Bayesian Hierarchical Curve Registration

Paolo Bighignoli, Federica Mattina, Federica Principe
Mathematical Engineering - Bayesian statistics project

Dataset

- Healthy people - Figure 1
- People who had physiotherapy - Figure 2
- People who had surgery - Figure 3

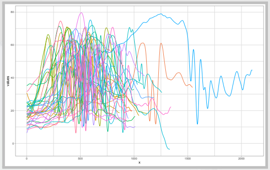


Figure 1

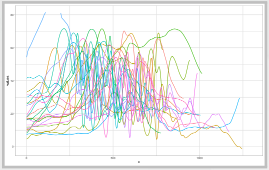


Figure 2

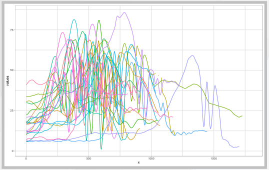


Figure 3

Goals

- Synchronize data through curve registration
  - Model both amplitude and timing of each individual curve
  - Gibbs sampler algorithm with one step of Metropolis-Hastings
- Comparison among groups

Model

$$\begin{cases} y_i(t) = c_i + a_i \boldsymbol{\mathcal{B}}'_m(\mu_i(t))\boldsymbol{\beta} + \epsilon_i & i = 1, .., N \\ \epsilon_i \overset{iid}{\sim} \mathcal{N}(0, \sigma^2_\epsilon) \end{cases}$$

Common shape function: $m(t; \boldsymbol{\beta}) = \boldsymbol{\mathcal{B}}'_m(t)\boldsymbol{\beta}$

Time transformation function: $\mu_i(t) = f_i + g_i \cdot t$

$$\Rightarrow \quad y^\star_i(t) = y_i(t) \circ \mu_i(t)$$

Model

$$\boldsymbol{\theta} = (\boldsymbol{a}, \boldsymbol{c}, \boldsymbol{f}, \boldsymbol{g}, \boldsymbol{\beta}, a_0, c_0, \sigma_a^2, \sigma_c^2, \sigma_\epsilon^2, \lambda)$$

- **Likelihood:**

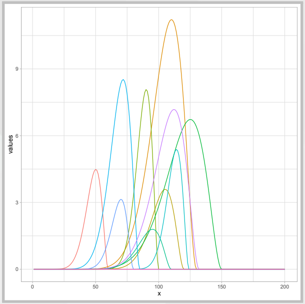
$$\boldsymbol{Y}_i \overset{ind}{\sim} \mathcal{N}(c_i \cdot 1_n + a_i \cdot \boldsymbol{\mathcal{B}}'_m(\mu_i(t))\boldsymbol{\beta}, \sigma_\epsilon^2 \boldsymbol{I}_n) \quad i = 1, .., N$$

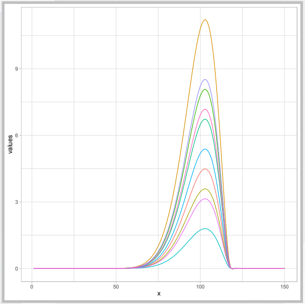
- **Prior:**
  - Gaussian for coefficients $a, c, f, g, a_0, c_0, \beta$.
  - Inverse-Gamma for variances $\sigma_a^2, \sigma_c^2, \sigma_\epsilon^2$, λ.

Testing on Fake Data

The computations on the original data are expensive due to:

- Thickness of sampling times
- Large number of parameters





Problems

- Setting of the proposal variance
- Identifiability of the chain

Adaptive Metropolis-Hastings

- **First approach** - Constant proposal variance
- **Last approach** - Adaptive Metropolis-Hastings

$$q_t(\cdot|X_1, \dots, X_{t-1}) \sim \mathcal{N}(X_{t-1}, \sigma_t^2)$$

where:

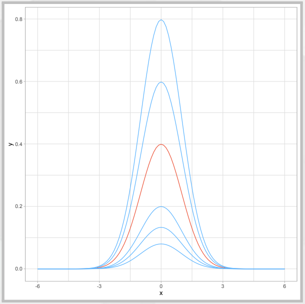
$$\sigma_t^2 = \begin{cases} \sigma_0^2 & t \leq t_0 \\ \frac{1}{t-1}\left(\sum_{i=0}^{t-1} x_i^2 - t \cdot \bar{x}^2\right) & t > t_0 \end{cases}$$

Identifiability of the chain

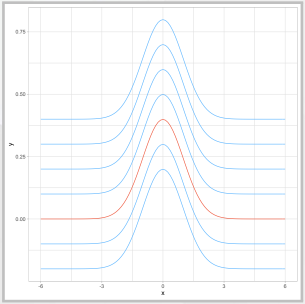
The model was over-parameterized:

- The product of the parameters a and β could be synthesized into a single parameter.
- The common shape function is defined just up to a constant.

Identifiability of the chain

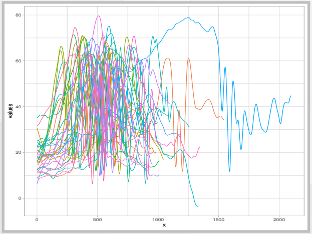


a_1=1

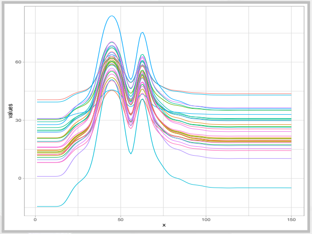


β_0=0

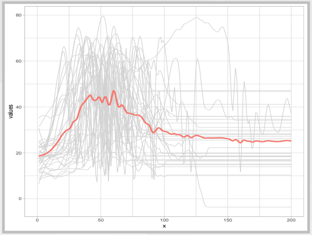
Healthy People



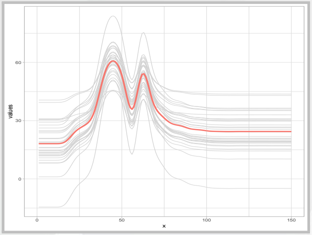
Original curves



Aligned curves

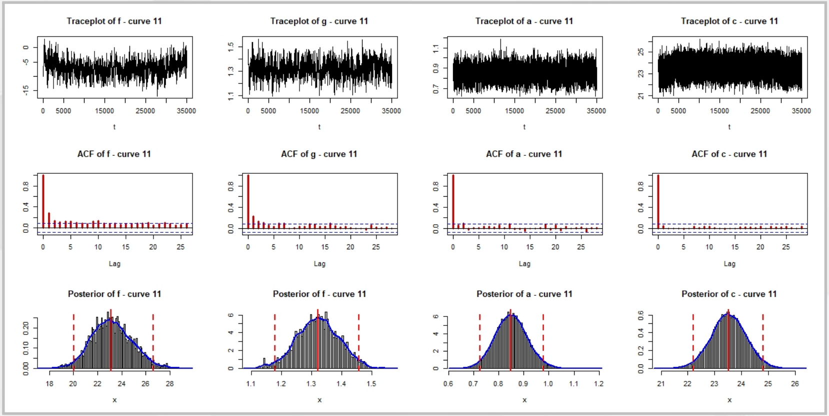


Cross-sectional mean

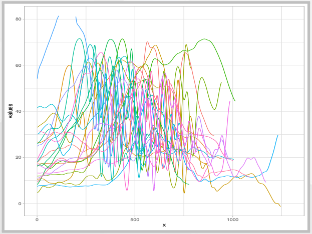


Posterior mean

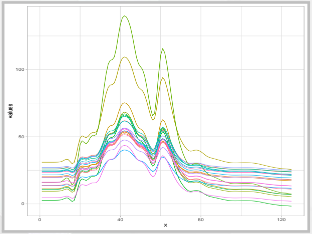
Posterior inference



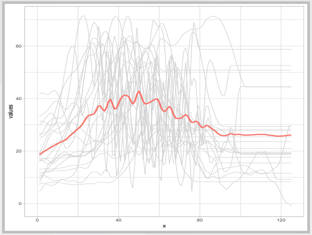
People who had physiotherapy



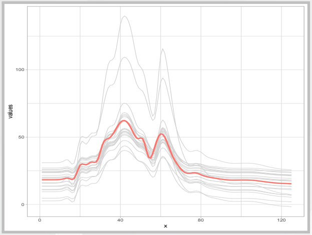
Original curves



Aligned curves

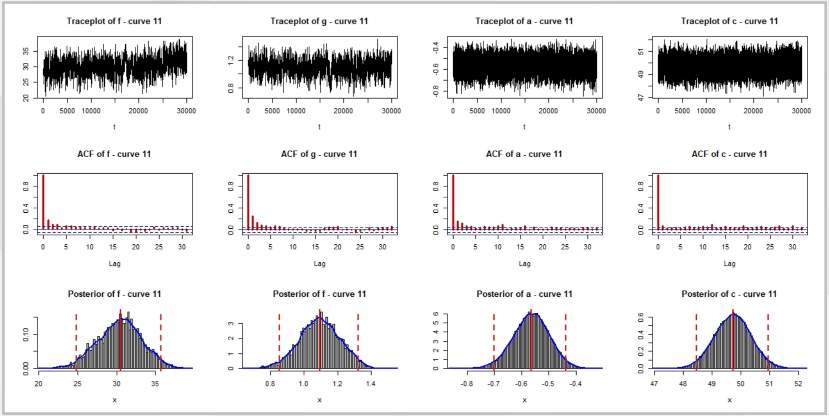


Cross-sectional mean

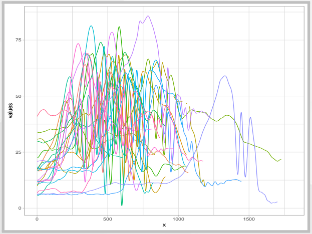


Posterior mean

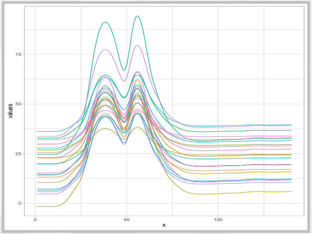
Posterior inference



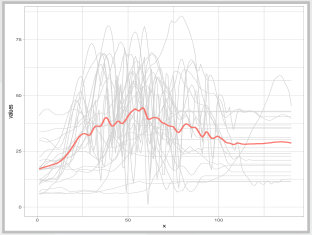
People who had surgery



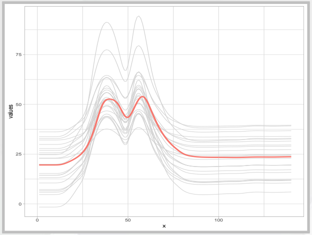
Original curves



Aligned curves

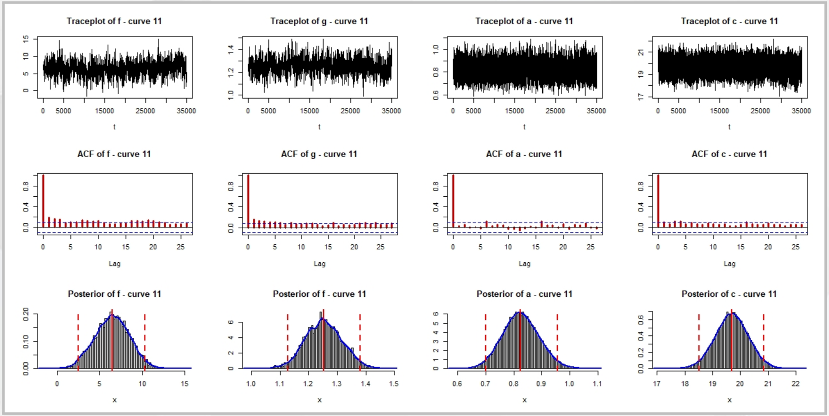


Cross-sectional mean



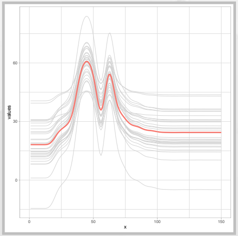
Posterior mean

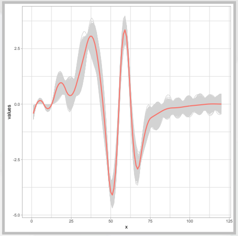
Posterior inference



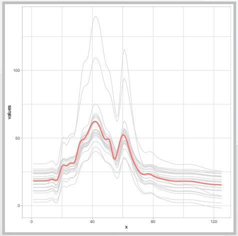
Comparison among groups

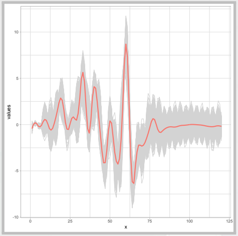
Healthy



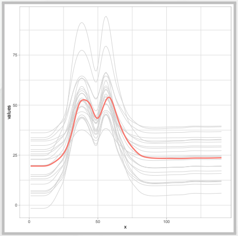


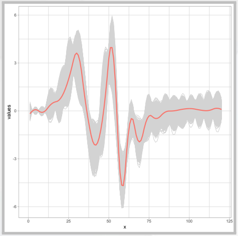
Physio





Surgery





Conclusions

- ☺ The algorithm alignes the curves effectively.
- ☺ Now that the problem of identifiability of the chain has been solved, the algorithm enables a good posterior inference.
- ☺ The method does not rely on presmoothing of the data.
- ☹ The choice of the proposal variance could be improved.

Visit our GitHub repository

`https://github.com/PrincipeFederica/Bayesian-Principe-Mattina-Bighignoli`

Bibliography

Donatello Telesca & Lurdes Y. T. Inoue. : *Bayesian Hierarchical Curve Registration.*
Journal of the American Statistical Association (2008).

Haario, Heikki; Saksman, Eero; Tamminen, Johanna. An adaptive Metropolis algorithm.
Bernoulli 7 (2001), no. 2, 223–242.

Eilers, P., and Marx, B. (1996) : *Flexible Smoothing Using B–Splines and Penalized Likelihood (with discussion).*
Statistical Science, 11, 1200–1224.

Lang, Stefan, and Andreas Brezger. "Bayesian P-splines."
Journal of computational and graphical statistics 13.1 (2004): 183-212.

Simon Jackman: *Bayesian Analysis for the Social Sciences*
Wiley, New York (2009).