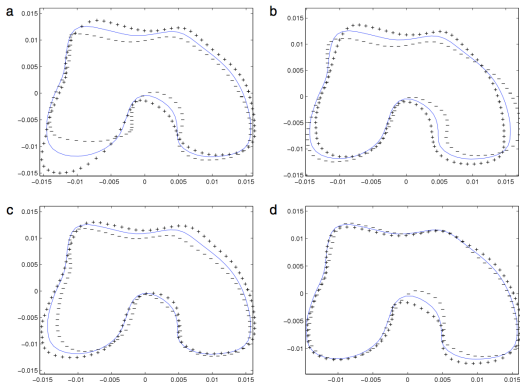
$$\int x_{YY}(s, t)\xi_X(t)dt + \int v_{XY}(s, t)\xi_Y(t)dt = \rho\xi_Y(s)$$

- ▶ To solve, $x_i(t)$ and $y_i(t)$ are replaced by vectors of points and we concatenate our X_i and Y_i into a single vector Z_i
- ▶ This is all accomplished by the *fda* library functions

- ▶ Data is mean centred, and the curves are evaluated on 100 equidistant points on $[0,1]$
- ▶ A 51 basis Fourier series is used to approximate its functional form
- ▶ Each curve is completely determined by the coefficients of its basis



- ▶ The first four PCs make up 55%, 10.59%, 9.02%, and 5.91% of the variance.



- ▶ For the discriminant analysis, we do the same as before, choosing only those components with a difference which is significant; components 1, 6, and 10
- ▶ Using LOOCV again, we select the classification with component 6 or 10, both with 16 errors. If two components are used, 1 and 6 give 11 errors.

- ▶ We find that each of these approaches are similarly effective in classification
- ▶ For the landmark data 84.62% of the variance is concentrated on a single component, and so the process is unable to identify fine details in the bone shape
- ▶ Functional PCA is not overly decomposed, but more decomposed than the landmark approach
- ▶ FDA also can be classified using other functional classification methods

ICA in Functional PCA Classification

- ▶ An alternative method for curve discrimination is the coefficients of the Independent Components Analysis (ICA) components
- ▶ Coefficients can be computed easily, and can be extended easily to the multivariate functional case

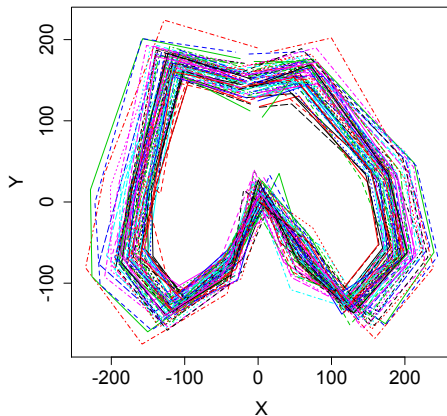
$$x_i(t) = \sum_{j=1}^n a_{ij} s_j(t)$$

- ▶ In practice we have discretized curves ($\mathbf{x}_i = \{x_{ik} : k = 1, \dots, m\}$), and can consider the $m \times n$ data matrix to be a linear combination of independent components, where \mathbf{S} contains the components and \mathbf{A} is a linear mixing matrix.

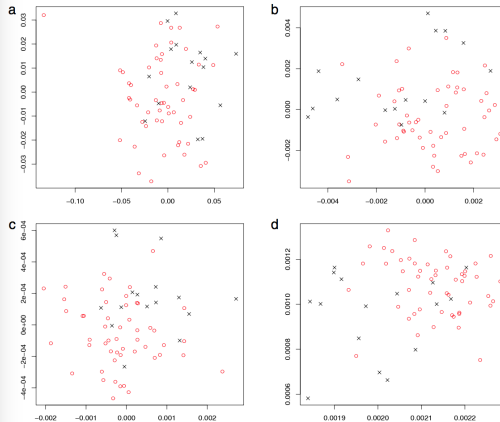
$$\mathbf{X} = \mathbf{S}\mathbf{A}$$

There are 3 assumptions of ICA

- ▶ Independence between signals
- ▶ Sums of the measured signals in the in \mathbf{X} tend to be more Gaussian than the source signals
- ▶ The mixture of signals is more complex than the simplest source signal



- ▶ We are looking for the un-mixing matrix \mathbf{W} such that $\mathbf{XW} = \mathbf{S}$.
- ▶ Since we assume that the mixture of signals is more complex and more Gaussian than the source signal, our independent components will identify the maximum non-Gaussian signals under the constraint that the components are orthonormal



- ▶ Applying the same methods as with the functional PCA, computing the PCA first, retain a certain number of components
- ▶ Then compute the ICA with all the data (using only non-negligible eigenvalues)
- ▶ The significantly different components are 1,2,7,10,12,14,16,18, and using LDA and LOOCV like before we select for single component and double component classification.
- ▶ For single component classification the first component was chosen and 12 errors were reported
- ▶ For double component classification the first and seventh were chosen, and 8 errors were reported

Final Thoughts

- ▶ The authors also considered other options for classification, but found that none performed as well as the functional PCA with ICA component selection
- ▶ LDA with a ridge penalty misclassified 10, but like in the other cases using the full dataset for the testing data so that is an over-optimistic estimate as well
- ▶ The double-cross validation method, where a single observation is removed and then another LOOCV takes place with the external LOOCV algorithm produced 15 errors, but suffers from the same problems the other methods in this paper.
- ▶ Overall, the data is sparse and difficult to fit a smooth function to without losing the characteristic shape of a femora bone.
- ▶ The landmark algorithm does not have a larger number of data points, particularly in the region of highest variability

- ▶ The authors suggest that components should be chosen when using functional shape analysis
- ▶ Often the component(s) of maximum variance are not the components with a significant difference between the eburnated and non-eburnated bone shapes