Elastic Full Procrustes Means for Sparse and Irregular Planar Curves

Manuel Pfeuffer

Masters Thesis Presentation
Chair of Statistics
Humboldt–Universität zu Berlin

Advisors: Lisa Steyer, Almond Stöcker, Prof. Dr. Sonja Greven

2nd Examiner: Prof. Dr. Nadja Klein

Motivation

Calculate shape means for 2D curves:

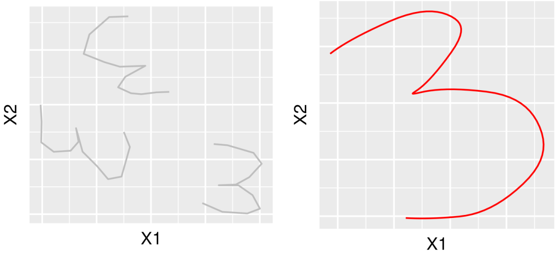

Figure: `digits3.dat` from the `shapes` package (Dryden 2019) with estimated elastic full Procrustes mean. Data: Anderson (1997)

Challenges:

Sparse and irregular, warping, translation/scaling/rotation

Outline

1. What is an Elastic Full Procrustes Mean?
2. Estimation Strategy
3. Results (so far), Problems, Outlook

Sparse and Irregular Planar Curves

How can we compare observations $\beta_i = (\beta_{i1}, \beta_{i2}, \ldots, \beta_{im_i})$?

$\rightarrow$ Treat β_i as functional data $\beta_i(t)$: $\beta_i : [0,1] \rightarrow \mathbb{R}^2$

$\beta_i(t)$ sampled at $t_i = (t_{i1}, \ldots, t_{im_i})$:

$$\beta_{i1} = \beta_i(t_{i1}), \ldots, \beta_{im} = \beta_i(t_{im_i})$$

How to construct $(t_{i1}, \ldots, t_{im_i})$?

- simple: arc–length
- better: same values of t relate to same "part" of curve

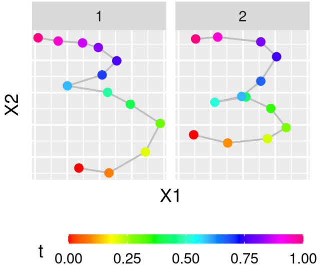

Figure: `digits3.dat` with arc–length parametrisation

Elastic Mean and Warping

Problem: Need to find an optimal re-parametrization (warping).

Well known problem in functional data analysis:

- Perform warping alignment on SRV curves

Square-Root-Velocity (SRV) Framework (Srivasta et al. 2011)

$$q : [0,1] \to \mathbb{R}^2, \quad q(t) = \frac{\dot{\beta}(t)}{\sqrt{||\dot{\beta}(t)||}} \quad \text{for } ||\dot{\beta}(t)|| \neq 0$$

- Use warping methods for sparse and irregular curves as implemented in `elasdics` (Steyer 2021)
- Mean under optimal re-parametrization is called **elastic**

Elastic Mean and Warping

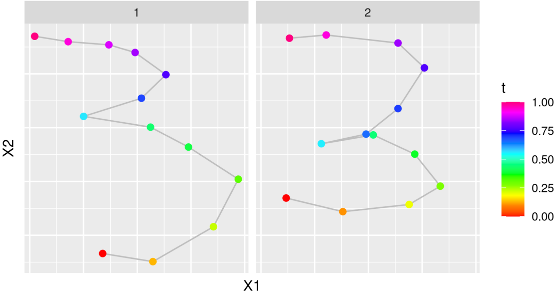

Figure: with `align_curves()` from package `elasdics` (Steyer 2021)

Problem: Methods are not invariant under rotation/scaling!

Full Procrustes Mean and Procrustes Fits

Basic idea:

1. Calculate mean that is invariant under rotation, scaling
2. Align rotation and scaling
3. Align parametrisation

$\rightarrow$ use estimated mean as reference for alignment

1+2 are known problems in **statistical shape analysis** (see e.g. Dryden and Mardia (2016)) $\rightarrow$ Procrustes mean and Procrustes fits

Note: More efficient to perform steps directly on the SRV curves!

Full Procrustes Mean and Procrustes Fits

Now: Need to derive the Procrustes mean for functional data

We can show using complex notation

$$q_i : [0,1] \to \mathbb{C}, \quad q_i(t) = x_i(t) + i\, y_i(t)$$

that the population level full Procrustes mean for normalized curves $\tilde{q}$ is given by:

$$\mu_q = \underset{z:[0,1]\to\mathbb{C},\, ||z||=1}{\operatorname{argmax}} \int_0^1 \int_0^1 \overline{z}(s)\, \mathbb{E}\left[\tilde{q}(s)\overline{\tilde{q}}(t)\right] z(t) ds dt$$

$\rightarrow$ $\mathbb{E}\left[\tilde{q}(s)\overline{\tilde{q}}(t)\right]$ is the complex covariance function $C(s,t)$

Full Procrustes Mean and Procrustes Fits

Population level full Procrustes mean:

$$\mu_q = \underset{z:[0,1]\to\mathbb{C},\,||z||=1}{\operatorname{argmax}} \int_0^1 \int_0^1 \overline{z}(s) C(s,t) z(t) ds dt$$

$\rightarrow$ Functional PCA problem (see Ramsay and Silverman (2005))

$\rightarrow$ Solution is the leading complex eigenfunction of $C(s,t)$

$\rightarrow$ We only need to find a good estimate $\hat{C}(s,t)$ to get $\hat{\mu}_q$!

Procrustes fits:

$$\hat{q}_i^P = \langle \tilde{q}_i, \hat{\mu}_q \rangle \tilde{q}_i$$

with $\langle f, g \rangle = \int_0^1 \overline{f(t)} g(t) dt$

Hermitian Covariance Smoothing

Treat estimation of $C(s,t) = \mathbb{E}[\tilde{q}(s)\overline{\tilde{q}}(t)]$ as a regression problem:

- we can build response $y_{ijk} = \tilde{q}_i(t_{ij})\overline{\tilde{q}}_i(t_{ik})$
- treat parametrisation t_{ij}, t_{ik} as "covariates" s and t
- non-parametric regression: $\mathbb{E}[y] = f(s,t)$
- use $\hat{C}(s,t) = \hat{f}(s,t)$ for functional PCA

Note: Using symmetry properties of $C(s,t)$ is important for efficient estimation (see Cederbaum, Scheipl, and Greven (2018)).

$\rightarrow$ use every combination (t_{ij}, t_{ik}) only once

Hermitian Covariance Smoothing

Here: Complex covariance function is hermitian $C(s,t) = \overline{C}(t,s)$

$$\mathbb{E}[Re(y)] = f_{symm}(s,t)$$

$$\mathbb{E}[Im(y)] = f_{skew}(s,t)$$

- model real and imaginary parts seperately using e.g. `mgcv` (Wood 2017)
- with symmetric and skew-symmetric tensor product P-splines from `sparseFLMM` (Cederbaum, Volkmann, and Stöcker 2021)

We get:

$$\hat{C}(s,t) = b(s)^T \hat{\Xi} b(t), \quad \text{with} \quad \hat{\Xi} = \hat{\Xi}_{symm} + i\,\hat{\Xi}_{skew}$$

Hermitian Covariance Smoothing

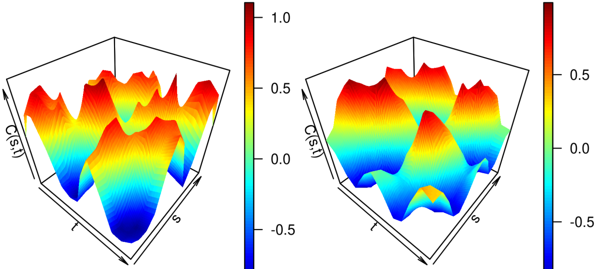

Figure: Estimated real (left) and imaginary (right) parts of $C(s,t)$

Empirical Procrustes Mean in a Fixed Basis

Idea: Estimate covariance function and mean in the same basis.

Then the optimization problem reduces to:

$$\hat{\theta}_\mu = \underset{\theta:\theta^H G\theta=1}{\operatorname{argmax}}\ \theta^H G \hat{\underline{\Xi}} G\theta \quad \text{with} \quad G_{kl} = \langle b_k, b_l \rangle$$

Eigenvalue problem treating $\theta = \theta_{Re} + i\,\theta_{Im}$ seperately:

$$\begin{pmatrix} \hat{\underline{\Xi}}_{symm} & -\hat{\underline{\Xi}}_{skew} \\ \hat{\underline{\Xi}}_{skew} & \hat{\underline{\Xi}}_{symm} \end{pmatrix} \begin{pmatrix} G & 0 \\ 0 & G \end{pmatrix} \begin{pmatrix} \theta_{Re} \\ \theta_{Im} \end{pmatrix} = \lambda \begin{pmatrix} \theta_{Re} \\ \theta_{Im} \end{pmatrix}$$

$\Rightarrow$ Solution is the leading normalized eigenvector of $\underline{\hat{\underline{\Xi}} G}$

Procrustes Fits

After obtaining $\hat{\mu}_q$ we can estimate the Procrustes fits:

$$\hat{q}_i^P = \langle \tilde{q}_i, \hat{\mu}_q \rangle \tilde{q}_i$$

- at the moment: integration with linear interpolation of $\tilde{q}_i$'s
- alternative: smoothing in mean basis with $\langle \hat{\tilde{q}}_i, \hat{\mu}_q \rangle = \hat{\theta}_i^H G \hat{\theta}_\mu$

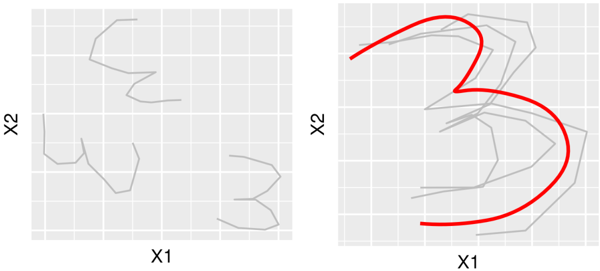

Figure: Procrustes fits and mean (on full dataset) before warping.

Results

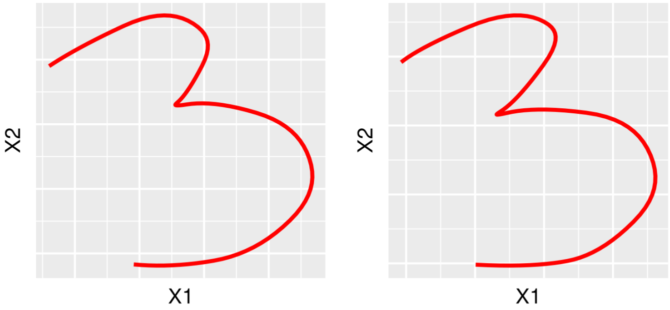

Figure: Full Procrustes mean (left) and elastic full Procrustes mean (right) on `digits3.dat`.

Results / Problems

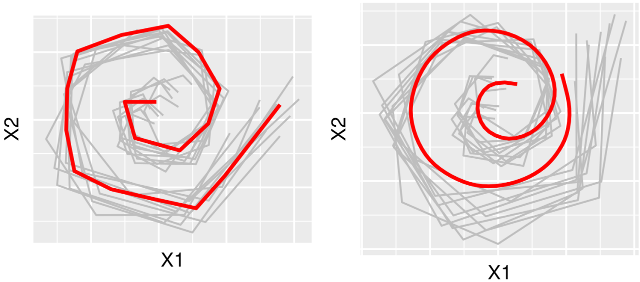

Figure: Polygon (left) and smooth (right) elastic full Procrustes mean with procrustes fits.

→ Parameters: **knots**, **spline degree**, **penalty**

→ Consistent results for piecewise constant splines (at SRV level) and zero order penalty (in the cov. estimation).

Outlook

Next steps:

- Real world data application (open curves)
- Better normalization / estimation of procrustes fits

Nice to have (maybe later):

- Mean for closed curves
- **Code**: bugs, testcases, faster

References

- Anderson, C. R. (1997). *Object recognition using statistical shape analysis*. PhD thesis. University of Leeds.
- Cederbaum, J., F. Scheipl, and S. Greven (2018). "Fast symmetric additive covariance smoothing". In: *Computational Statistics & Data Analysis* 120, pp. 25–41.
- Cederbaum, J., A. Volkmann, and A. Stöcker (2021). *sparseFLMM: Functional Linear Mixed Models for Irregularly or Sparsely Sampled Data*. R package version 0.4.0. URL: https://CRAN.R-project.org/package=sparseFLMM.
- Dryden, I. L. (2019). *shapes package*. Contributed package, Version 1.2.5. R Foundation for Statistical Computing. Vienna, Austria. URL: http://www.R-project.org.

- Dryden, I. L. and K. V. Mardia (2016). *Statistical shape analysis. With applications in R*. Vol. 995. John Wiley & Sons.
- Ramsay, J. and B. W. Silverman (2005). *Functional Data Analysis*. Springer series in statistics.
- Srivasta, A. et al. (2011). "Shape analysis of elastic curves in euclidean spaces". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 33.7, pp. 1415–1428.
- Steyer, L. (2021). *elasdics: Elastic Analysis of Sparse, Dense and Irregular Curves*. R package version 0.1.1. URL: https://CRAN.R-project.org/package=elasdics.
- Steyer, L., A. Stöcker, and S. Greven (2021). "Elastic analysis of irregularly or sparsely sampled curves". In: Preprint.
- Wood, S.N (2017). *Generalized Additive Models: An Introduction with R*. 2nd ed. Chapman and Hall/CRC.

Putting it all together

Algorithm Elastic Full Procrustes Mean

Input: Data curves $\beta_1, \dots, \beta_N$

Output: Procrustes mean $\hat{\mu}$ and Procrustes fits $\hat{\beta}_1^P, \dots, \hat{\beta}_N^P$

1: initialize arc–length parametrisation t_i for all β_i
2: calculate normalized SRV curves $\tilde{q}_i = \frac{q_i}{||q_i||}$
3: **while** convergence not reached **do**
4: estimate $C(s,t)$ using t_i
5: calculate $\hat{\mu}_q$ as the leading eigenfunction of $\hat{C}(s,t)$
6: estimate Procrustes fits $\tilde{q}_i^P$
7: update $t_i \leftarrow t_i^{optim}$ using warping alignment on $\hat{\tilde{q}}_i^P$
8: **end while**
9: **return** integrated Procrustes mean and fits

Problems

Normalization: $\tilde{q}_i = \frac{q_i}{||q_i||}$ is itself an estimate ($||q_i|| = \sqrt{\langle q_i, q_i \rangle}$)

$\rightarrow$ can either restrict β_i's to unit–length or normalize q_i directly

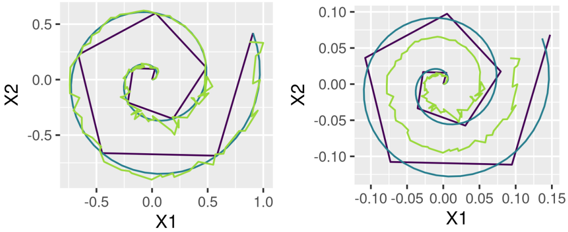

Figure: Spirals (left) and unit–length spirals (right).

$\rightarrow$ likely need smoothing (in the mean basis?) for this

$\rightarrow$ only ok, as long as all curves have same "amount" of sparsity

Derivation: Empirical Full Procrustes Mean

Full Procrustes Distance and Mean

$$d_F(q_1, q_2) = \inf_{\Gamma, b} ||q_1 - b\Gamma q_2||$$

$$\mu_q : [0,1] \to \mathbb{R}^2, \quad \hat{\mu}_q = \underset{z:[0,1]\to\mathbb{R}^2}{\operatorname{argmin}} \sum_{i=1}^{N} d_F(z, q_i)^2$$

One can show that (Dryden and Mardia 2016, see Ch. 8):

$$\hat{\mu}_q = \underset{z:[0,1]\to\mathbb{C}}{\operatorname{argmin}} \sum_{i=1}^{N} \underbrace{1 - \frac{\langle z, q_i\rangle\langle q_i, z\rangle}{\langle z, z\rangle\langle q_i, q_i\rangle}}_{=d_F(z,q_i)^2}$$

Derivation: Empirical Full Procrustes Mean

Note: $\langle q_1, q_2 \rangle = \int_0^1 \overline{q}_1(t) q_2(t) dt$, and $||q_i|| = \sqrt{\langle q_i, q_i \rangle}$

$$\hat{\mu}_q = \underset{z:||z||=1}{\operatorname{argmax}} \sum_{i=1}^{N} \langle z, \tilde{q}_i \rangle \langle \tilde{q}_i, z \rangle, \quad \text{with} \quad \tilde{q}_i = \frac{q_i}{||q_i||}$$

$$= \underset{z:||z||=1}{\operatorname{argmax}} \sum_{i=1}^{N} \int_0^1 \overline{z}(s) \tilde{q}_i(s) ds \int_0^1 \overline{\tilde{q}}_i(t) z(t) dt$$

$$= \underset{z:||z||=1}{\operatorname{argmax}} \int_0^1 \int_0^1 \overline{z}(s) \left(\sum_{i=1}^{N} \tilde{q}_i(s) \overline{\tilde{q}}_i(t) \right) z(t) ds dt$$

$\rightarrow$ sample analouge to $C(s,t) = \mathbb{E}[\tilde{q}(s)\overline{\tilde{q}}(t)]$

Derivation: Mean in a Fixed Basis

$$\mu_q = \underset{z:[0,1]\to\mathbb{C},\, ||z||=1}{\operatorname{argmax}} \int_0^1 \int_0^1 \overline{z}(s) C(s,t) z(t) ds dt$$

with $z(t) = b(t)^T \theta$ and $\hat{C}(s,t) = b(s)^T \hat{\Xi} b(t)$. Then

$$\begin{aligned}
\hat{\theta}_\mu &= \underset{\theta:||b^T\theta||=1}{\operatorname{argmax}} \int_0^1 \int_0^1 \overline{(b(s)^T\theta)} b(s)^T \hat{\Xi} b(t) b(t)^T \theta \, ds dt \\
&= \underset{\theta:||b^T\theta||=1}{\operatorname{argmax}} \theta^H \left(\int_0^1 b(s) b(s)^T ds \right) \hat{\Xi} \left(\int_0^1 b(t) b(t)^T dt \right) \theta \\
&= \underset{\theta:||b^T\theta||=1}{\operatorname{argmax}} \theta^H G \hat{\Xi} G \theta
\end{aligned}$$

Note: $||b^T\theta|| = \sqrt{\theta^H G \theta}$