Soutenance de thèse de doctorat

Modèles statistiques et algorithmes stochastiques pour l'analyse de données longitudinales à dynamiques multiples et à valeurs sur des variétés riemanniennes

Juliette Chevallier

26 septembre 2019

Sous la direction de
Stéphanie Allassonnière

Centre de mathématiques appliquées (CMAP, UMR 7641 CNRS)

Motivation: Computational Anatomy

- ***A wide range of datasets:*** Images, tensor, meshed surfaces, clinical variables (age, diagnosis, physiological parameters, etc.)

- ***Geometric deformations:*** Deformable template model from [Grenander, 1993], based on the work of [D'Arcy Thompson, 1942].

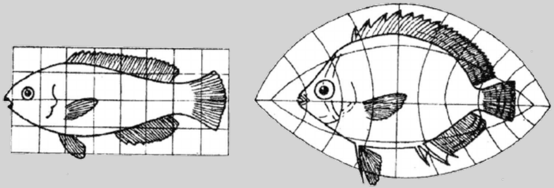


Illustration taken from the book *On Growth and Form* of D'Arcy Thompson.

Motivation: Computational Anatomy

- ***A wide range of datasets:*** Images, tensor, meshed surfaces, clinical variables (age, diagnosis, physiological parameters, etc.)

  $\rightarrow$ Shape spaces

- ***Geometric deformations:*** Deformable template model from [Grenander, 1993], based on the work of [D'Arcy Thompson, 1942].

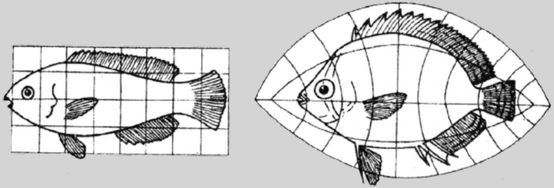


Illustration taken from the book *On Growth and Form* of D'Arcy Thompson.

Motivation: Computational Anatomy

- ***A wide range of datasets:*** Images, tensor, meshed surfaces, clinical variables (age, diagnosis, physiological parameters, etc.)

  $\longrightarrow$ Shape spaces

- ***Geometric deformations:*** Deformable template model from [Grenander, 1993], based on the work of [D'Arcy Thompson, 1942].

  $\longrightarrow$ Riemannian metric: $\inf_{g \in \mathcal{G}} \{ d_{\mathcal{G}}(Id, g) \mid g \cdot x = y \}$.

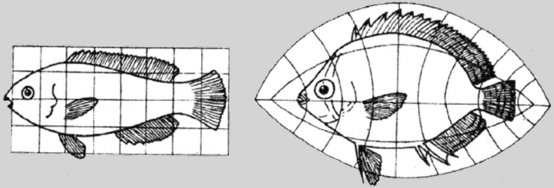


Illustration taken from the book *On Growth and Form* of D'Arcy Thompson.

Outline of the Defense

1. The Usefulness of Studying Longitudinal Riemannian Manifold-Valued Data
2. A Coherent Framework for Longitudinal Observations on a Riemannian Manifold
3. Application to Chemotherapy Monitoring
4. A New Class of EM Algorithms
5. Conclusion and Perspectives

Acknowledgments. *This work was supported by a public grant as part of the Investissement d'avenir, project reference ANR-11-LABX-0056-LMH.*

The Usefulness of Studying Longitudinal Riemannian Manifold-Valued Data

Dataset: Repeated observations of a phenomenon $(t_i, y_i) \in \mathbb{R}^{k_i} \times M^{k_i}$, $i \in [\![1, n]\!]$.

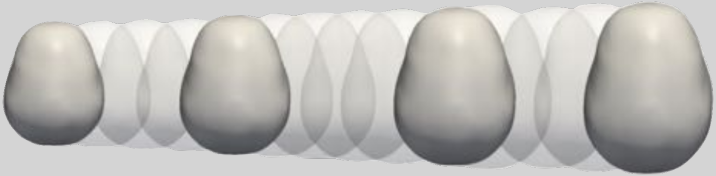


Illustration taken from [Fishbaugh et al., 2017].

Dataset: Repeated observations of a phenomenon $(t_i, y_i) \in \mathbb{R}^{k_i} \times M^{k_i}$, $i \in [\![1, n]\!]$.

Geodesic regression model: [Fletcher, 2011] Let $p \in M$ and $v \in TM$:

$$y_i = \mathcal{E}xp\left(\mathcal{E}xp\left(p\,; t_i v\right); \varepsilon\right), \quad \text{where} \quad \varepsilon \quad \text{is a r.v. valued in} \quad T_{\mathcal{E}xp\,(p\,;t_i v)}M\,.$$

- Estimation performed through a *least squares* method,
- Generalization for multivariate regression [Kim et al., 2014].

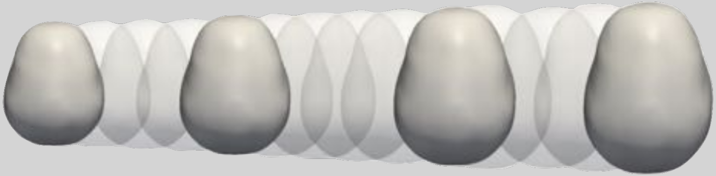


Illustration taken from [Fishbaugh et al., 2017].

A multitude of regression models:

- [Trouvé and Vialard, 2012] Completely different methodology: random perturbation in the Hamiltonian equations that determine the geodesic flow.
  $\longrightarrow$ ***Non-parametric spline regression model***.

- [Fishbaugh et al., 2017] Based on the *deformable template* model.
  Given M_0 and a deformation morphism χ, at each time $M_t = \chi_t(M_0)$.
  $\longrightarrow$ ***Geodesic regression of shapes model*** in the framework of LDDMM.

The Usefulness of Studying Longitudinal Riemannian Manifold-Valued Data

Dataset: Repeated observations of a phenomenon $(t_i, y_i) \in \mathbb{R}^{k_i} \times \mathbb{R}^{k_i}$, $i \in [\![1, n]\!]$.

Basic idea: Two different types of effects:

- *fixed effects* shared by all of the individuals in the population,
- *random effects* specific to each individual.

Dataset: Repeated observations of a phenomenon $(t_i, y_i) \in \mathbb{R}^{k_i} \times \mathbb{R}^{k_i}$, $i \in [\![1, n]\!]$.

Basic idea: Two different types of effects:

- *fixed effects* shared by all of the individuals in the population,
- *random effects* specific to each individual.

Linear mixed effects models: [Laird and Ware, 1982]

$$\forall i \in [\![1, n]\!] , \qquad y_i = A_i \, \alpha + B_i \, \beta_i + \varepsilon_i \ , \quad \text{where} \quad \varepsilon_i \sim \mathcal{N}(0, \Sigma) \, .$$

Dataset: Repeated observations of a phenomenon $(t_i, y_i) \in \mathbb{R}^{k_i} \times \mathbb{R}^{k_i}$, $i \in [\![1, n]\!]$.

Basic idea: Two different types of effects:

- *fixed effects* shared by all of the individuals in the population,
- *random effects* specific to each individual.

Nonlinear mixed effects models: [Sheiner and Beal, 1980, Bates and Watts, 1988]

$$\forall i \in [\![1, n]\!], \quad \forall j \in [\![1, k_i]\!], \qquad \begin{cases} y_{i,j} = f(z_i; t_{i,j}) + \varepsilon_{i,j}, & \text{where} \quad \varepsilon_{i,j} \sim \mathcal{N}(0, \sigma). \\ z_i = A_i\, \alpha + B_i\, \beta_i \end{cases}$$

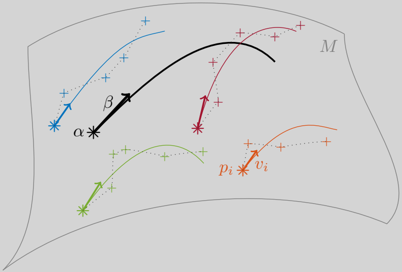
Dataset: Repeated observations of a phenomenon $(t_i, y_i) \in \mathbb{R}^{k_i} \times M^{k_i}$, $i \in [\![1, n]\!]$.

Geodesic hierarchical regression model: [Muralidharan and Fletcher, 2012]
Individual *geodesic* trajectories, themselves random perturbations of a mean *geodesic* path.

$$\begin{cases} y_{i,j} = \mathcal{E}xp\left(\mathcal{E}xp\left(p_i\,;t_{i,j}v_i\right);\,\varepsilon_{i,j}\right) \\ (p_i, v_i) = \mathcal{E}xp_S\left((\alpha,\beta)\,;\,(q_i,w_i)\right) \end{cases},$$

where $\mathcal{E}xp_S =$ exponential map associated with Sasaki's metric on TM.

- Estimation performed *via* a *least squares* method,
- High dependence on the 1st time of measurement.



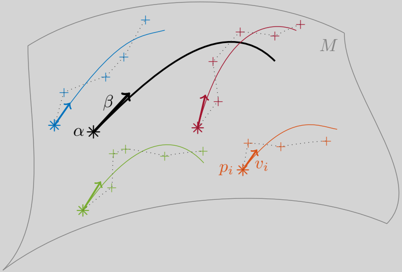
Dataset: Repeated observations of a phenomenon $(t_i, y_i) \in \mathbb{R}^{k_i} \times M^{k_i}$, $i \in [\![1, n]\!]$.

Geodesic hierarchical regression model: [Muralidharan and Fletcher, 2012]
Individual *geodesic* trajectories, themselves random perturbations of a mean *geodesic* path.

$$\begin{cases} y_{i,j} &= \mathcal{E}xp\left(\mathcal{E}xp\left(p_i\,;t_{i,j}v_i\right);\,\varepsilon_{i,j}\right) \\ (p_i, v_i) &= \mathcal{E}xp_{\mathcal{S}}\left((\alpha,\beta)\,;\,(q_i,w_i)\right) \end{cases},$$

where $\mathcal{E}xp_{\mathcal{S}} =$ exponential map associated with Sasaki's metric on TM.

- Estimation performed *via* a *least squares* method,
- High dependence on the 1st time of measurement.



Geodesic hierarchical model for diffeomorphisms: [Singh et al., 2013]
Close link between *groups of deformations* and *shape spaces*.

The Usefulness of Studying Longitudinal Riemannian Manifold-Valued Data

Dataset: n subjects $(S^i)_{i\in[\![1,N]\!]}$ at the corresponding times $(t_j^i)_{i\in[\![1,N]\!],j\in[\![1,k_i]\!]}$.

Spatio-temporal atlas: [Durrleman et al., 2009]

1. *Mean trajectory,*
2. *Individual trajectories.*

Dataset: n subjects $(S^i)_{i\in[\![1,N]\!]}$ at the corresponding times $(t^i_j)_{i\in[\![1,N]\!],j\in[\![1,k_i]\!]}$.

Spatio-temporal atlas: [Durrleman et al., 2009]

1. *Mean trajectory.* Continuous deformation of a template shape: $M_t = \chi_t(M_0)$,
2. *Individual trajectories.*

Dataset: n subjects $(S^i)_{i\in[\![1,N]\!]}$ at the corresponding times $(t^i_j)_{i\in[\![1,N]\!],j\in[\![1,k_i]\!]}$.

Spatio-temporal atlas: [Durrleman et al., 2009]

1. *Mean trajectory.* Continuous deformation of a template shape: $M_t = \chi_t(M_0)$,
2. *Individual trajectories.* Spatio-temporal deformation of the mean trajectory: Given (ψ^i, ϕ^i) , we enforce the matching of $\phi^i\left(M(\psi^i(t^i_j))\right)$ and $S^i(t^i_j)$.

Spatio-temporal atlas: [Durrleman et al., 2009]

1. *Mean trajectory.* Continuous deformation of a template shape: $M_t = \chi_t(M_0)$,
2. *Individual trajectories.* Spatio-temporal deformation of the mean trajectory: Given (ψ^i, ϕ^i), we enforce the matching of $\phi^i\Big(M\big(\psi^i(t_j^i)\big)\Big)$ and $S^i(t_j^i)$:

$$J(\chi, M_0, \psi^i, \phi^i) = \sum_{i=1}^{n} \left\{ \sum_{t_j^i} d\left(\phi^i \left(\chi_{\psi^i(t_j^i)}(M_0) \right) , S^i(t_j^i) \right)^2 + \gamma \operatorname{Reg}(\chi, \phi^i, \psi^i) \right\} .$$

- [Durrleman et al., 2013] Generalization to obtain a generative model,
- Non-parametric model $\longrightarrow$ Difficulties for the estimation,
- [Devilliers et al., 2017] The estimated solution is biased due to noise.

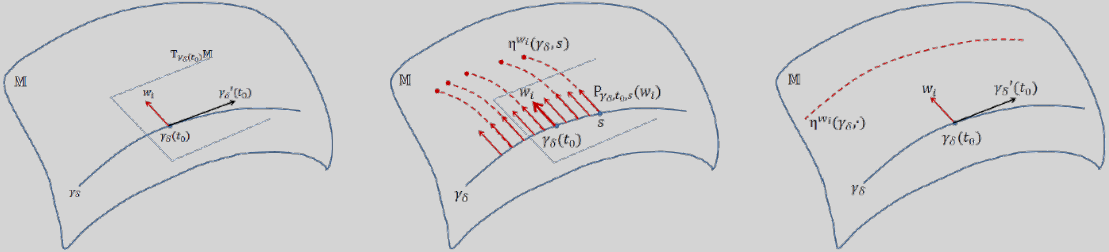
A growing interest in spatio-temporal models:

- [Yang et al., 2011, Delor et al., 2013] Notion of time warps.
- [Hong et al., 2014] Parametric temporal deformations.
  $\longrightarrow$ ***Geodesic regression with parametric time warp model***.
  - Simplified and so efficient algorithmic,
  - Only for geodesic regression.
- [Su et al., 2014] Parametrization to *geometrically* align trajectories.
  - Efficient comparison of the trajectories,
  - Parametrization does not make sense from a modeling perspective.
- [Schiratti et al., 2015] Hierarchical model for the study of longitudinal data.
  $\longrightarrow$ ***Generic spatio-temporal model***.

Parallel Variation of a Curve

Juliette Chevallier

September, 26th 2019

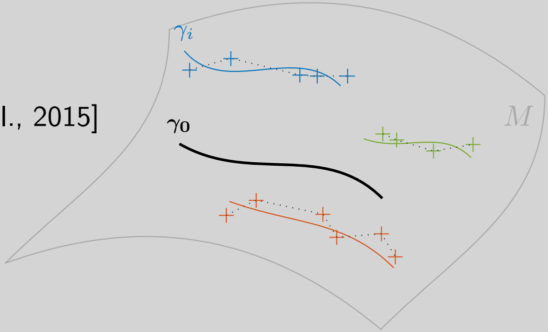


Dataset: $(t_i, y_i) \in \mathbb{R}^{k_i} \times \mathbb{R}^{k_i d}$, $i \in [\![1, n]\!]$.

Generic spatio-temporal model: [Schiratti et al., 2015]
Nonlinear and parametric mixed effects model.

1. *Representative trajectory,*

2. *Individual trajectories.*



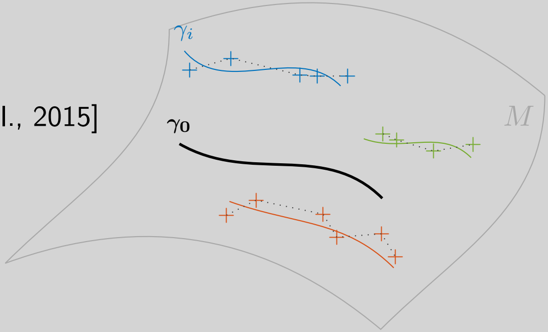
Dataset: $(t_i, y_i) \in \mathbb{R}^{k_i} \times \mathbb{R}^{k_i d}$, $i \in [\![1, n]\!]$.

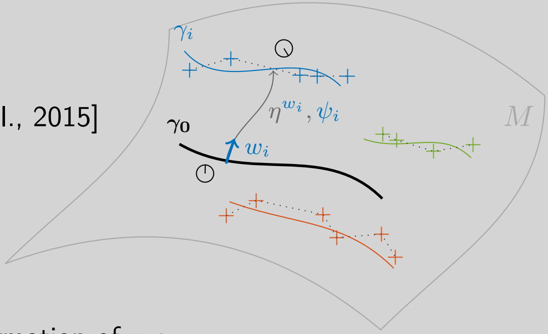
Generic spatio-temporal model: [Schiratti et al., 2015]
Nonlinear and parametric mixed effects model.

1. *Representative trajectory. Geodesic* path:

$$\gamma_0 \colon t \mapsto \mathcal{E}xp_{t_0}(p_0, v_0)(t) ,$$

2. *Individual trajectories.*





Dataset: $(t_i, y_i) \in \mathbb{R}^{k_i} \times \mathbb{R}^{k_i d}$, $i \in [\![1, n]\!]$.

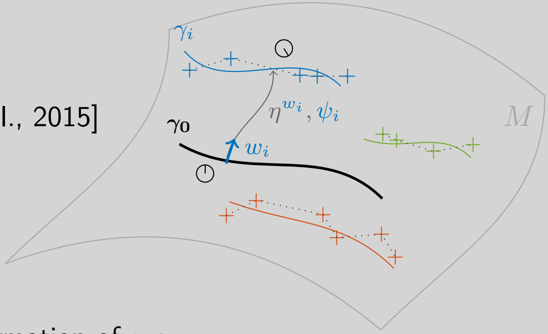
Generic spatio-temporal model: [Schiratti et al., 2015]
Nonlinear and parametric mixed effects model.

1. *Representative trajectory. Geodesic* path:

$$\gamma_0 \colon \; t \; \mapsto \; \mathcal{E}xp_{t_0}\left(p_0, v_0\right)(t) \, ,$$

2. *Individual trajectories.* Spatio-temporal deformation of γ_0:

$$\gamma_i \colon \; t \; \mapsto \; \eta^{w_i}\big(\gamma_0 \, ; \, \psi_i(t)\big) \, , \quad \text{where} \quad \psi_i \colon \; t \; \mapsto \; \alpha_i(t - t_0 - \tau_i) + t_0 \, .$$



Dataset: $(t_i, y_i) \in \mathbb{R}^{k_i} \times \mathbb{R}^{k_i d}$, $i \in [\![1, n]\!]$.

Generic spatio-temporal model: [Schiratti et al., 2015]
Nonlinear and parametric mixed effects model.

1. *Representative trajectory. Geodesic* path:

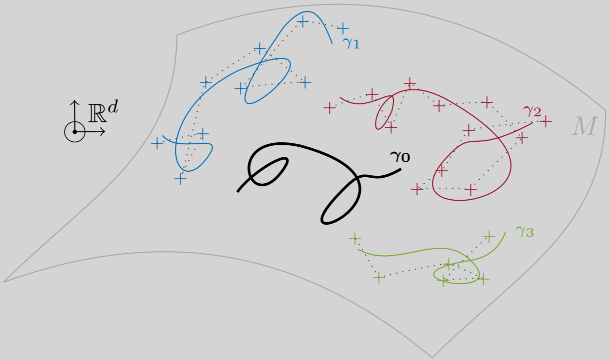
$$\gamma_0 \colon t \mapsto \mathcal{E}xp_{t_0}(p_0, v_0)(t),$$

2. *Individual trajectories.* Spatio-temporal deformation of γ_0:

$$\gamma_i \colon t \mapsto \eta^{w_i}(\gamma_0 \,;\, \psi_i(t)), \quad \text{where} \quad \psi_i \colon t \mapsto \alpha_i(t - t_0 - \tau_i) + t_0.$$

$$y_{i,j} = \eta^{w_i}\Big(\mathcal{E}xp_{t_0}(p_0, v_0) \,;\, \alpha_i(t_{i,j} - t_0 - \tau_i) + t_0\Big) + \varepsilon_{i,j} \quad \text{where} \quad \varepsilon_{i,j} \sim \mathcal{N}(0, \sigma).$$

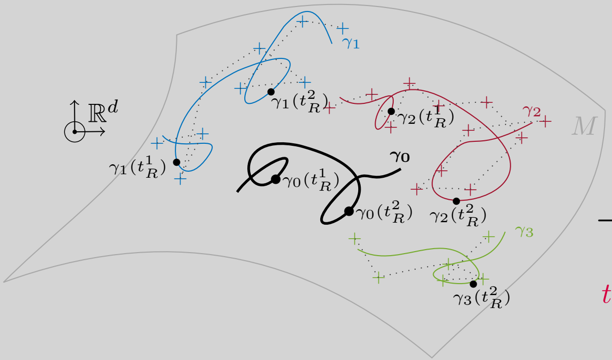
A Coherent Framework for Longitudinal Observations on a Riemannian Manifold



And when the data are only piecewise-geodesically distributed ?

Dataset: $(t_i, y_i) \in \mathbb{R}^{k_i} \times \mathbb{R}^{k_i d}$, $i \in [\![1, n]\!]$.

Number of components m known.
Here, $m = 2$.



And when the data are only piecewise-geodesically distributed ?

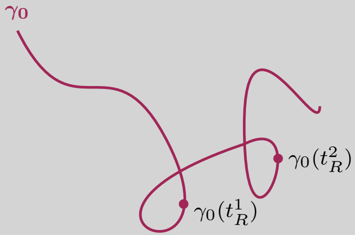
$\longrightarrow$ *Breaking-up* times sequence:

$$t_R = \left(-\infty < t_R^1 < \ldots < t_R^{m-1} < +\infty\right).$$

Dataset: $(t_i, y_i) \in \mathbb{R}^{k_i} \times \mathbb{R}^{k_i d}$, $i \in [\![1, n]\!]$.

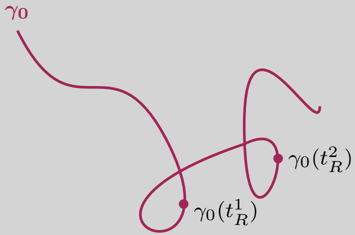
Number of components m known.
Here, $m = 2$.

1. *Breaking-up times* sequence: $t_R = \left(-\infty < t_R^1 < \ldots < t_R^{m-1} < +\infty\right)$, $m \in \mathbb{N}$.



$\rightarrow$ We build γ_0 to be *geodesic* on each $]t_R^{\ell-1}, t_R^{\ell}]$.

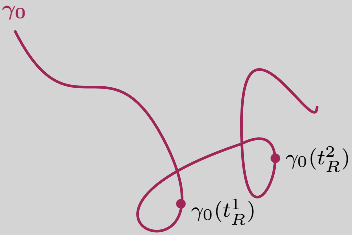
1. *Breaking-up times* sequence: $t_R = \left(-\infty < t_R^1 < \ldots < t_R^{m-1} < +\infty\right)$, $m \in \mathbb{N}$.



$\rightarrow$ We build γ_0 to be *geodesic* on each $]t_R^{\ell-1}, t_R^{\ell}]$.

2. *Template manifold* $M_0 \subset \mathbb{R}^d$ geodesically complete,
   - $\left(\bar{\gamma}_0^{\ell}\right)_{\ell \in [\![1,m]\!]}$ geodesics on M_0 ,
   - $\left(\phi_0^{\ell}\right)_{\ell \in [\![1,m]\!]}$ isometries defined on M_0 ;
3. $\forall \ell \in [\![1, m]\!]$, $M_0^{\ell} = \phi_0^{\ell}(M_0)$ and $\gamma_0^{\ell} = \phi_0^{\ell} \circ \bar{\gamma}_0^{\ell}$;

1. *Breaking-up times* sequence: $t_R = \left(-\infty < t_R^1 < \ldots < t_R^{m-1} < +\infty\right)$, $m \in \mathbb{N}$.



$\rightarrow$ We build γ_0 to be *geodesic* on each $]t_R^{\ell-1}, t_R^{\ell}]$.

2. *Template manifold* $M_0 \subset \mathbb{R}^d$ geodesically complete,
   - $\left(\bar{\gamma}_0^\ell\right)_{\ell \in [\![1,m]\!]}$ geodesics on M_0 ,
   - $\left(\phi_0^\ell\right)_{\ell \in [\![1,m]\!]}$ isometries defined on M_0 ;
3. $\forall \ell \in [\![1, m]\!], \quad M_0^\ell = \phi_0^\ell(M_0) \quad$ and $\quad \gamma_0^\ell = \phi_0^\ell \circ \bar{\gamma}_0^\ell$;
4. *Piecewise-geodesic curve*:

$$\gamma_0 = \gamma_0^1 \, \mathbb{1}_{]-\infty, t_R^1]} + \sum_{\ell=2}^{m-1} \gamma_0^\ell \, \mathbb{1}_{]t_R^{\ell-1}, t_R^\ell]} + \gamma_0^m \, \mathbb{1}_{]t_R^{m-1}, +\infty[}$$

5. *Boundary conditions* on the rupture times to ensure continuity.

A Geometric Model

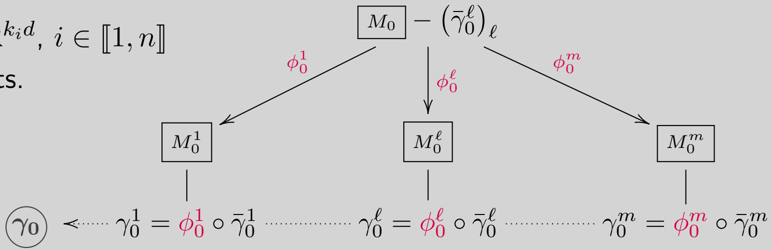
$\boxed{M_0} - (\bar{\gamma}_0^\ell)_\ell$

Dataset: $(t_i, y_i) \in \mathbb{R}^{k_i} \times \mathbb{R}^{k_i d}$, $i \in [\![1, n]\!]$

m diffeomorphic components.

Dataset: $(t_i, y_i) \in \mathbb{R}^{k_i} \times \mathbb{R}^{k_i d}$, $i \in [\![1, n]\!]$

m diffeomorphic components.



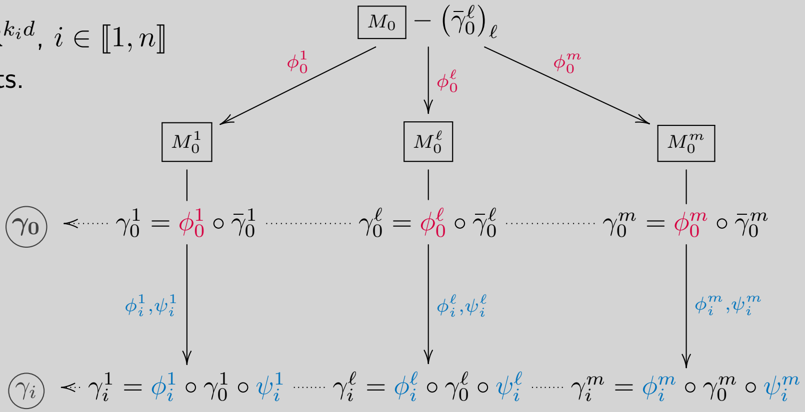
1. *Population variable:*
   $z_{\text{pop}} \longleftrightarrow (\phi_0^\ell)_{\ell \in [\![1,m]\!]}$;

Juliette Chevallier

September, 26th 2019

Dataset: $(t_i, y_i) \in \mathbb{R}^{k_i} \times \mathbb{R}^{k_i d}$, $i \in [\![1, n]\!]$

m diffeomorphic components.



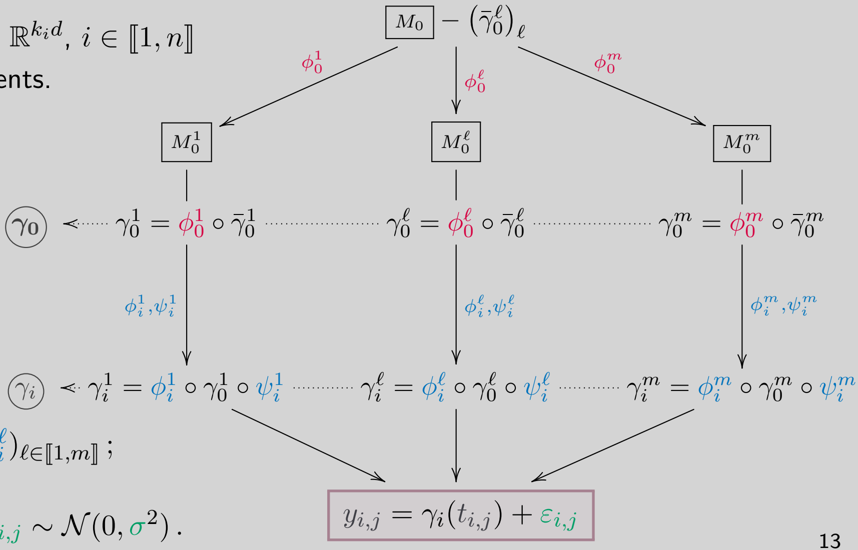
1. *Population variable:*
   $z_{\text{pop}} \longleftrightarrow (\phi_0^\ell)_{\ell \in [\![1,m]\!]}$;

2. *Individual (spatial and temporal) variables:*
   $z_i = (z_i^\phi, z_i^\psi) \longleftrightarrow (\phi_i^\ell, \psi_i^\ell)_{\ell \in [\![1,m]\!]}$;

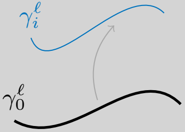
Dataset: $(t_i, y_i) \in \mathbb{R}^{k_i} \times \mathbb{R}^{k_i d}$, $i \in [\![1, n]\!]$

m diffeomorphic components.

1. *Population variable:*
   $z_{\text{pop}} \longleftrightarrow (\phi_0^\ell)_{\ell \in [\![1,m]\!]}$;
2. *Individual (spatial and temporal) variables:*
   $z_i = (z_i^\phi, z_i^\psi) \longleftrightarrow (\phi_i^\ell, \psi_i^\ell)_{\ell \in [\![1,m]\!]}$;
3. *Gaussian noise:* $\sigma \longleftrightarrow \varepsilon_{i,j} \sim \mathcal{N}(0, \sigma^2)$.

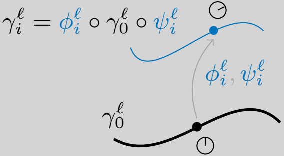


Let $i \in [\![1, n]\!]$. We build γ_i to derive from γ_0 through spatiotemporal transformations.



Individual Trajectories γ_i – Space and Time Warping

Let $i \in [\![1, n]\!]$. We build γ_i to derive from γ_0 through spatiotemporal transformations.

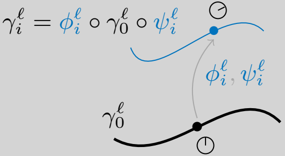


1. *Space warps* $(\phi_i^\ell)_{\ell \in [\![1,m]\!]}$ defined in view of applications, diffeomorphics, satisfy $\phi_i^\ell \circ \gamma_0^\ell(t_R^\ell) = \phi_i^\ell \circ \gamma_0^{\ell+1}(t_R^\ell)$;
2. *Time warps* $(\psi_i^\ell)_{\ell \in [\![1,m]\!]}$: Let $(\alpha_i^\ell)_{\ell \in [\![1,m]\!]} \in \mathbb{R}_+^m$ and $\tau_i^1 \in \mathbb{R}$,

$$t_{R,i}^\ell := t_R^{\ell-1} + \tau_i^\ell + \frac{t_R^\ell - t_R^{\ell-1}}{\alpha_i^\ell}, \quad \tau_i^{\ell+1} := t_{R,i}^\ell - t_R^\ell \quad \text{and} \quad \psi_i^\ell \colon t \mapsto \alpha_i^\ell \big(t - t_R^{\ell-1} - \tau_i^\ell\big) + t_R^{\ell-1} \, ;$$

Individual Trajectories γ_i – Space and Time Warping

Let $i \in [\![1, n]\!]$. We build γ_i to derive from γ_0 through spatiotemporal transformations.



1. *Space warps* $(\phi_i^\ell)_{\ell \in [\![1,m]\!]}$ defined in view of applications, diffeomorphics, satisfy $\phi_i^\ell \circ \gamma_0^\ell(t_R^\ell) = \phi_i^\ell \circ \gamma_0^{\ell+1}(t_R^\ell)$;

2. *Time warps* $(\psi_i^\ell)_{\ell \in [\![1,m]\!]}$: Let $(\alpha_i^\ell)_{\ell \in [\![1,m]\!]} \in \mathbb{R}_+^m$ and $\tau_i^1 \in \mathbb{R}$,

$$t_{R,i}^\ell := t_R^{\ell-1} + \tau_i^\ell + \frac{t_R^\ell - t_R^{\ell-1}}{\alpha_i^\ell}, \quad \tau_i^{\ell+1} := t_{R,i}^\ell - t_R^\ell \quad \text{and} \quad \psi_i^\ell : t \mapsto \alpha_i^\ell \left(t - t_R^{\ell-1} - \tau_i^\ell\right) + t_R^{\ell-1} ;$$

3. $\gamma_i^\ell := \phi_i^\ell \circ \gamma_0^\ell \circ \psi_i^\ell$ and

$$\gamma_i = \gamma_i^1 \, \mathbb{1}_{]-\infty, t_{R,i}^1]} + \sum_{\ell=2}^{m-1} \gamma_i^\ell \, \mathbb{1}_{]t_{R,i}^{\ell-1}, t_{R,i}^\ell]} + \gamma_i^m \, \mathbb{1}_{]t_{R,i}^{m-1}, +\infty[}$$

4. *Gaussian noise:* $y_{i,j} = \gamma_i(t_{i,j}) + \varepsilon_{i,j}$ where $\varepsilon_{i,j} \sim \mathcal{N}(0, \sigma^2)$ and $\sigma \in \mathbb{R}^+$.

A Coherent Framework for Longitudinal Observations on a Riemannian Manifold

Dataset: $(t_i, y_i) \in \mathbb{R}^{k_i} \times \mathbb{R}^{k_i d}$, $i \in [\![1, n]\!]$, m diffeomorphic components.

A parametric model:

- *Population variable:* $z_{\text{pop}} \in \mathcal{Z}_{\text{pop}} \subset \mathbb{R}^{p_{\text{pop}}}$;
- *Individual variables:* $z_i \in \mathcal{Z}_i \subset \mathbb{R}^{p_{\text{ind}}}$.

Parameters: $\theta = (z_{\text{pop}}, z_i, \sigma)$;

$$\Theta = \mathcal{Z}_{\text{pop}} \times \mathcal{Z}_i \times \mathbb{R}^+.$$

Dataset: $(t_i, y_i) \in \mathbb{R}^{k_i} \times \mathbb{R}^{k_i d}$, $i \in [\![1, n]\!]$, m diffeomorphic components.

A parametric model:

- *Population variable:* $z_{\text{pop}} \in \mathcal{Z}_{\text{pop}} \subset \mathbb{R}^{p_{\text{pop}}}$; [Kuhn and Lavielle, 2005]
  Problem: Our model doesn't belong to the *curved exponential family*.
  Solution: $z_{\text{pop}} \sim \mathcal{N}(\overline{z_{\text{pop}}}, \mathcal{D}_{\text{pop}}^{-1})$, with $\mathcal{D}_{\text{pop}} = \sigma_{\text{pop}}^2 I_{p_{\text{pop}}}$.
- *Individual variables:* $z_i \in \mathcal{Z}_i \subset \mathbb{R}^{p_{\text{ind}}}$.

Parameters: $\theta = (\overline{z_{\text{pop}}}\, , \, z_i \, , \, \sigma)$;
$\Theta = \mathbb{R}^{p_{\text{pop}}} \times \mathcal{Z}_i \times \mathbb{R}^+$.

Hyper-parameter: $\sigma_{\text{pop}} \in \mathbb{R}^+$.

Dataset: $(t_i, y_i) \in \mathbb{R}^{k_i} \times \mathbb{R}^{k_i d}$, $i \in [\![1, n]\!]$, m diffeomorphic components.

A parametric model:

- *Population variable:* $z_{\text{pop}} \sim \mathcal{N}(\overline{z_{\text{pop}}}, \mathcal{D}_{\text{pop}}^{-1}) \in \mathcal{Z}_{\text{pop}}$, with $\mathcal{D}_{\text{pop}} = \sigma_{\text{pop}}^2 I_{p_{\text{pop}}}$.
- *Individual variables:* $z_i \in \mathcal{Z}_i \subset \mathbb{R}^{p_{\text{ind}}}$;

  Problem: interested in the individual behaviours w.r.t. the characteristic one.

  Solution: $z_i \sim \mathcal{N}(0, \Sigma)$, where $\Sigma \in \mathcal{S}_{p_{\text{ind}}}^{+}(\mathbb{R})$.

Parameters: $\theta = (\overline{z_{\text{pop}}} , \Sigma , \sigma)$;

$\Theta = \mathbb{R}^{p_{\text{pop}}} \times \mathcal{S}_{p_{\text{ind}}}^{+}(\mathbb{R}) \times \mathbb{R}^{+}$.

Hyper-parameter: $\sigma_{\text{pop}} \in \mathbb{R}^{+}$.

Dataset: $(t_i, y_i) \in \mathbb{R}^{k_i} \times \mathbb{R}^{k_i d}$, $i \in [\![1, n]\!]$, m diffeomorphic components.

A parametric model:

- *Population latent variable:* $z_{\text{pop}} \sim \mathcal{N}(\overline{z_{\text{pop}}}, \mathcal{D}_{\text{pop}}^{-1}) \in \mathcal{Z}_{\text{pop}}$, with $\mathcal{D}_{\text{pop}} = \sigma_{\text{pop}}^2 I_{p_{\text{pop}}}$.
- *Individual latent variables:* $z_i \sim \mathcal{N}(0, \Sigma)$, where $\Sigma \in \mathcal{S}_{p_{\text{ind}}}^{+}(\mathbb{R})$.

Parameters: $\theta = (\overline{z_{\text{pop}}}, \Sigma, \sigma)$;

$$\Theta = \mathbb{R}^{p_{\text{pop}}} \times \mathcal{S}_{p_{\text{ind}}}^{+}(\mathbb{R}) \times \mathbb{R}^{+}.$$

Hyper-parameter: $\sigma_{\text{pop}} \in \mathbb{R}^{+}$.

A hierarchical model: $z = (z_{\text{pop}}, z_i)$

$$\begin{cases} y \,|\, z, \theta \sim \displaystyle\bigotimes_{i=1}^{n} \bigotimes_{j=1}^{k_i} \mathcal{N}\left(\gamma_i(t_{i,j}), \sigma^2\right) \\ z \,|\, \theta \sim \mathcal{N}(\overline{z_{\text{pop}}}, \mathcal{D}_{\text{pop}}^{-1}) \displaystyle\bigotimes_{i=1}^{n} \mathcal{N}(0, \Sigma) \end{cases}$$

Dataset: $(t_i, y_i) \in \mathbb{R}^{k_i} \times \mathbb{R}^{k_i d}$, $i \in [\![1, n]\!]$, m diffeomorphic components.

A parametric model:

- *Population latent variable:* $z_{\text{pop}} \sim \mathcal{N}(\overline{z_{\text{pop}}}, \mathcal{D}_{\text{pop}}^{-1}) \in \mathcal{Z}_{\text{pop}}$, with $\mathcal{D}_{\text{pop}} = \sigma_{\text{pop}}^2 I_{p_{\text{pop}}}$.
- *Individual latent variables:* $z_i \sim \mathcal{N}(0, \Sigma)$, where $\Sigma \in \mathcal{S}_{p_{\text{ind}}}^+(\mathbb{R})$.

Parameters: $\theta = (\overline{z_{\text{pop}}}\, , \Sigma\, , \sigma)$;

$$\Theta = \mathbb{R}^{p_{\text{pop}}} \times \mathcal{S}_{p_{\text{ind}}}^+(\mathbb{R}) \times \mathbb{R}^+.$$

Hyper-parameters: $\sigma_{\text{pop}} \in \mathbb{R}^+$,

$$V \in \mathcal{S}_{p_{\text{ind}}}^+(\mathbb{R}), \quad v,\, m_\Sigma,\, m_\sigma \in \mathbb{R}\,.$$

A hierarchical model: $z = (z_{\text{pop}}, (z_i)_i)$

$$\begin{cases} y \,|\, z, \theta \sim \bigotimes_{i=1}^{n} \bigotimes_{j=1}^{k_i} \mathcal{N}\left(\gamma_i(t_{i,j}), \sigma^2\right) \\ z \,|\, \theta \sim \mathcal{N}(\overline{z_{\text{pop}}}, \mathcal{D}_{\text{pop}}^{-1}) \bigotimes_{i=1}^{n} \mathcal{N}(0, \Sigma) \\ (\Sigma, \sigma) \sim \mathcal{W}^{-1}(V, m_\Sigma) \otimes \mathcal{W}^{-1}(v, m_\sigma) \end{cases}$$

A Coherent Framework for Longitudinal Observations on a Riemannian Manifold

Theorem: Existence of the MAP estimator

Given a piecewise-geodesic model and the choice of probability distributions for the parameters and latent variables of the model, for any dataset $(t_{i,j}, y_{i,j})_{(i,j) \in [\![1,n]\!] \times [\![1,k_i]\!]}$, there exists $\widehat{\theta}_{\mathrm{MAP}} \in \underset{\theta \in \Theta}{\operatorname{argmax}}\ q(\theta | y)$.

Admissible parameters: For all $\omega \in \mathbb{R}$, let $\mathbb{E}^*(\omega) = \sup_{\theta \in \Theta^\omega} \mathbb{E}_{P(\mathrm{d}y)} \left[\log q(y|\theta) \right]$ and

$$\Theta_*^\omega = \left\{ \theta \in \Theta^\omega \;\middle|\; \| \overline{z_{\text{pop}}} \|_2 \leqslant \omega \;\wedge\; \mathbb{E}_{P(\mathrm{d}y)} \left[\log q(y|\theta) \right] = \mathbb{E}^*(\omega) \right\}.$$

Admissible parameters: For all $\omega \in \mathbb{R}$, let $\mathbb{E}^*(\omega) = \sup_{\theta \in \Theta^\omega} \mathbb{E}_{P(\mathrm{d}y)}\left[\log q(y|\theta)\right]$ and

$$\Theta^\omega_* = \left\{ \theta \in \Theta^\omega \;\middle|\; \| \overline{z_{\mathsf{pop}}} \|_2 \leqslant \omega \wedge \mathbb{E}_{P(\mathrm{d}y)}\left[\log q(y|\theta)\right] = \mathbb{E}^*(\omega) \right\}.$$

Two kinds of latent variables: critical vs regular: $z = \left(z_{\mathsf{pop}}^{\mathsf{crit}}, z_{\mathsf{pop}}^{\mathsf{reg}}, \left(z_i^{\mathsf{crit}}, z_i^{\mathsf{reg}} \right)_i \right)$

Admissible parameters: For all $\omega \in \mathbb{R}$, let $\mathbb{E}^*(\omega) = \sup_{\theta \in \Theta^\omega} \mathbb{E}_{P(\mathrm{d}y)} \left[\log q(y|\theta) \right]$ and

$$\Theta_*^\omega = \left\{ \theta \in \Theta^\omega \;\middle|\; \| \overline{z_{\text{pop}}} \|_2 \leqslant \omega \wedge \mathbb{E}_{P(\mathrm{d}y)} \left[\log q(y|\theta) \right] = \mathbb{E}^*(\omega) \right\}.$$

Two kinds of latent variables: critical vs regular: $z = \left(z_{\text{pop}}^{\text{crit}}, z_{\text{pop}}^{\text{reg}}, \left(z_i^{\text{crit}}, z_i^{\text{reg}} \right)_i \right)$

Critical variables induce critical trajectories

$\forall v \in [\![1, p_{\text{pop}}^{\text{crit}} + p_{\text{ind}}^{\text{crit}}]\!]$, there exists $\gamma_{i,v}^{\text{crit}}$ s.t.

$$\lim_{|(z_{\text{pop}}^{\text{crit}}, z_i^{\text{crit}})_v| \to +\infty} \vec{\gamma}_i(z_{\text{pop}}, z_i) = \gamma_{i,v}^{\text{crit}}$$

and $\mathcal{L}_{k_i}(\{y_i = \gamma_{i,v}^{\text{crit}}\}) = 0\,.$

Consistency of the Maximum A Posteriori estimator

Juliette Chevallier
September, 26th 2019

Admissible parameters: For all $\omega \in \mathbb{R}$, let $\mathbb{E}^*(\omega) = \sup_{\theta \in \Theta^\omega} \mathbb{E}_{P(\mathrm{d}y)}\left[\log q(y|\theta)\right]$ and

$$\Theta_*^\omega = \left\{ \theta \in \Theta^\omega \;\middle|\; \| \overline{z_{\text{pop}}} \|_2 \leqslant \omega \wedge \mathbb{E}_{P(\mathrm{d}y)}\left[\log q(y|\theta)\right] = \mathbb{E}^*(\omega) \right\}.$$

Two kinds of latent variables: critical vs regular: $z = \left(z_{\text{pop}}^{\text{crit}}, z_{\text{pop}}^{\text{reg}}, \left(z_i^{\text{crit}}, z_i^{\text{reg}} \right)_i \right)$

Critical variables induce critical trajectories	Ind. traj. grow super-linearly w.r.t. *regular variables*
$\forall v \in [\![1, p_{\text{pop}}^{\text{crit}} + p_{\text{ind}}^{\text{crit}}]\!]$, there exists $\gamma_{i,v}^{\text{crit}}$ s.t. $$\lim_{\lvert (z_{\text{pop}}^{\text{crit}}, z_i^{\text{crit}})_v \rvert \to +\infty} \vec{\gamma}_i(z_{\text{pop}}, z_i) = \gamma_{i,v}^{\text{crit}}$$ and $$\mathcal{L}_{k_i}(\{y_i = \gamma_{i,v}^{\text{crit}}\}) = 0.$$	$\forall v \in [\![1, p_{\text{pop}}^{\text{reg}} + p_{\text{ind}}^{\text{reg}}]\!]$, there exists $a_{i,v}$, $b_{i,v}$ • depend only of $(z_{\text{pop}}^{\text{reg}}, z_i^{\text{reg}})_{-v}$, • $a_{i,v}(z_{-v}^{\text{reg}}) \geqslant 0$; $a_{i,v}(z_{-v}^{\text{reg}}) = 0$ iff $z_{-v}^{\text{reg}} = 0$. $$\|\vec{\gamma}_i(z_{\text{pop}}, z_i)\|_\infty \geqslant a_{i,v}(\cdot) \lvert (z_{\text{pop}}^{\text{reg}}, z_i^{\text{reg}})_v \rvert + b_{i,v}(\cdot).$$

(H 1) $p^\ell = p_{\text{pop}} + \ell\, p_{\text{ind}} < \sum_{i=1}^{\ell} k_i$;

(H 2) Times of registration t_i i.i.d;

(H 3) $P(\mathrm{d}y^\ell)$ is continuous with polynomial tail decay of degree bigger than $p^\ell + 1$, apart from a compact $K \subset \mathbb{R}^{k^\ell}$;

(H 4) Latent variables are either *critical* or *regular*.

Theorem: Consistency of the MAP estimator

Assume that there exists $\ell \in [\![1, n]\!]$ s.t. (H 1-4) hold. Let $(\hat{\theta}_m)_{m\in\mathbb{N}}$ denote any MAP estimator. Then $\Theta_*^\omega \neq \varnothing$ and for any $\varepsilon \in \mathbb{R}_+^*$,

$$\lim_{m\to\infty} \mathbb{P}\left[\delta(\hat{\theta}_m, \Theta_*^\omega) \geqslant \varepsilon\right] = 0$$

where δ in any metric compatible with the topology on Θ^ω.

Consistency of the Maximum A Posteriori estimator

Sketch of the proof: Consider the Alexandrov one-point compactification $\overline{\Theta^\omega} = \Theta^\omega \cup \{\infty\}$.

1. *Main difficulty:* To prove that $\Theta_*^\omega \neq \varnothing$.
2. Given $\Theta_*^\omega \neq \varnothing$, we follow the classical analysis of [van der Vaart, 2000].

Sketch of the proof: Consider the Alexandrov one-point compactification $\overline{\Theta^\omega} = \Theta^\omega \cup \{\infty\}$.

1. *Main difficulty:* To prove that $\Theta_*^\omega \neq \varnothing$.

   1.1 $\theta \mapsto \mathbb{E}_{P(\mathrm{d}\boldsymbol{y}^\ell)}\left[\log q(y|\theta)\right]$ is continuous on Θ^ω. So, if for any sequence $(\theta_\kappa)_\kappa$ s.t.

   $$\lim_{\kappa\to\infty} \theta_\kappa \in \overline{\Theta^\omega} \setminus \Theta^\omega, \; \lim_{\kappa\to+\infty} \mathbb{E}_{P(\mathrm{d}\boldsymbol{y}^\ell)}\left[\sum_{i=1}^{\ell} \log q(\boldsymbol{y}_i|\theta_\kappa)\right] = -\infty \text{ then } \Theta_*^\omega \neq \varnothing\,.$$

2. Given $\Theta_*^\omega \neq \varnothing$, we follow the classical analysis of [van der Vaart, 2000].

Sketch of the proof: Consider the Alexandrov one-point compactification $\overline{\Theta^\omega} = \Theta^\omega \cup \{\infty\}$.

1. *Main difficulty:* To prove that $\Theta_*^\omega \neq \varnothing$.

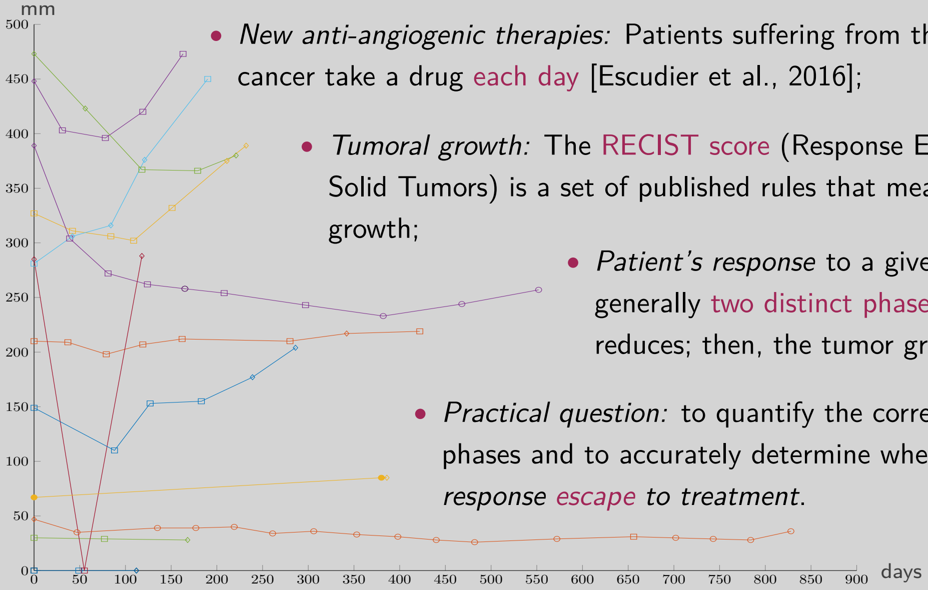
   1.1 $\theta \mapsto \mathbb{E}_{P(\mathrm{d}\boldsymbol{y}^\ell)}\left[\log q(y|\theta)\right]$ is continuous on Θ^ω. So, if for any sequence $(\theta_\kappa)_\kappa$ s.t.
   $$\lim_{\kappa\to\infty} \theta_\kappa \in \overline{\Theta^\omega} \setminus \Theta^\omega, \quad \lim_{\kappa\to+\infty} \mathbb{E}_{P(\mathrm{d}\boldsymbol{y}^\ell)}\left[\sum_{i=1}^{\ell} \log q(\boldsymbol{y}_i|\theta_\kappa)\right] = -\infty \text{ then } \Theta_*^\omega \neq \varnothing\,.$$

   1.2 ***Negative part:*** $\displaystyle\lim_{\kappa\to\infty} \sum_{i=1}^{\ell} \log q\left(\boldsymbol{y}_i|\theta_\kappa\right) = -\infty \quad P(\mathrm{d}\boldsymbol{y}^\ell)$ a.s. for such any sequence $(\theta_\kappa)_\kappa$.
   From the monotone convergence theorem we then have $\displaystyle\liminf_{\kappa\to+\infty} \mathbb{E}_{P(\mathrm{d}\boldsymbol{y}^\ell)}\left[\left(f_\kappa(\boldsymbol{y}^\ell)\right)^-\right] = +\infty\,;$

   1.3 ***Positive part:*** $\displaystyle\mathbb{E}_{P(\mathrm{d}\boldsymbol{y}^\ell)}\left[\sup_{\theta\in\Theta}\left(\sum_{i=1}^{\ell}\log q(\boldsymbol{y}_i|\theta)\right)^+\right] < +\infty.$
   So, according to the dominated convergences theorem, $\displaystyle\lim_{\kappa\to+\infty} \mathbb{E}_{P(\mathrm{d}\boldsymbol{y}^\ell)}\left[\left(f_\kappa(\boldsymbol{y}^\ell)\right)^+\right] = 0\,;$

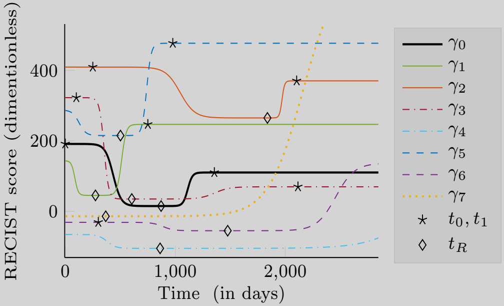
2. Given $\Theta_*^\omega \neq \varnothing$, we follow the classical analysis of [van der Vaart, 2000].

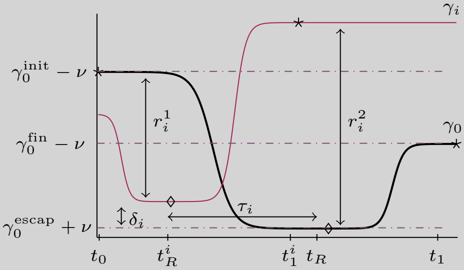
Application to Chemotherapy Monitoring



- *New anti-angiogenic therapies:* Patients suffering from the metastatic kidney cancer take a drug each day [Escudier et al., 2016];
- *Tumoral growth:* The RECIST score (Response Evaluation Criteria In Solid Tumors) is a set of published rules that measures the tumoral growth;
- *Patient's response* to a given treatment has generally two distinct phases: first, tumor's size reduces; then, the tumor grows again;
- *Practical question:* to quantify the correlation between both phases and to accurately determine when *the patient's response escape to treatment*.

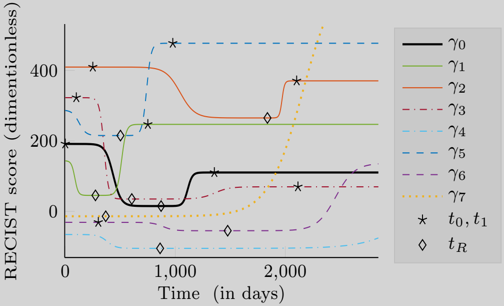
The Piecewise-Logistic Curve Model

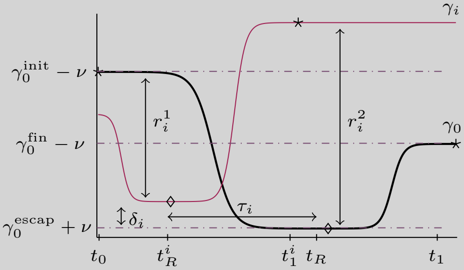




We set $m = 2$, $d = 1$ and $M_0 =]0, 1[$ equipped with the logistic metric. Let $\nu \in \mathbb{R}$.

1. *Representative path γ_0:*
   - M_0 is map onto $]\gamma_0^{\text{escap}}, \gamma_0^{\text{init}}[$ and $]\gamma_0^{\text{escap}}, \gamma_0^{\text{fin}}[$ through affine transformations,
   - require that $\gamma_0^1(t_R) = \gamma_0^2(t_R) = \gamma_0^{\text{escap}} + \nu$, $\gamma_0^1(t_0) = \gamma_0^{\text{init}} - \nu$ and $\gamma_0^2(t_1) = \gamma_0^{\text{fin}} - \nu$;





We set $m = 2$, $d = 1$ and $M_0 =]0, 1[$ equipped with the logistic metric. Let $\nu \in \mathbb{R}$.

1. *Representative path γ_0:*
   - M_0 is map onto $]\gamma_0^{\text{escap}}, \gamma_0^{\text{init}}[$ and $]\gamma_0^{\text{escap}}, \gamma_0^{\text{fin}}[$ through affine transformations,
   - require that $\gamma_0^1(t_R) = \gamma_0^2(t_R) = \gamma_0^{\text{escap}} + \nu$, $\gamma_0^1(t_0) = \gamma_0^{\text{init}} - \nu$ and $\gamma_0^2(t_1) = \gamma_0^{\text{fin}} - \nu$;

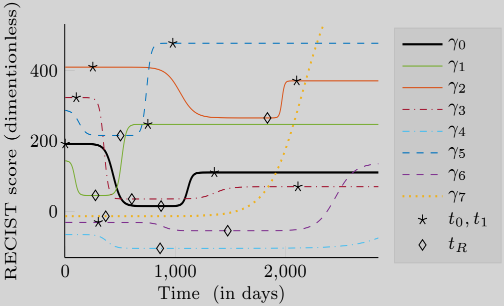
2. *Individual trajectories γ_i:*
   - *Time warps (ψ_i^1, ψ_i^2):* We set $\alpha_i^\ell = e^{\xi_i^\ell}$,

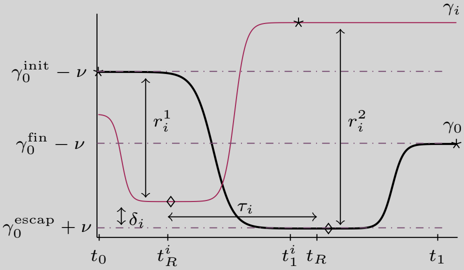
     $$\psi_i^\ell : t \mapsto e^{\xi_i^\ell} \left(t - t_0\right) + t_0 + \tau_i \,;$$

   - *Space warps (ϕ_i^1, ϕ_i^2):* Given $(\rho_i^1, \rho_i^2, \delta_i) \in \mathbb{R}^3$,

     $$\phi_i^\ell : x \mapsto e^{\rho_i^\ell} \left(x - \gamma_0(t_R)\right) + \gamma_0(t_R) + \delta_i \,.$$

The Piecewise-Logistic Curve Model



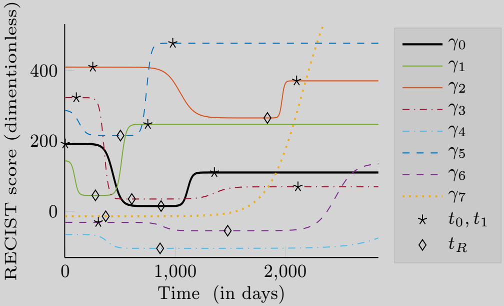


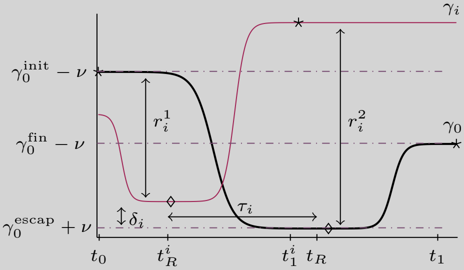
We set $m = 2$, $d = 1$ and $M_0 =]0, 1[$.

1. *Representative path γ_0:* succession of two logistics curves between γ_0^{init}, γ_0^{escap} and γ_0^{fin};
2. *Individual trajectories γ_i:* Space and time warps

$$\psi_i^\ell : t \mapsto e^{\xi_i^\ell} (t - t_0) + t_0 + \tau_i \,,$$

$$\phi_i^\ell : x \mapsto e^{\rho_i^\ell} (x - \gamma_0(t_R)) + \gamma_0(t_R) + \delta_i \,;$$





We set $m = 2$, $d = 1$ and $M_0 =]0, 1[$.

1. *Representative path* γ_0*:* succession of two logistics curves between γ_0^{init}, γ_0^{escap} and γ_0^{fin};

2. *Individual trajectories* γ_i*:* Space and time warps

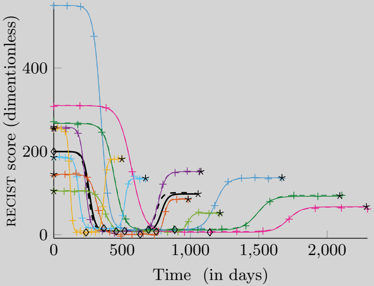
$$\psi_i^\ell : t \mapsto e^{\xi_i^\ell} (t - t_0) + t_0 + \tau_i \,,$$

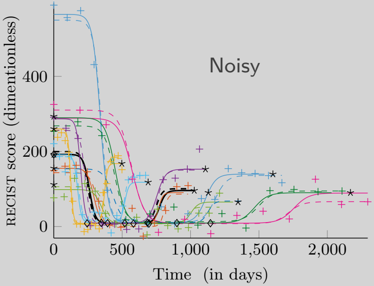
$$\phi_i^\ell : x \mapsto e^{\rho_i^\ell} (x - \gamma_0(t_R)) + \gamma_0(t_R) + \delta_i \,;$$

3. *Latent variables:* $z_{\text{pop}} = \left(\gamma_0^{\text{init}}, \gamma_0^{\text{escap}}, \gamma_0^{\text{fin}}, t_R, t_1\right)$ and $z_i = (\xi_i^1, \xi_i^2, \tau_i, \rho_i^1, \rho_i^2, \delta_i)$;

4. *Parameters:* $\theta = \left(\overline{\gamma_0^{\text{init}}}, \overline{\gamma_0^{\text{escap}}}, \overline{\gamma_0^{\text{fin}}}, \overline{t_R}, \overline{t_1}, \Sigma, \sigma\right)$.

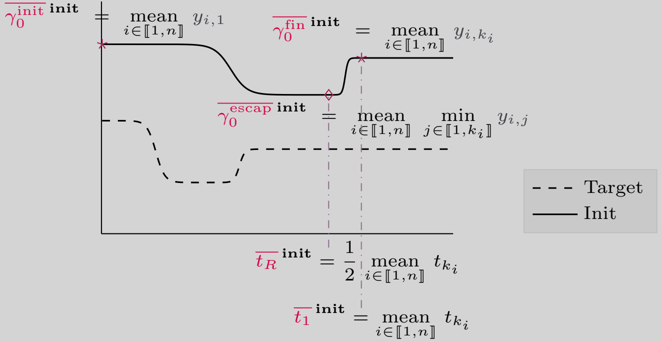
Qualitative Performance of the Estimation ($n = 250$)





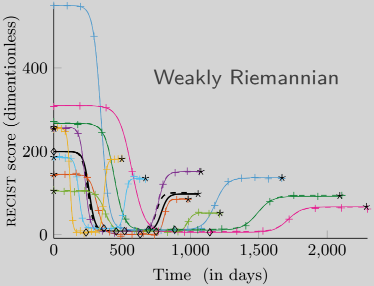
- Estimation performed through the SAEM algorithm ;
- Initialization: $\langle$ true γ_0 $\rangle$ *vs* $\left\langle \begin{array}{c}\text{mean curve in} \\ \text{Euclidean setting}\end{array} \right\rangle$

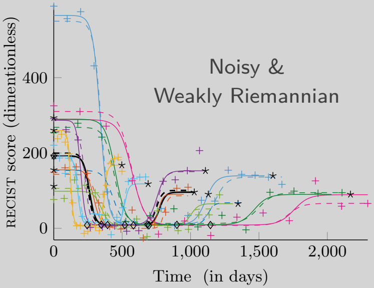
  $\longleftrightarrow$ Weakly or Strongly Riemannian.

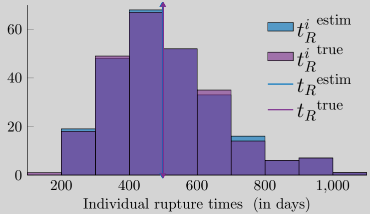


Qualitative Performance of the Estimation ($n = 250$)

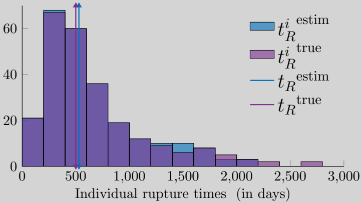
Juliette Chevallier
September, 26th 2019



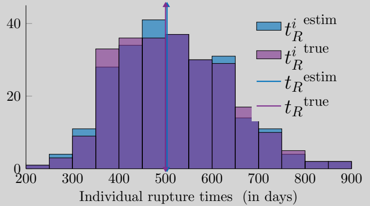




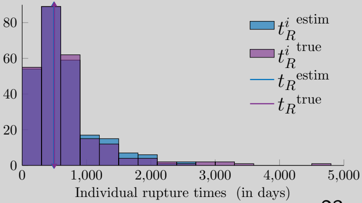
Weakly Riemannian



Strongly Riemannian



Noisy & Weakly Riemannian



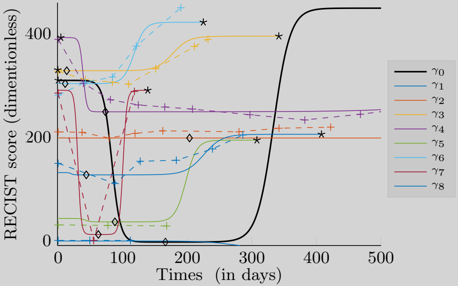
Noisy & Strongly Riemannian

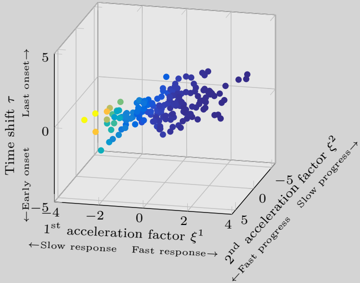
Quantitative Performance of the Estimation ($n = 250$)

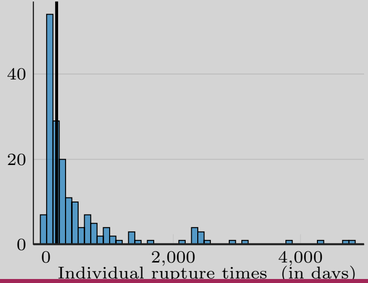
Dataset	Sample size n	$\overline{\gamma_0^{\text{init}}}$	$\overline{\gamma_0^{\text{escap}}}$	$\overline{\gamma_0^{\text{fin}}}$	$\overline{t_R}$	$\overline{t_1}$
Quasi Euclidean	50	6.03 (0.32)	10.25 (0.50)	3.69 (0.25)	1.95 (0.13)	2.43 (0.18)
	100	2.19 (0.17)	3.28 (0.22)	2.07 (0.18)	1.69 (0.11)	1.86 (0.17)
	250	1.30 (0.10)	1.96 (0.13)	1.53 (0.08)	0.78 (0.06)	1.67 (0.09)
Noisy & Weakly Riemannian	50	3.74 (0.26)	25.73 (1.64)	6.84 (0.40)	3.32 (0.26)	3.73 (0.26)
	100	2.35 (0.15)	12.20 (0.64)	1.35 (0.09)	2.98 (0.22)	2.29 (0.18)
	250	1.70 (0.12)	3.94 (0.29)	1.33 (0.09)	1.36 (0.10)	1.51 (0.10)
Strongly Riemannian	50	71.13 (1.33)	100.24 (8.09)	90.73 (2.54)	7.78 (0.56)	46.39 (1.32)
	100	58.73 (0.98)	58.88 (3.00)	84.99 (1.42)	8.13 (0.57)	42.06 (1.04)
	250	67.49 (0.47)	23.12 (1.54)	57.82 (0.74)	6.01 (0.33)	38.09 (0.36)
Noisy & Strongly Riemannian	50	41.61 (1.26)	29.86 (2.53)	46.38 (1.60)	9.04 (0.58)	29.90 (0.58)
	100	60.39 (0.81)	28.43 (2.06)	58.35 (1.07)	8.11 (0.54)	29.75 (0.50)
	250	55.89 (0.74)	15.56 (0.98)	59.90 (0.58)	3.26 (0.25)	39.28 (0.43)

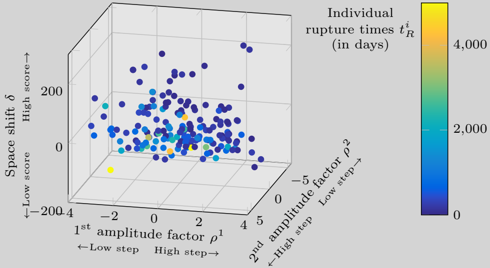
RECIST Score (After 600 Iterations)

Juliette Chevallier





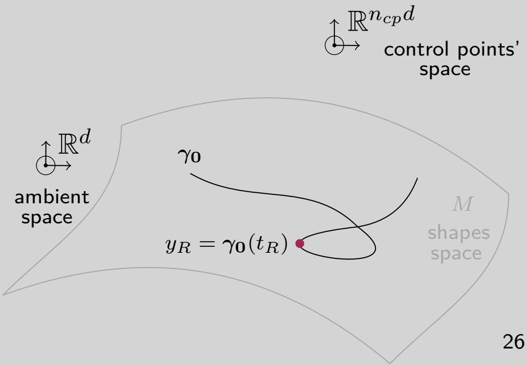




Application to Chemotherapy Monitoring

Build on the work of [Bône et al., 2018] ; Applicable either for currents, varifolds, normal cycles.

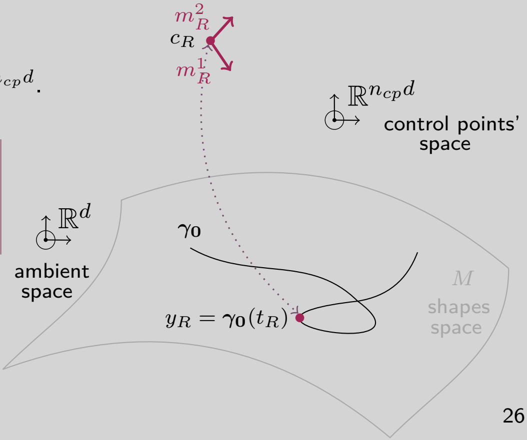
1. *Representative path γ_0:* Given
   - Rupture shape $y_R \in M \subset \mathbb{R}^d$ and time t_R ,



Build on the work of [Bône et al., 2018] ; Applicable either for currents, varifolds, normal cycles.

1. *Representative path γ_0:* Given
   - Rupture shape $y_R \in M \subset \mathbb{R}^d$ and time t_R ,
   - Set of n_{cp} rupture control points $c_R \in \mathbb{R}^{n_{cp}d}$,
   - Backward and forward momenta $m_R^1, m_R^2 \in \mathbb{R}^{n_{cp}d}$.

$$\gamma_0 :\ t \ \mapsto \ \mathcal{E}xp_{c_R,t_R,-t}(m_R^1) \circ y_R \ \mathbb{1}_{]-\infty,t_R]}(t) + \mathcal{E}xp_{c_R,t_R,t}(m_R^2) \circ y_R \ \mathbb{1}_{[t_R,+\infty[}(t)\,,$$



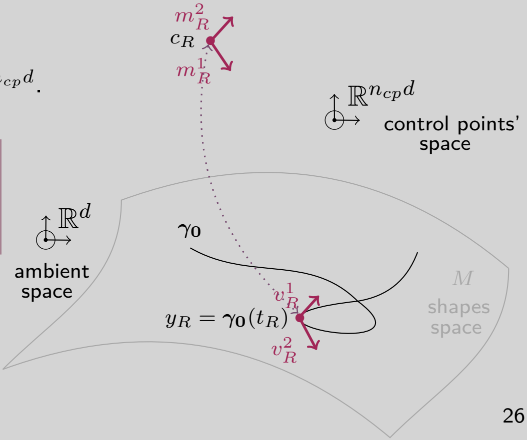
Build on the work of [Bône et al., 2018] ; Applicable either for currents, varifolds, normal cycles.

1. *Representative path* γ_0*:* Given
   - Rupture shape $y_R \in M \subset \mathbb{R}^d$ and time t_R ,
   - Set of n_{cp} rupture control points $c_R \in \mathbb{R}^{n_{cp}d}$,
   - Backward and forward momenta $m_R^1, m_R^2 \in \mathbb{R}^{n_{cp}d}$.

$$\gamma_0 : t \mapsto \mathcal{E}xp_{c_R,t_R,-t}(m_R^1) \circ y_R \; \mathbb{1}_{]-\infty,t_R]}(t) + \mathcal{E}xp_{c_R,t_R,t}(m_R^2) \circ y_R \; \mathbb{1}_{[t_R,+\infty[}(t) \,,$$

Velocity vectors:

$v_R^1 = \langle \; c_R \mid m_R^1 \; \rangle$ and $v_R^2 = \langle \; c_R \mid m_R^2 \; \rangle$.

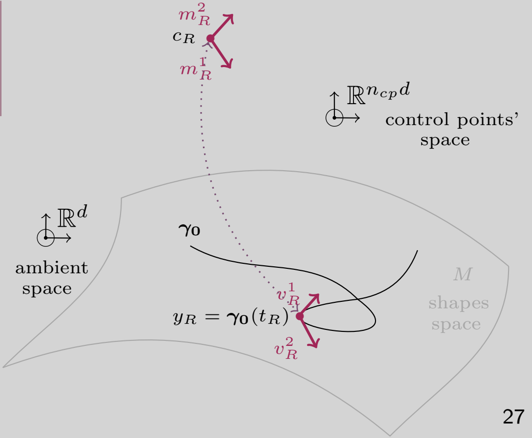


Build on the work of [Bône et al., 2018] ; Applicable either for currents, varifolds, normal cycles.

1. *Representative path γ_0:* Given y_R, t_R , c_R, m_R^1 and m_R^2.

$$\gamma_0 : t \mapsto \mathcal{E}xp_{c_R,t_R,-t}(m_R^1) \circ y_R \; \mathbb{1}_{]-\infty,t_R]}(t) + \mathcal{E}xp_{c_R,t_R,t}(m_R^2) \circ y_R \; \mathbb{1}_{[t_R,+\infty[}(t) \,,$$

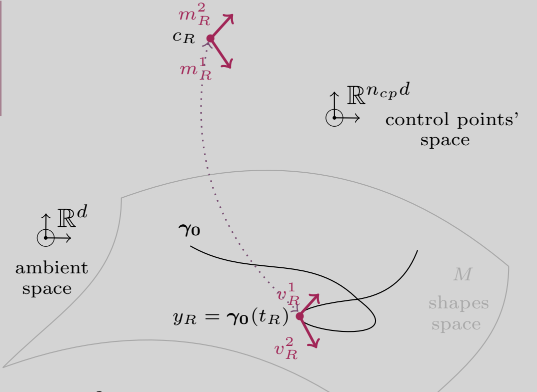
2. *Individual trajectories γ_i:*
   - Time warp: $\psi_i^1 : t \mapsto e^{\xi_i^1}(t - t_R - \tau_i) + t_R$ and $\psi_i^2 : t \mapsto e^{\xi_i^2}(t - t_R - \tau_i) + t_R$,



Build on the work of [Bône et al., 2018] ; Applicable either for currents, varifolds, normal cycles.

1. *Representative path γ_0:* Given y_R, t_R , c_R, m_R^1 and m_R^2.

$$\gamma_0 : t \mapsto \mathcal{E}xp_{c_R,t_R,-t}(m_R^1) \circ y_R \ \mathbb{1}_{]-\infty,t_R]}(t) + \mathcal{E}xp_{c_R,t_R,t}(m_R^2) \circ y_R \ \mathbb{1}_{[t_R,+\infty[}(t) \,,$$



2. *Individual trajectories γ_i:*
   - Time warp: $\psi_i^1 : t \mapsto e^{\xi_i^1}(t - t_R - \tau_i) + t_R$ and $\psi_i^2 : t \mapsto e^{\xi_i^2}(t - t_R - \tau_i) + t_R$,
   - Space warp: $\mathcal{E}xp$-parallelism of γ_0

$$\eta_w : t \mapsto \mathcal{E}xp_{c(t),0,1}\left(P_t(w)\right)$$

where $c(t) = \mathcal{E}xp_{c_R,t_R,-t}(m_R^1) \circ c_R \ \mathbb{1}_{]-\infty,t_R]}(t) + \mathcal{E}xp_{c_R,t_R,t}(m_R^2) \circ c_R \ \mathbb{1}_{[t_R,+\infty[}(t)$

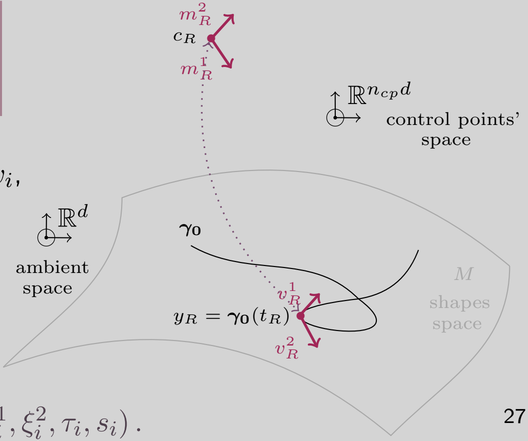
Build on the work of [Bône et al., 2018] ; Applicable either for currents, varifolds, normal cycles.

1. *Representative path γ_0:* Given y_R, t_R , c_R, m_R^1 and m_R^2.

$$\gamma_0 : t \mapsto \mathcal{E}xp_{c_R,t_R,-t}(m_R^1) \circ y_R \, \mathbb{1}_{]-\infty,t_R]}(t) + \mathcal{E}xp_{c_R,t_R,t}(m_R^2) \circ y_R \, \mathbb{1}_{[t_R,+\infty[}(t) \,,$$

2. *Individual trajectories γ_i:* Given ξ_i^1, ξ_i^2, τ_i and w_i,

$$\gamma_i : t \mapsto \eta_{w_i}(\psi_i^1(t)) \circ y_R \, \mathbb{1}_{]-\infty,t_R^i]}(t) + \eta_{w_i}(\psi_i^2(t)) \circ y_R \, \mathbb{1}_{[t_R^i,+\infty[}(t)$$

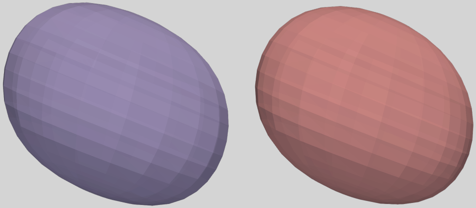


3. Independent Component Analysis: $w_i = As_i$.

$$z_{\text{pop}} = (y_R, c_R, m_R^1, m_R^2, t_R, A) \quad \text{and} \quad z_i = (\xi_i^1, \xi_i^2, \tau_i, s_i) \,.$$

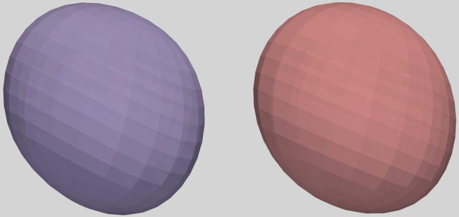
Performance of the Estimation

$\overline{y_R}$	$\overline{t_R}$	Template reconstruction	t_R^i	Individuals reconstruction
1.30	0.01	9.72	0.31 (0.41)	7.94 (5.91)



(a) Simulated tumor (b) Reconstructed tumor

(a) Template



(a) Simulated tumor (b) Reconstructed tumor

(b) One subject

Sample vs Reconstruction – 3D Shapes

A New Class of EM Algorithms

The Expectation-Maximization algorithm

Let $\mathcal{Y} \subset \mathbb{R}^{n_y}$, $\mathcal{Z} \subset \mathbb{R}^{n_z}$ and $\Theta \subset \mathbb{R}^{n_\theta}$.

MLE: Given $y_1^n = (y_1, \ldots, y_n) \in \mathcal{Y}^n$,

$$\widehat{\theta}_n^{MLE} \in \operatorname*{argmax}_{\theta \in \Theta} q\big(y_1^n; \theta\big)$$

E-step: Conditional expected log-likelihood

$$Q(\theta|\theta_k) = \int_{\mathcal{Z}} \log q(y, z; \theta)\, q(z|y; \theta_k)\, \mathrm{d}\mu(z)\,;$$

M-step: Maximize $Q(\,\cdot\,|\theta_k)$ in Θ:

$$\theta_{k+1} \in \operatorname*{argmax}_{\theta \in \Theta} Q(\theta|\theta_k)\,.$$

The Expectation-Maximization algorithm

Let $\mathcal{Y} \subset \mathbb{R}^{n_y}$, $\mathcal{Z} \subset \mathbb{R}^{n_z}$ and $\Theta \subset \mathbb{R}^{n_\theta}$.

MLE: Given $y_1^n = (y_1, \dots, y_n) \in \mathcal{Y}^n$,

$$\widehat{\theta}_n^{MLE} \in \operatorname*{argmax}_{\theta \in \Theta} q\big(y_1^n; \theta\big)$$

E-step: Conditional expected log-likelihood

$$Q(\theta|\theta_k) = \int_{\mathcal{Z}} \log q(y, z; \theta)\, q(z|y; \theta_k)\, \mathrm{d}\mu(z)\,;$$

M-step: Maximize $Q(\,\cdot\,|\theta_k)$ in Θ:

$$\theta_{k+1} \in \operatorname*{argmax}_{\theta \in \Theta} Q(\theta|\theta_k)\,.$$

Convergence for curved exponential families

(M1) $\exists\, S : \mathbb{R}^{n_y} \times \mathbb{R}^{n_z} \to \mathcal{S} \subset \mathbb{R}^{n_s}$ Borel function

$\mathcal{C}onv(S) \subset \mathcal{S}$, $\int_{\mathcal{Z}} \|S(y,z)\|\, q(z|y;\theta)\, \mathrm{d}\mu(z) \;<\; +\infty$

$$q(y, z; \theta) = \exp\big(-\psi(\theta) + \langle\, S(y,z) \mid \phi(\theta)\, \rangle\big)$$

Juliette Chevallier

The Expectation-Maximization algorithm

Let $\mathcal{Y} \subset \mathbb{R}^{n_y}$, $\mathcal{Z} \subset \mathbb{R}^{n_z}$ and $\Theta \subset \mathbb{R}^{n_\theta}$.

MLE: Given $y_1^n = (y_1, \ldots, y_n) \in \mathcal{Y}^n$,

$$\widehat{\theta}_n^{MLE} \in \operatorname*{argmax}_{\theta \in \Theta} q\big(y_1^n; \theta\big)$$

E-step: Conditional expected log-likelihood

$$Q(\theta|\theta_k) = \int_{\mathcal{Z}} \log q(y, z; \theta)\, q(z|y; \theta_k)\, \mathrm{d}\mu(z)\,;$$

M-step: Maximize $Q(\,\cdot\,|\theta_k)$ in Θ:

$$\theta_{k+1} \in \operatorname*{argmax}_{\theta \in \Theta} Q(\theta|\theta_k)\,.$$

Convergence for curved exponential families

(M1) $\exists\, S : \mathbb{R}^{n_y} \times \mathbb{R}^{n_z} \to \mathcal{S} \subset \mathbb{R}^{n_s}$ Borel function $\mathcal{C}onv(S) \subset \mathcal{S}$, $\int_{\mathcal{Z}} \|S(y,z)\|\, q(z|y;\theta)\, \mathrm{d}\mu(z) < +\infty$

$$q(y, z; \theta) = \exp\big(-\psi(\theta) + \langle\, S(y,z) \mid \phi(\theta)\, \rangle\big)$$

(M2) $\psi \in \mathcal{C}^2(\Theta, \mathbb{R})$ and $\phi \in \mathcal{C}^2(\Theta, \mathcal{S})$;

(M3) $\theta \mapsto \int_{\mathcal{Z}} S(y,z) q(z|y;\theta)\, \mathrm{d}\mu(z) \in \mathcal{C}^1(\Theta, \mathcal{S})$;

(M4) $\ell : \theta \mapsto \int_{\mathcal{Z}} q(y,z;\theta)\, \mathrm{d}\mu(z) \in \mathcal{C}^1(\Theta, \mathbb{R})$ and

$$\partial_\theta \int_{\mathcal{Z}} q(y,z;\theta)\, \mathrm{d}\mu(z) = \int_{\mathcal{Z}} \partial_\theta\, q(y,z;\theta)\, \mathrm{d}\mu(z)\,;$$

(M5) $\exists\, \hat{\theta} \in \mathcal{C}^1(\theta, \mathcal{S})$ s.t.

$$\psi(\hat{\theta}(s)) + \langle s | \phi(\hat{\theta}(s)) \rangle \geqslant \psi(\theta) + \langle s | \phi(\theta) \rangle\,.$$

Convergence EM – [Delyon et al., 1999]

Assume (M1-5) and that $(\theta_k)_{k\in\mathbb{N}}$ remains in a compact subset. Then, for any initial point,

$$\lim_{k\to\infty} d(\theta_k, \mathcal{L}) = 0\,,$$

where $\mathcal{L} = \{\,\theta \in \Theta \,|\, \partial_\theta \ell(\theta) = 0\,\}$.

E-step: Conditional expected log-likelihood

$$Q(\theta|\theta_k) \;=\; \int_{\mathcal{Z}} \log q(y,z;\theta)\, q(z|y;\theta_k)\; \mathrm{d}\mu(z)\,;$$

M-step: Maximize $Q(\,\cdot\,|\theta_k)$ in Θ:

$$\theta_{k+1} \;\in\; \operatorname*{argmax}_{\theta\in\Theta} Q(\theta|\theta_k)\,.$$

Convergence for curved exponential families

(M1) $\exists\, S : \mathbb{R}^{n_y} \times \mathbb{R}^{n_z} \to \mathcal{S} \subset \mathbb{R}^{n_s}$ Borel function $\mathcal{C}onv(S) \subset \mathcal{S}$, $\int_{\mathcal{Z}} \|S(y,z)\|\, q(z|y;\theta)\, \mathrm{d}\mu(z) \;<\; +\infty$

$$q(y,z;\theta) \;=\; \exp\big(-\psi(\theta) + \langle\; S(y,z) \;|\; \phi(\theta)\;\rangle\big)$$

(M2) $\psi \in \mathcal{C}^2(\Theta,\mathbb{R})$ and $\phi \in \mathcal{C}^2(\Theta,\mathcal{S})$;

(M3) $\theta \;\mapsto\; \int_{\mathcal{Z}} S(y,z) q(z|y;\theta)\, \mathrm{d}\mu(z) \in \mathcal{C}^1(\Theta,\mathcal{S})$;

(M4) $\ell\colon \theta \;\mapsto\; \int_{\mathcal{Z}} q(y,z;\theta)\, \mathrm{d}\mu(z) \in \mathcal{C}^1(\Theta,\mathbb{R})$ and

$$\partial_\theta \int_{\mathcal{Z}} q(y,z;\theta)\, \mathrm{d}\mu(z) \;=\; \int_{\mathcal{Z}} \partial_\theta\, q(y,z;\theta)\, \mathrm{d}\mu(z)\,;$$

(M5) $\exists\, \hat{\theta} \in \mathcal{C}^1(\theta,\mathcal{S})$ s.t.
$\psi(\hat{\theta}(s)) + \langle s|\phi(\hat{\theta}(s))\rangle \;\geqslant\; \psi(\theta) + \langle s|\phi(\theta)\rangle\,.$

Speeding-up ⟵ EM

Speeding-up ← EM

M-step

GEM – Generalized EM

E-step: Compute

$$Q(\theta|\theta_k) = \mathbb{E}\left[\log q(Z|y,\theta_k)\right] ;$$

M-step: Find $\theta_{k+1} \in \Theta$ s.t.

$$Q(\theta_{k+1}|\theta_k) \geqslant Q(\theta_k|\theta_k),$$

[Delyon et al., 1999]

See also *Gradient EM* [Lange, 1995]

Speeding-up ← EM → (E-step) *SAEM*; EM → (M-step) GEM; EM → (E-step) SEM; EM → (E-step) MCEM

GEM – Generalized EM	*SEM – Stochastic EM*	*MCEM – Monte-Carlo EM*
E-step: Compute $Q(\theta\|\theta_k) = \mathbb{E}\left[\log q(Z\|y,\theta_k)\right]$;	**S-step:** Draw an unobserved sample z_k from $q(\,\cdot\,\|y;\theta_k)$;	**S-step:** Draw m samples $z_k^j \sim q(\,\cdot\,\|y;\theta_k)$;
M-step: Find $\theta_{k+1} \in \Theta$ s.t. $Q(\theta_{k+1}\|\theta_k) \geqslant Q(\theta_k\|\theta_k)$,	**M-step:** Maximize Q_{k+1}: $\theta_{k+1} \in \operatorname{argmax}_{\theta\in\Theta} Q_{k+1}(\theta)$.	**E-step:** Monte-Carlo estim. $Q_k(\theta) = \frac{1}{m}\sum_{j=1}^{m} \log q(y, z_k^j; \theta)$;
		M-step: Maximize Q_{k+1}.
[Delyon et al., 1999] See also *Gradient EM* [Lange, 1995]	[Celeux and Diebolt, 1985]	[Wei and Tanner, 1990]

The SAEM algorithm

- *Idea:* Replace the E-step by a *stochastic approximation*,
- Sequence of positive step-size $(\gamma_k)_{k \in \mathbb{N}}$.

S-step: Draw $z_k \sim q(\,\cdot\,|y;\theta_k)\,;$

SA-step: Update $Q_k(\theta)$ as

$$Q_{k+1}(\theta) = Q_k(\theta) + \gamma_k\big(\,\log q(y, z_k; \theta) - Q_k(\theta)\,\big)\,;$$

M-step: Maximize Q_{k+1} in Θ:

$$\theta_{k+1} \in \operatorname*{argmax}_{\theta \in \Theta} Q_{k+1}(\theta)\,.$$

The SAEM algorithm

- *Idea:* Replace the E-step by a *stochastic approximation*,
- Sequence of positive step-size $(\gamma_k)_{k\in\mathbb{N}}$.

S-step: Draw $z_k \sim q(\,\cdot\,|y;\theta_k)\,;$

SA-step: Update $Q_k(\theta)$ as

$$\begin{aligned} Q_{k+1}(\theta) \;&=\; Q_k(\theta) \\ &\quad + \gamma_k\big(\,\log q(y,z_k;\theta) - Q_k(\theta)\,\big)\,; \end{aligned}$$

M-step: Maximize Q_{k+1} in Θ:

$$\theta_{k+1} \;\in\; \operatorname*{argmax}_{\theta\in\Theta} Q_{k+1}(\theta)\,.$$

Convergence for curved exponential families

(SAEM1) $\gamma_k \in [0,1]$, $\displaystyle\sum_{k=1}^{\infty} \gamma_k = \infty$ and $\displaystyle\sum_{k=1}^{\infty} \gamma_k^2 < \infty\,;$

(SAEM2) $\psi \in \mathcal{C}^{n_s}(\Theta,\mathbb{R})$ and $\phi \in \mathcal{C}^{n_s}(\Theta,\mathcal{S})\,;$

(SAEM3) $\mathbb{E}\big[\phi(Z_{k+1})\big|\mathcal{F}_k\big] = \displaystyle\int_{\mathcal{Z}} \phi(z)q(z|y;\theta_k)\,\mathrm{d}\mu(z)\,;$

(SAEM4) $\displaystyle\int_{\mathcal{Z}} \|S(y,z)\|^2\, q(y,z;\theta)\,\mathrm{d}\mu(z) \;<\; +\infty\,.$

Cvgce SAEM – [Delyon et al., 1999]

Assume (M1-5), (SAEM1-4) and that $(s_k)_{k\in\mathbb{N}}$ remains in a compact subset. Then, for any initial point,

$$\lim_{k\to\infty} d(\theta_k, \mathcal{L}) = 0\,,$$

where $\mathcal{L} = \{\,\theta\in\Theta \,|\, \partial_\theta \ell(\theta) = 0\,\}$.

S-step: Draw $z_k \sim q(\,\cdot\,|y;\theta_k)$;

SA-step: Update $s_k(\theta)$ as

$$s_{k+1}(\theta) \;=\; s_k(\theta) + \gamma_k\big(S(y,z_k) - s_k(\theta)\big)\,;$$

M-step: Maximize Q_{k+1} in Θ:

$$\theta_{k+1} \;\in\; \underset{\theta\in\Theta}{\operatorname{argmax}}\; Q_{k+1}(\theta)\,.$$

Convergence for curved exponential families

(SAEM1) $\gamma_k \in [0,1]$, $\sum_{k=1}^{\infty}\gamma_k = \infty$ and $\sum_{k=1}^{\infty}\gamma_k^2 < \infty$;

(SAEM2) $\psi \in \mathcal{C}^{n_s}(\Theta,\mathbb{R})$ and $\phi \in \mathcal{C}^{n_s}(\Theta,\mathcal{S})$;

(SAEM3) $\mathbb{E}\big[\phi(Z_{k+1})\big|\mathcal{F}_k\big] = \int_{\mathcal{Z}} \phi(z) q(z|y;\theta_k)\,\mathrm{d}\mu(z)$;

(SAEM4) $\int_{\mathcal{Z}} \|S(y,z)\|^2\, q(y,z;\theta)\,\mathrm{d}\mu(z) \;<\; +\infty$.

MCMC-SAEM: Monte-Carlo Markov chain procedure in the S-step

[Kuhn and Lavielle, 2004]

[Allassonnière et al., 2010]

A New Class of EM Algorithms

The SAEM algorithm

- Sequence of positive step-size $(\gamma_k)_{k\in\mathbb{N}}$.

S-step: Sample z_k under the *posterior* density $q(\cdot|y;\theta)$;

SA-step: Update $s_k(\theta)$ as

$$s_{k+1}(\theta) = s_k(\theta) + \gamma_k\big(S(y,z_k) - s_k(\theta)\big) \,;$$

M-step: Maximize Q_{k+1} in Θ:

$$\theta_{k+1} \in \operatorname*{argmax}_{\theta\in\Theta} Q_{k+1}(\theta) \,.$$

The SAEM algorithm

- Sequence of positive step-size $(\gamma_k)_{k\in\mathbb{N}}$.

S-step: Sample z_k under the *posterior* density $q(\cdot|y;\theta)$;

SA-step: Update $s_k(\theta)$ as

$$s_{k+1}(\theta) = s_k(\theta) + \gamma_k\big(S(y, z_k) - s_k(\theta)\big) ;$$

M-step: Maximize Q_{k+1} in Θ:

$$\theta_{k+1} \in \operatorname*{argmax}_{\theta\in\Theta} Q_{k+1}(\theta) .$$

The approximated-SAEM algorithm

- Sequence of approximated-distributions on $\mathcal{Z} \times \Theta$: $\tilde{q} = (\tilde{q}_k)_{k\in\mathbb{N}}$.

S-step: Sample $\tilde{z}_k$ under the *approximated* density $\tilde{q}_k(\cdot;\theta_k)$

SA-step: Update $s_k(\theta)$ as

$$s_{k+1}(\theta) = s_k(\theta) + \gamma_k\big(S(y, \tilde{z}_k) - s_k(\theta)\big) ;$$

M-step: Maximize Q_{k+1} in Θ:

$$\theta_{k+1} \in \operatorname*{argmax}_{\theta\in\Theta} Q_{k+1}(\theta) .$$

Convergence for curved exponential families

We adapt (M1), (SAEM3), (SAEM4) to have regularity against q *and* $\tilde{q}$.

Approx: $\forall y \in \mathcal{Y}$, $\forall \mathcal{K} \subset \Theta$ compact,

$$\lim_{k\to\infty} \left\{ \sup_{\theta\in\mathcal{K}} \int_{\mathcal{Z}} S(y,z)\big(\tilde{q}_k(z;\theta) - q(z|y;\theta)\big)\, \mathrm{d}\mu(z) \right\} = 0\,.$$

The approximated-SAEM algorithm

- Sequence of positive step-size $(\gamma_k)_{k\in\mathbb{N}}$.
- Sequence of approximated-distributions on $\mathcal{Z} \times \Theta$: $\tilde{q} = (\tilde{q}_k)_{k\in\mathbb{N}}$.

S-step: Sample $\tilde{z}_k$ under the *approximated* density $\tilde{q}_k(\cdot;\theta_k)$

SA-step: Update $s_k(\theta)$ as

$$s_{k+1}(\theta) = s_k(\theta) + \gamma_k\big(S(y,\tilde{z}_k) - s_k(\theta)\big)\,;$$

M-step: Maximize Q_{k+1} in Θ:

$$\theta_{k+1} \in \operatorname*{argmax}_{\theta\in\Theta} Q_{k+1}(\theta)\,.$$

Convergence for curved exponential families

We adapt (M1), (SAEM3), (SAEM4) to have regularity against q *and* $\tilde{q}$.

Approx: $\forall y \in \mathcal{Y}$, $\forall \mathcal{K} \subset \Theta$ compact,

$$\lim_{k\to\infty} \left\{ \sup_{\theta\in\mathcal{K}} \int_{\mathcal{Z}} S(y,z)\big(\tilde{q}_k(z;\theta) - q(z|y;\theta)\big)\, \mathrm{d}\mu(z) \right\} = 0\,.$$

Convergence *approximated*-SAEM

Assume (M⋆1-5), (SAEM⋆1-4) and the compactness condition. Then, for any initial point,

$$\lim_{k\to\infty} d(\theta_k, \mathcal{L}) = 0\,,$$

The approximated-SAEM algorithm

- Sequence of positive step-size $(\gamma_k)_{k\in\mathbb{N}}$.
- Sequence of approximated-distributions on $\mathcal{Z}\times\Theta$: $\tilde{q} = (\tilde{q}_k)_{k\in\mathbb{N}}$.

S-step: Sample $\tilde{z}_k$ under the *approximated* density $\tilde{q}_k(\cdot;\theta_k)$

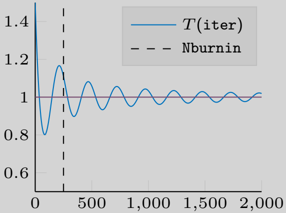
SA-step: Update $s_k(\theta)$ as

$$s_{k+1}(\theta) = s_k(\theta) + \gamma_k\big(S(y,\tilde{z}_k) - s_k(\theta)\big)\,;$$

M-step: Maximize Q_{k+1} in Θ:

$$\theta_{k+1} \in \operatorname*{argmax}_{\theta\in\Theta} Q_{k+1}(\theta)\,.$$

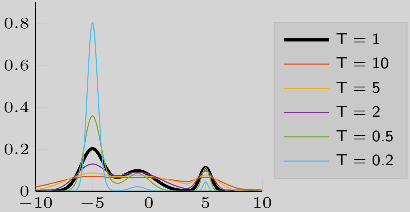
A New Class of EM Algorithms



(a) Tempering scheme

Temperatures: $T = (T_k)_{k \in \mathbb{N}}$ sequence of positive numbers s.t. $\lim_{k \to \infty} T_k = 1$. Let $c_\theta(T_k)$ is a scaling constant.

$$\tilde{q}_k(z;\theta) = \frac{1}{c_\theta(T_k)}\, q(z|y;\theta)^{1/T_k}$$

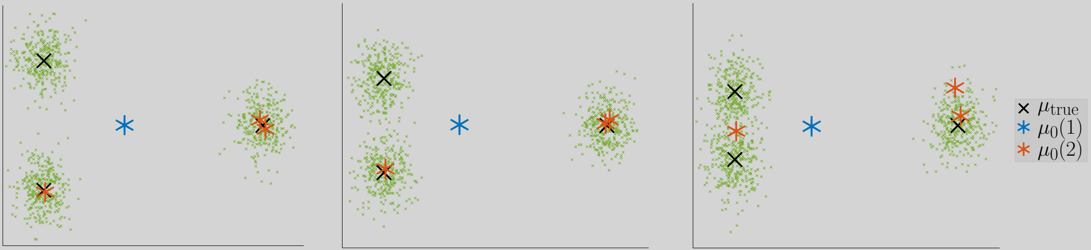


(b) Tempering distributions

In practice: $a \in [0,1[,\ b, c \in \mathbb{R}$.

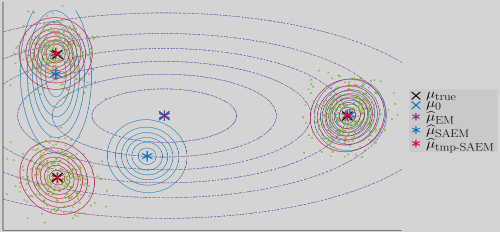
$$T_k = 1 + a^\kappa + b\,\frac{\sin(\kappa)}{\kappa}$$

where $\kappa = \dfrac{k + c \times r}{r}$.

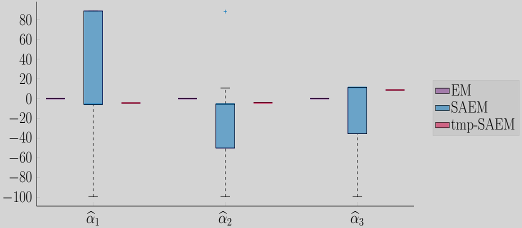


(a) Dataset I **(b)** Dataset II **(c)** Dataset III

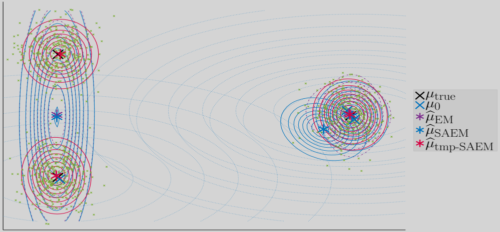
Performance of the Estimation for the Dataset I



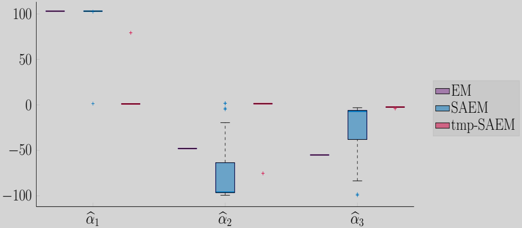
(a) Init. 1 – Qualitative performance



(b) Init. 1 – Relative error for α (in %)

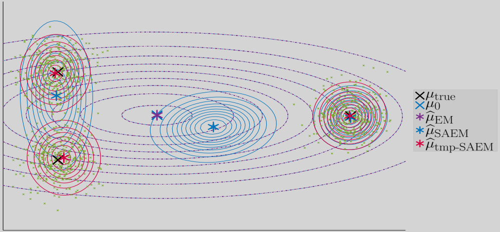


(c) Init. 2 – Qualitative performance

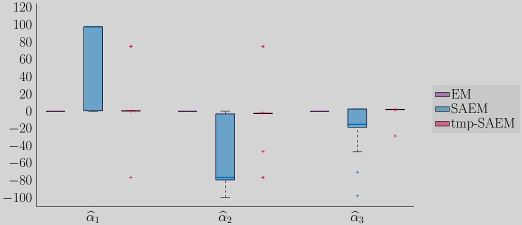


(d) Init. 2 – Relative error for α (in %)

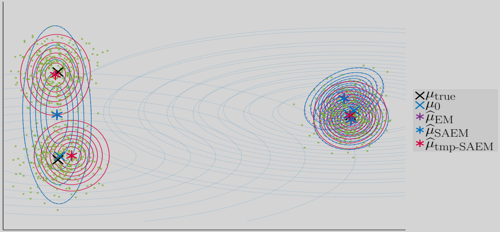
Performance of the Estimation for the Dataset II



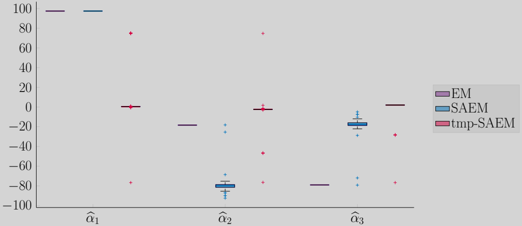
(a) Init. 1 – Qualitative performance



(b) Init. 1 – Relative error for α (in %)

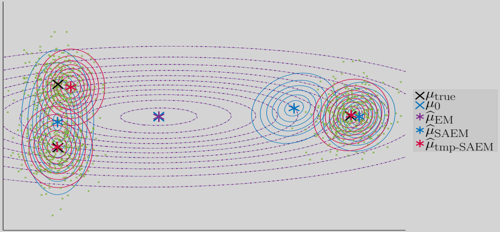


(c) Init. 2 – Qualitative performance

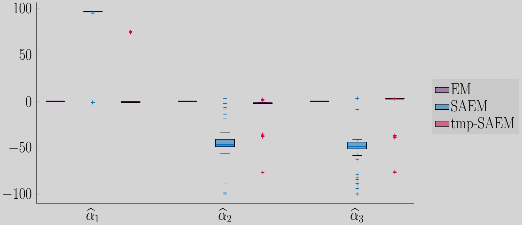


(d) Init. 2 – Relative error for α (in %)

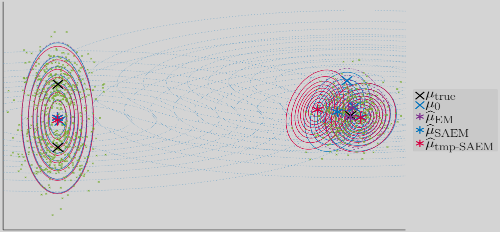
Performance of the Estimation for the Dataset III



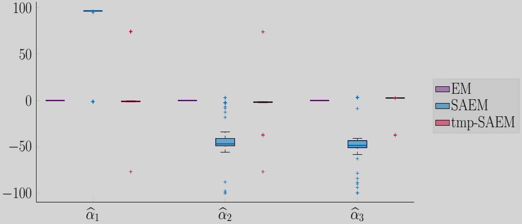
(a) Init. 1 – Qualitative performance



(b) Init. 1 – Relative error for α (in %)



(c) Init. 2 – Qualitative performance



(d) Init. 2 – Relative error for α (in %)

Contributions

1. Generic approach to study non-monotonous dynamics on Riemannian manifolds:
   - Nonlinear mixed effects model,
   - Spatio-temporal deformation of a group-representative trajectory;
2. Demonstration of the *existence* and the *consistency* of the MAP estimator for this generic model;
3. Application to chemotherapy monitoring:
   - through RECIST score $\longleftrightarrow$ Piecewise-logistic curve model,
   - through Anatomical Shapes $\longleftrightarrow$ Piecewise-geodesic shape model;
4. New stochastic approximation version of the EM algorithm and demonstration of the convergence toward local maxima;
5. Building on simulated annealing techniques, an instantiation of this general procedure to favor convergence toward global maxima.

Merci de votre attention

References i

Allassonnière, S., Kuhn, E., and Trouvé, A. (2010). **Construction of Bayesian deformable models via a stochastic approximation algorithm: A convergence study.** Bernoulli, 16(3):641–678.

Bates, D. M. and Watts, D. G. (1988). **Nonlinear regression analysis and its applications.** Wiley Series in Probability and Mathematical Statistics. Wiley.

Beg, M. F., Miller, M. I., Trouvé, A., and Younes, L. (2005). **Computing large deformation metric mappings via geodesic flows of diffeomorphisms.** International journal of computer vision, 61(2):139–157.

Bône, A., Colliot, O., and Durrleman, S. (2018). **Learning distributions of shape trajectories from longitudinal datasets: A hierarchical model on a manifold of diffeomorphisms.** In Computer Vision and Pattern Recognition, Salt Lake City, United States.

Celeux, G. and Diebolt, J. (1985). **The SEM algorithm: a probabilistic teacher algorithm derived from the EM algorithm for the mixture problem.** Computational statistics quarterly, 2:73–82.

D'Arcy Thompson, W. (1942). **On Growth and Form.** Cambridge university press.

Debavelaere, V., Bône, A., Durrleman, S., and Allassonnière, S. (2019). **Clustering of longitudinal shape data sets using mixture of separate or branching trajectories.** to appear in MICAI 2019, Hal-02103355.

Delor, I., Charoin, J.-E., Gieschke, R., Retout, S., and Jacqmin, P. (2013). **Modeling Alzheimer's disease progression using disease onset time and disease trajectory concepts applied to cdr-sob scores from adni.** CPT: pharmacometrics & systems pharmacology, 2(10):1–10.

Delyon, B., Lavielle, M., and Moulines, E. (1999). **Convergence of a stochastic approximation version of the EM algorithm.** The Annals of Statistics, 27(1):94–128.

Devilliers, L., Allassonnière, S., Trouvé, A., and Pennec, X. (2017). **Template estimation in computational anatomy: Fréchet means top and quotient spaces are not consistent.** SIAM Journal on Imaging Sciences, 10(3):1139–1169.

Dupuis, P., Grenander, U., and Miller, M. I. (1998). **Variational problems on flows of diffeomorphisms for image matching.** Quarterly of applied mathematics, pages 587–600.

Durrleman, S., Pennec, X., Gerig, G., Trouvé, A., and Ayache, N. (2009). **Spatiotemporal atlas estimation for developmental delay detection in longitudinal datasets.** Technical report, INRIA.

Durrleman, S., Pennec, X., Trouvé, A., Braga, J., Gerig, G., and Ayache, N. (2013). **Toward a comprehensive framework for the spatiotemporal statistical analysis of longitudinal shape data.** International journal of computer vision, 103(1):22–59.

Escudier, B., Porta, C., Schmidinger, M., Rioux-Leclercq, N., Bex, A., Khoo, V. S., Gruenvald, V., and Horwich, A. (2016). **Renal cell carcinoma: ESMO clinical practice guidelines for diagnosis, treatment and follow-up.** Annals of Oncology, 27(suppl 5):v58–v68.

Fishbaugh, J., Durrleman, S., Prastawa, M., and Gerig, G. (2017). **Geodesic shape regression with multiple geometries and sparse parameters.** Medical image analysis, 39:1–17.

Fletcher, T. (2011). **Geodesic regression on Riemannian manifolds.** In Proceedings of the Third International Workshop on Mathematical Foundations of Computational Anatomy-Geometrical and Statistical Methods for Modelling Biological Shape Variability, pages 75–86.

Grenander, U. (1993). **General pattern theory-A mathematical study of regular structures.** Clarendon Press.

References iv

Hong, Y., Singh, N., Kwitt, R., and Niethammer, M. (2014). **Time-warped geodesic regression.** In International Conference on Medical Image Computing and Computer-Assisted Intervention, pages 105–112. Springer.

Kim, H. J., Adluru, N., Collins, M. D., Chung, M. K., Bendlin, B. B., Johnson, S. C., Davidson, R. J., and Singh, V. (2014). **Multivariate general linear models (MGLM) on Riemannian manifolds with applications to statistical analysis of diffusion weighted images.** In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 2705–2712.

Koval, I., Schiratti, J.-B., Routier, A., Bacci, M., Colliot, O., Allassonniere, S., and Durrleman, S. (2018). **Spatiotemporal propagation of the cortical atrophy: Population and individual patterns.** Frontiers in Neurology, 9.

Kuhn, E. and Lavielle, M. (2004). **Coupling a stochastic approximation version of EM with an MCMC procedure.** ESAIM: Probability and Statistics, 8:115–131.

Kuhn, E. and Lavielle, M. (2005). **Maximum likelihood estimation in nonlinear mixed effects models.** Computational Statistics & Data Analysis, 49(4):1020–1038.

Laird, N. M. and Ware, J. H. (1982). **Random-effects models for longitudinal data.** Biometrics, 38(4):963–974.

References v

Lange, K. (1995). **A gradient algorithm locally equivalent to the EM algorithm.** Journal of the Royal Statistical Society: Series B, 57(2):425–437.

Muralidharan, P. and Fletcher, P. T. (2012). **Sasaki metrics for analysis of longitudinal data on manifolds.** In IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 1027–1034. IEEE.

Schiratti, J.-B., Allassonnière, S., Colliot, O., and Durrleman, S. (2015). **Learning spatiotemporal trajectories from manifold-valued longitudinal data.** In Neural Information Processing Systems, number 28 in Advances in Neural Information Processing Systems, Montréal, Canada.

Sheiner, L. B. and Beal, S. L. (1980). **Evaluation of methods for estimating population pharmacokinetic parameters. i. michaelis-menten model: routine clinical pharmacokinetic data.** Journal of pharmacokinetics and biopharmaceutics, 8(6):553–571.

Singh, N., Hinkle, J., Joshi, S., and Fletcher, P. T. (2013). **A hierarchical geodesic model for diffeomorphic longitudinal shape analysis.** In International Conference on Information Processing in Medical Imaging (IPMI), pages 560–571. Springer.

References vi

Su, J., Kurtek, S., Klassen, E., and Srivastava, A. (2014). **Statistical analysis of trajectories on Riemannian manifolds: bird migration, hurricane tracking and video surveillance.** The Annals of Applied Statistics, 8(1):530–552.

Trouvé, A. and Vialard, F.-X. (2012). **Shape splines and stochastic shape evolutions: A second order point of view.** Quarterly of Applied Mathematics, pages 219–251.

van der Vaart, A. W. (2000). **Asymptotic Statistics.** Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press.

Wei, G. C. and Tanner, M. A. (1990). **A Monte Carlo implementation of the EM algorithm and the poor man's data augmentation algorithms.** Journal of the American statistical Association, 85(411):699–704.

Yang, E., Farnum, M., Lobanov, V., Schultz, T., Raghavan, N., Samtani, M. N., Novak, G., Narayan, V., and DiBernardo, A. (2011). **Quantifying the pathophysiological timeline of Alzheimer's disease.** Journal of Alzheimer's Disease, 26(4):745–753.

Deformable template model: [D'Arcy Thompson, 1942, Grenander, 1993].

Let $M \subset \mathbb{R}^d$, $\boldsymbol{\mathcal{G}}$ a group of deformations and $\boldsymbol{x_0} \in M$ a template shape.

$\longrightarrow$ Transitive action of $\mathcal{G}$ over M $\longrightarrow$ *Shape space* $\boldsymbol{\mathcal{G} \cdot x_0}$.

Deformation metric mapping: [Dupuis et al., 1998, Beg et al., 2005]

$d_{\mathcal{G}}$ a right-invariant metric on $\mathcal{G}$ $\longrightarrow$ d_M pseudo-metric on M:

$$\forall\, x, y \in M, \quad d_M(x, y) \;=\; \inf_{g \in \mathcal{G}} \{d_{\mathcal{G}}(Id, g) \mid g \cdot x = y\}\,.$$

The large deformation diffeomorphic metric mapping or LDDMM framework endows (a restriction of) $\mathcal{G}$ with a tractable metric.

Models in constant improvement:

- [Kim et al., 2014] Generalization of the geodesic hierarchical model.

  $\rightarrow$ ***Riemannian nonlinear mixed effects model.***

$$\begin{cases} y_{i,j} &= \mathcal{E}xp\Big(\mathcal{E}xp\Big(b_i\,;\Gamma_{b,b_i}(v)\big(\alpha_i(t_{i,j}-t_0-\tau_i)\big)\Big)\,;\varepsilon_{i,j}\Big) \\ b_i &= \mathcal{E}xp\,(b\,;u_i) \end{cases},$$

  where $\Gamma_{b,b_i}(v) \in T_{b_i}M =$ parallel transport of $v \in T_bM$ from b to b_i.

  High complexity of the model $\rightarrow$ impossible to estimate exactly the parameters.

- [Koval et al., 2018] Generic spatio-temporal model for the study of networks.

- [Bône et al., 2018] Generic spatio-temporal model for the LDDMM framework.

- [Debavelaere et al., 2019] Generic spatio-temporal model into a mixture model.