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 observed at discrete points:

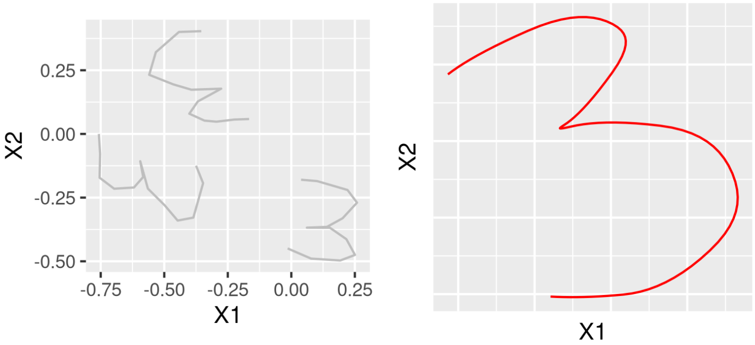

Figure: `digits3.dat` from the `shapes` package (Dryden 2019) with estimated elastic full Procrustes mean. Data: Anderson (1997)

Outline

1. What is an Elastic Full Procrustes Mean?
   - $\rightarrow$ Sparse and Irregular Planar Curves
   - $\rightarrow$ Elastic Mean and Warping
   - $\rightarrow$ Full Procrustes Mean and Procrustes Fits
2. Estimation Strategy
   - $\rightarrow$ Hermitian Covariance Smoothing
   - $\rightarrow$ Estimation of the Procrustes Mean in a Fixed Basis
   - $\rightarrow$ Procrustes Fits
3. Results (so far), Problems, Outlook

Sparse and Irregular Planar Curves

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

Treat as functional data

$$\beta_i : [0,1] \to \mathbb{R}^2$$

observed on $t_i = (t_{i1}, \dots, t_{im_i})$

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

Functional Mean

$$\hat{\mu}(t) = N^{-1} \sum_{i=1}^{N} \beta_i(t), \; t \in [0,1]$$

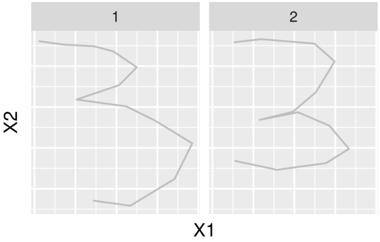

Figure: `digits3.dat` from the `shapes` package (Dryden 2019). Data: Anderson (1997)

Elastic Mean and Warping

- Parametrisation $t_i = (t_{i1}, \dots, t_{im})$ usually unknown.
- Initialize as $t_{ij} = \frac{\text{length of curve i up to point j}}{\text{length of curve i}} \in [0, 1]$

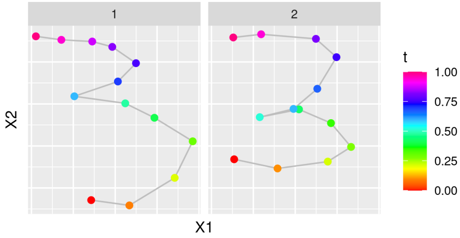

Figure: with `get_arc_length_param()` from `elasdics` (Steyer 2021)

Functional Mean

$$\hat{\mu}(t) = N^{-1} \sum_{i=1}^{N} \beta_i(t)$$

Problem: $\beta_1(0.5)$, $\beta_2(0.5)$ relate to different "parts" of the curve

Elastic Mean and Warping

Problem: $\beta_1(0.5)$ and $\beta_2(0.5)$ relate to different "parts" along curve
$\rightarrow$ Find a re-parametrization (warping) that aligns these parts.

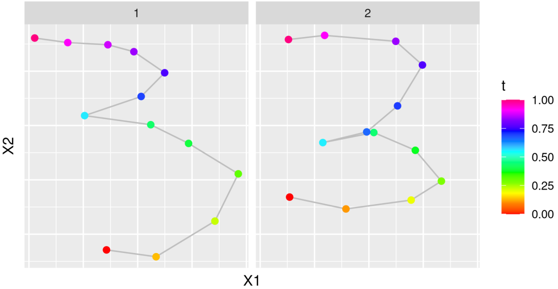

Figure: with `align_curves()` from package `elasdics` (Steyer 2021)

Note: Parametrisation relates to speed at which curve is traversed!

Elastic Mean and Warping

Note: Parametrisation relates to speed at which curve is traversed!

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

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

- Perform warping alignment on SRV curves
- We can recover original curves β_i up to translation
- Use warping methods for sparse and irregular curves as outlined in Steyer, Stöcker, and Greven (2021) and implemented in `elasdics` (Steyer 2021)

Full Procrustes Mean and Procrustes Fits

Problem: Methods are not invariant under scaling and rotation.

Idea:

1. Calculate SRV mean that is invariant under rotation, scaling
2. Align rotation and scaling of SRV curves to fit the mean
3. Perform warping on aligned SRV curves

Well known problem in **statistical shape analysis** (see e.g. Dryden and Mardia (2016)):

Full Procrustes Distance and Mean

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

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

Full Procrustes Mean and Procrustes Fits

Calculation simplifies for 2D shapes when using complex notation:

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

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

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

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

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

Full Procrustes Mean and Procrustes Fits

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

$$
\begin{aligned}
\hat{\mu}_q &= \underset{z:||z||=1}{\operatorname{argmax}} \sum_{i=1}^{N} \langle z, \tilde{q}_i \rangle \langle \tilde{q}_i, z \rangle \\
&= \underset{z:||z||=1}{\operatorname{argmax}} \sum_{i=1}^{N} \int_0^1 \overline{z}(s) \tilde{q}_i(s) ds \int_0^1 \overline{\tilde{q}}_i(t) z(t) dt \\
&= \underset{z:||z||=1}{\operatorname{argmax}} \int_0^1 \int_0^1 \overline{z}(s) \left(\sum_{i=1}^{N} \tilde{q}_i(s) \overline{\tilde{q}}_i(t) \right) z(t) ds dt
\end{aligned}
$$

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

→ tough to estimate in a sparse setting

Full Procrustes Mean and Procrustes Fits

Population level full Procrustes mean:

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

→ Functional PCA problem (see Ramsay and Silverman (2005))
→ Solution is the complex leading eigenfunction of $C(s,t)$
→ We only need to find a good estimate $\hat{C}(s,t)$ to get $\hat{\mu}_q$!

We can then estimate the rotation and scaling aligned procrustes fits as:

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

Note: For $\vec{x}, \vec{y} \in \mathbb{R}^d$: $\cos(\theta) = \frac{\langle \vec{x}, \vec{y} \rangle}{||\vec{x}|| \cdot ||\vec{y}||}$

Hermitian Covariance Smoothing

Problem: $\check{C}(s,t) = \frac{1}{N}\sum_{i=1}^{N} \tilde{q}_i(s)\overline{\tilde{q}_i}(t)$ is not a good estimator

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

- we have observations $y_{ijk} = \tilde{q}_i(t_{ij})\overline{\tilde{q}}_i(t_{ik})$
- 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: Difference to $\check{C}(s,t)$: Smoothing once with all the data vs. smoothing N times with little data.

Hermitian Covariance Smoothing

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

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

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

- model real and imaginary parts seperately
- use `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

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

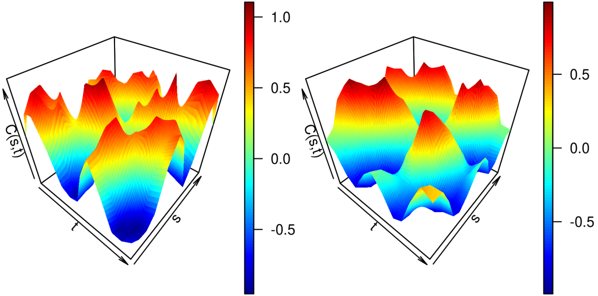

Figure: Estimated real (left) and imaginary (right) parts of $C(s,t)$ using P-splines of degree 1, 13 knots and a zero order penalty.

$\rightarrow$ How to choose knots, degree and penalty?

Empirical Procrustes Mean in a Fixed Basis

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

$$\hat{\mu}_q(t) = b(t)^T \hat{\theta}_\mu, \quad \hat{C}(s,t) = b(s)^T \hat{\Xi} b(t)$$

with real spline basis $b = (b_1, \dots, b_k)$, $\hat{\theta}_\mu \in \mathbb{C}^k$, $\hat{\Xi} \in \mathbb{C}^{k \times k}$

Then we can solve the optimization problem directly on $\hat{\Xi}$:

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

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

Procrustes Fits

Using $\hat{\mu}_q(t) = b(t)^T \hat{\theta}_q$ 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 $\langle \hat{\tilde{q}}_i, \hat{\mu}_q \rangle = \hat{\theta}_i^H G \hat{\theta}_\mu$

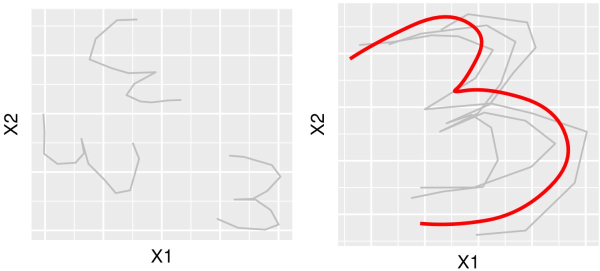

Figure: Procrustes fits and mean (on full dataset) before warping.

Results

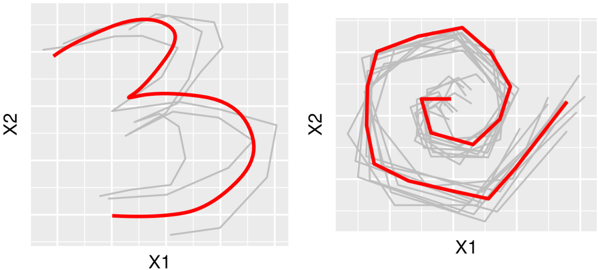

Figure: Elastic full Procrustes mean and procrustes fits for piecewise linear (left) and piecewise constant (right) splines on SRV level.

→ Consistent results for piecewise constant splines and zero order penalty.

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

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

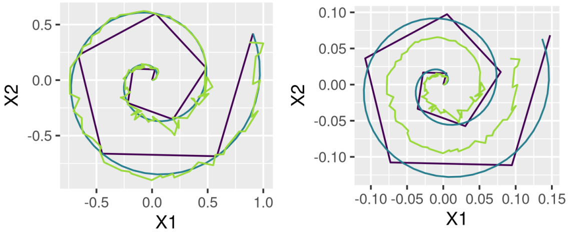

Figure: Spirals (left) and unit–length spirals (right).

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

Outlook

ToDo:

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

Nice to have (maybe later):

- Cover codebase with testcases (make it *really* solid)
- Disentangle codebase from `elasdics` and package it
- Reduce bloat / make it faster

Timeline: Would like to be finished by July

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.