Models for warping functions and registration

Karthik Bharath

University of Nottingham

- Work with Sebastian Kurtek (with much help from Anuj, Eric and Huiling Le)
- Supported by NSF (DMS 1613054); NIH (R01-CA214955).

NSF-CBMS workshop on Elastic Functional Data Analysis 2018

This talk. . .

- Statistical models on function spaces for elastic functional data and issues.
- Model-based pairwise registration of curves/functions.
- How to define and sample from distribution on warping functions of $[0,1]$ and $\mathbb{S}^1$?
- One or two applications.

Where to define models for elastic functional data?

- **Original data space**

  Functions: $\mathbb{L}^2([0,1],\mathbb{R})$;
  Curves: $AC(D,\mathbb{R}^n)$, $\mathbb{L}^2(D,\mathbb{R}^n)$ where $D \in \{[0,1],\mathbb{S}^1\}$.

Where to define models for elastic functional data?

- **Original data space**

  Functions: $\mathbb{L}^2([0,1],\mathbb{R})$;
  Curves: $AC(D,\mathbb{R}^n)$, $\mathbb{L}^2(D,\mathbb{R}^n)$ where $D \in \{[0,1],\mathbb{S}^1\}$.

- **Ambient/Top/Pre-shape space**

  Functions: $\mathbb{L}^2([0,1],\mathbb{R})$ (or sometimes $\mathbb{S}^\infty$);
  Curves: $\mathbb{S}^\infty$ (open) or its subset (closed).

Where to define models for elastic functional data?

- **Original data space**

  Functions: $\mathbb{L}^2([0,1],\mathbb{R})$;
  Curves: $AC(D,\mathbb{R}^n)$, $\mathbb{L}^2(D,\mathbb{R}^n)$ where $D \in \{[0,1],\mathbb{S}^1\}$.

- **Ambient/Top/Pre-shape space**

  Functions: $\mathbb{L}^2([0,1],\mathbb{R})$ (or sometimes $\mathbb{S}^\infty$);
  Curves: $\mathbb{S}^\infty$ (open) or its subset (closed).

- **Quotient/Shape space**

  Functions: $\mathbb{L}^2([0,1],\mathbb{R})/\Gamma$ where Γ is group of warp maps of $[0,1]$;
  Curves: $\mathbb{S}^\infty/(\Gamma \times SO(n))$ for open curves.

Where to define models for elastic functional data?

Laws of stochastic processes are natural models:

Where to define models for elastic functional data?

Laws of stochastic processes are natural models:

- **Original data space**. Easy.

  For e.g. Gaussian process on $\mathbb{L}^2(D, \mathbb{R}^n)$; diffusion bridges as solutions of SDEs.

Where to define models for elastic functional data?

Laws of stochastic processes are natural models:

- **Original data space**. Easy.

  For e.g. Gaussian process on $\mathbb{L}^2(D, \mathbb{R}^n)$; diffusion bridges as solutions of SDEs.

- **Ambient/Top/Pre-shape space**. Hard.

  For e.g. Marginalization/projection of Gaussian measure on $\mathbb{L}^2$ to get one on $\mathbb{S}^\infty$ is not straightforward.

Where to define models for elastic functional data?

Laws of stochastic processes are natural models:

- **Original data space**. Easy.

  For e.g. Gaussian process on $\mathbb{L}^2(D, \mathbb{R}^n)$; diffusion bridges as solutions of SDEs.

- **Ambient/Top/Pre-shape space**. Hard.

  For e.g. Marginalization/projection of Gaussian measure on $\mathbb{L}^2$ to get one on $\mathbb{S}^\infty$ is not straightforward.

- **Quotient/Shape space**. Impossibly hard.

  No invariant measure for infinite-dimensional warping group.

Where to define models for elastic functional data?

Laws of stochastic processes are natural models:

- **Original data space**. Easy.

  For e.g. Gaussian process on $\mathbb{L}^2(D, \mathbb{R}^n)$; diffusion bridges as solutions of SDEs.

- **Ambient/Top/Pre-shape space**. Hard.

  For e.g. Marginalization/projection of Gaussian measure on $\mathbb{L}^2$ to get one on $\mathbb{S}^\infty$ is not straightforward.

- **Quotient/Shape space**. Impossibly hard.

  No invariant measure for infinite-dimensional warping group.

In principle, we can have $p(\text{data}) = \int p(\text{data}|\text{group}) dp(\text{group})$.

Example: mean estimation with marginal models

Models for a random curve that are roughly of the form

$$X(t) = (\mu(t), g(t)) \quad \text{``}+\text{''} \quad \epsilon(t), \quad g \in G$$

where:

- ϵ is a stochastic process (likelihood given g);
- (μ, g) is action of group element g on μ;
- μ is a deterministic mean/template;
- g is a stochastic process (random effect/prior on G)

Integrate out g over G.

Focus of this talk: Stochastic pairwise registration

Let $\Gamma := \{\gamma : D \to D\}$ be a class of warping functions.

Match $f_i : D \to \mathbb{R}^k, i = 1, 2$ with:

Focus of this talk: Stochastic pairwise registration

Let $\Gamma := \{\gamma : D \to D\}$ be a class of warping functions.

Match $f_i : D \to \mathbb{R}^k, i = 1, 2$ with:

- $$\arg\min_{\gamma \in \Gamma} d(f_1, (f_2, \gamma)),$$

  with a distribution on Γ;

Focus of this talk: Stochastic pairwise registration

Let $\Gamma := \{\gamma : D \to D\}$ be a class of warping functions.

Match $f_i : D \to \mathbb{R}^k, i = 1, 2$ with:

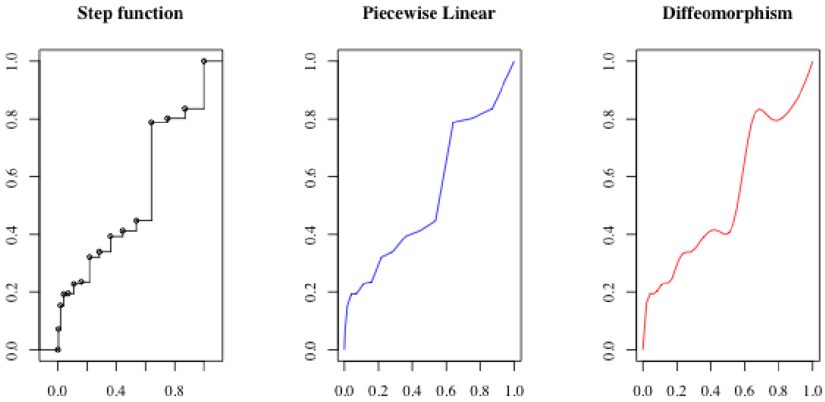
- $$\arg\min_{\gamma \in \Gamma} d(f_1, (f_2, \gamma)),$$

  with a distribution on Γ;

- or a posterior summary based on a Bayesian model:

  likelihood on $\{f_1(t) - (f_2, \gamma) | \gamma\}$ and prior distribution on Γ.

Candidate classes of warping functions



Distributions on warping functions on $[0, 1]$

We wish to have a random function $\gamma : [0, 1] \to [0, 1]$, increasing with $\gamma(0) = 0$ and $\gamma(1) = 1$.

[1] E. Shavgulidze. Trudy Math. Inst. Steklov., 217 (1997),189-208

[2] E. Bardelli and A. G. Menucci. Journal of Geometric Mechanics. 9 (2017), 291-316.

Distributions on warping functions on $[0, 1]$

We wish to have a random function $\gamma : [0, 1] \to [0, 1]$, increasing with $\gamma(0) = 0$ and $\gamma(1) = 1$.

- No 'Haar' measure on Γ—only those quasi-invariant with respect to composition exist. [1].

[1] E. Shavgulidze. Trudy Math. Inst. Steklov., 217 (1997),189-208

[2] E. Bardelli and A. G. Menucci. Journal of Geometric Mechanics. 9 (2017), 291-316.

Distributions on warping functions on $[0,1]$

We wish to have a random function $\gamma : [0,1] \to [0,1]$, increasing with $\gamma(0) = 0$ and $\gamma(1) = 1$.

- No 'Haar' measure on Γ—only those quasi-invariant with respect to composition exist. [1].
- Pushforwards of equivalent Gaussian measures on $\mathbb{L}^2$ tangent spaces of $\mathbb{S}^\infty$ under Exponential maps under are mutually singular. [2].

[1] E. Shavgulidze. Trudy Math. Inst. Steklov., 217 (1997),189-208

[2] E. Bardelli and A. G. Menucci. Journal of Geometric Mechanics. 9 (2017), 291-316.

Distributions on warping functions on $[0, 1]$

We wish to have a random function $\gamma : [0, 1] \to [0, 1]$, increasing with $\gamma(0) = 0$ and $\gamma(1) = 1$.

- No 'Haar' measure on Γ—only those quasi-invariant with respect to composition exist. [1].
- Pushforwards of equivalent Gaussian measures on $\mathbb{L}^2$ tangent spaces of $\mathbb{S}^\infty$ under Exponential maps under are mutually singular. [2].
- Subordinator processes (a.s. non-negative, non-decreasing Levy process) are natural candidates (e.g. Gamma process).

[1] E. Shavgulidze. Trudy Math. Inst. Steklov., 217 (1997),189-208

[2] E. Bardelli and A. G. Menucci. Journal of Geometric Mechanics. 9 (2017), 291-316.

Distributions on warping functions on $[0, 1]$

We wish to have a random function $\gamma : [0, 1] \to [0, 1]$, increasing with $\gamma(0) = 0$ and $\gamma(1) = 1$.

- No 'Haar' measure on Γ—only those quasi-invariant with respect to composition exist. [1].
- Pushforwards of equivalent Gaussian measures on $\mathbb{L}^2$ tangent spaces of $\mathbb{S}^\infty$ under Exponential maps under are mutually singular. [2].
- Subordinator processes (a.s. non-negative, non-decreasing Levy process) are natural candidates (e.g. Gamma process).
- Define a map $\Phi : C_0[0, 1] \to \text{Diff}^1[0, 1]$ as

$$g(t) \mapsto \Phi(g(t)) := \frac{\int_0^t e^{g(s)ds}}{\int_0^1 e^{g(s)}ds}.$$

Choose g to be a Gaussian process for e.g.

[1] E. Shavgulidze. Trudy Math. Inst. Steklov., 217 (1997),189-208

[2] E. Bardelli and A. G. Menucci. Journal of Geometric Mechanics. 9 (2017), 291-316.

Distributions on warping functions on $[0, 1]$

We wish to have a random function $\gamma : [0, 1] \to [0, 1]$, increasing with $\gamma(0) = 0$ and $\gamma(1) = 1$.

- No 'Haar' measure on Γ—only those quasi-invariant with respect to composition exist. [1].
- Pushforwards of equivalent Gaussian measures on $\mathbb{L}^2$ tangent spaces of $\mathbb{S}^\infty$ under Exponential maps under are mutually singular. [2].
- Subordinator processes (a.s. non-negative, non-decreasing Levy process) are natural candidates (e.g. Gamma process).
- Define a map $\Phi : C_0[0, 1] \to \mathrm{Diff}^1[0, 1]$ as

$$g(t) \mapsto \Phi(g(t)) := \frac{\int_0^t e^{g(s)ds}}{\int_0^1 e^{g(s)} ds}.$$

  Choose g to be a Gaussian process for e.g.

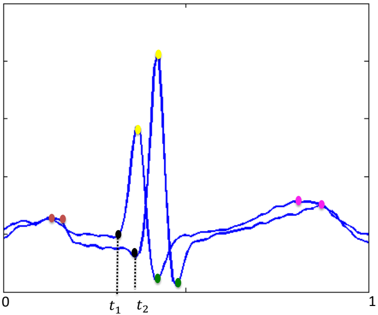
Statistical issue: Not easy to centre these distributions at desired γ or sample from.

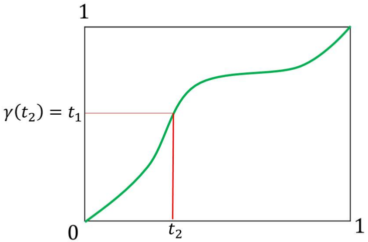
[1] E. Shavgulidze. Trudy Math. Inst. Steklov., 217 (1997),189-208

[2] E. Bardelli and A. G. Menucci. Journal of Geometric Mechanics. 9 (2017), 291-316.

A new variant of the curve/function registration problem is common and motivates the class of warping functions.

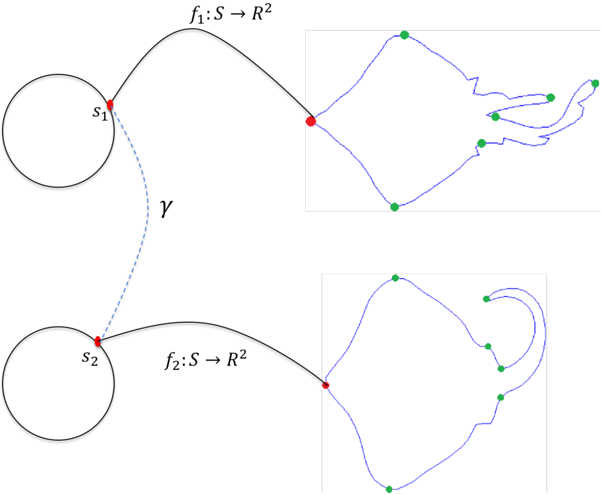
Pairwise alignment with landmarks





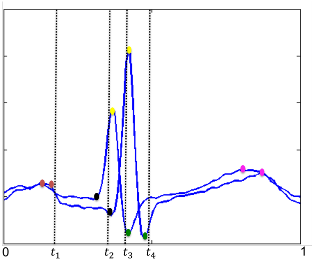
- If on f_1 a landmark is present at $t_1 \in [0, 1]$ and correspondingly at t_2 on f_2, then there is a warping constraint: $\gamma(t_2) = t_1$.

Pairwise alignment with landmarks



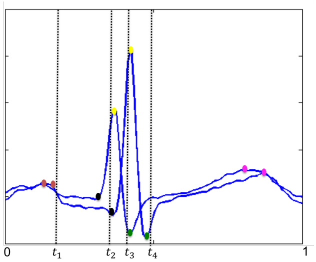
Desiderata in a $\mathbb{P}_\theta$ on Γ

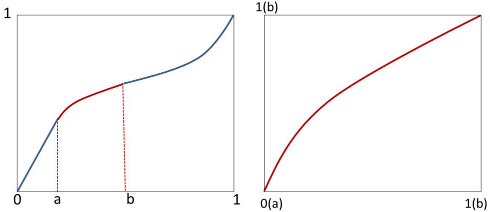
Desiderata in a $\mathbb{P}_\theta$ on Γ



Decompose global alignment into multiple local ones: $\mathbb{P}_\theta$ restricted to $\gamma_{|[t_1,t_2]}$ is independent of $\mathbb{P}_\theta$ restricted to $\gamma_{|[0,1]\setminus[t_1,t_2]}$ (Markov like)

Desiderata in a $\mathbb{P}_\theta$ on Γ





Decompose global alignment into multiple local ones: $\mathbb{P}_\theta$ restricted to $\gamma_{|[t_1,t_2]}$ is independent of $\mathbb{P}_\theta$ restricted to $\gamma_{|[0,1]\setminus[t_1,t_2]}$ (Markov like)

Optimal local alignment 'matches' optimal global one [a]: $\mathbb{P}_{\theta(t_2-t_1)}$ (Self-similarity)

---

[a] *Focus Invariance*: A. Trouve and L. Younes. SIAM Journal on Control and Optimization, 39:1112-1135, 2000

Program

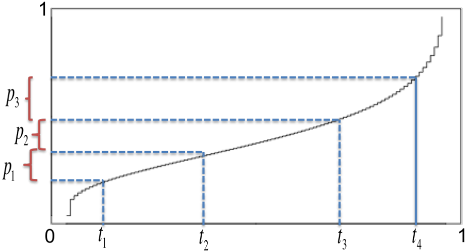
- Landmark constraints imply that diffeomorphism group for Γ is not appropriate.
- Decomposability and Markov-like property point towards a process with independent increments. (Levy)

Program

- Landmark constraints imply that diffeomorphism group for Γ is not appropriate.
- Decomposability and Markov-like property point towards a process with independent increments. (Levy)
- We first examine a simple sampling algorithm, and then pose the question:

  "Does there exists a distribution $\mathbb{P}_\theta$ from which the warping functions are being sampled from? Does it possess the desiderata?"

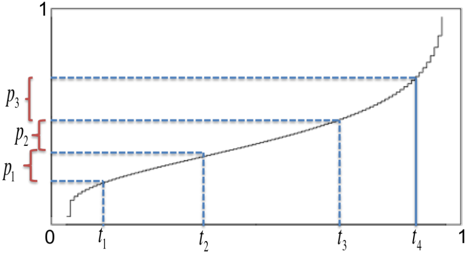
(Remarkably simple) Sampling algorithm [3]



1. Choose a (non-random) set of ordered points $0 =: t_0 < t_1 < t_2 < \cdots < t_{n-1} < t_n := 1$;
2. sample $(p_1, \ldots, p_n)$ from Dirichlet distribution with parameters $\theta(1, \ldots, 1)$;
3. construct a warp map on $[0, 1]$ by linear interpolation.

[3] W. Cheng, I. L. Dryden and X. Huang. Bayesian Analysis, 11, 447-475, 2015

Sampling algorithm: A generalisation



1. Choose a (non-random) set of ordered points $0 =: t_0 < t_1 < t_2 < \cdots < t_{n-1} < t_n := 1$;
2. set $p_i = x_{i:n} - x_{i-1:n}$, $i = 1, \ldots, n$, where $x_{i:n}$ are order statistics from F. ($x_i \sim U[0,1]$ implies $(p_1, \ldots, p_n) \sim$ Dirichlet $(1,\ldots,1)$);
3. construct a warp map on $[0,1]$ by linear interpolation.

'Nice' distribution as $n \to \infty$?

Consider the process

$$Y_n(t) := \sum_{i=1}^{\lfloor nt \rfloor} p_i + (nt - \lfloor nt \rfloor) p_{\lfloor nt \rfloor + 1}, \quad t \in [0, 1].$$

'Nice' distribution as $n \to \infty$?

Consider the process

$$Y_n(t) := \sum_{i=1}^{\lfloor nt \rfloor} p_i + (nt - \lfloor nt \rfloor) p_{\lfloor nt \rfloor + 1}, \quad t \in [0,1].$$

Proposition (Degenerate distribution)

1. ***Algorithm of Chen et al.***
   Under a uniform partition, in $C([0,1])$ with uniform topology: Y_n converges in probability to $\gamma(t) = t$; the process $\sqrt{n}(Y_n(t) - t)$ converges in distribution to a standard Brownian Bridge process.

'Nice' distribution as $n \to \infty$?

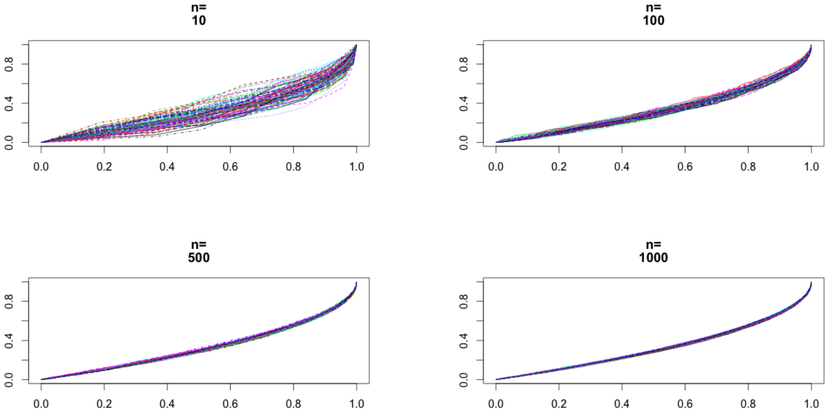
Consider the process

$$Y_n(t) := \sum_{i=1}^{\lfloor nt \rfloor} p_i + (nt - \lfloor nt \rfloor) p_{\lfloor nt \rfloor + 1}, \quad t \in [0, 1].$$

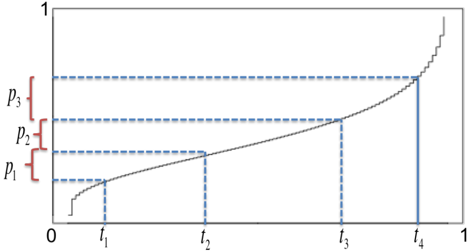
Proposition (Degenerate distribution)

1. ***Algorithm of Chen et al.***
   Under a uniform partition, in $C([0,1])$ with uniform topology: Y_n converges in probability to $\gamma(t) = t$; the process $\sqrt{n}(Y_n(t) - t)$ converges in distribution to a standard Brownian Bridge process.
2. ***Generalised version***
   Under some conditions on F, Y_n converges in probability to $F^{-1}(t)$ in $C[0,1]$ with uniform topology.

Example with F as Beta(1,2)



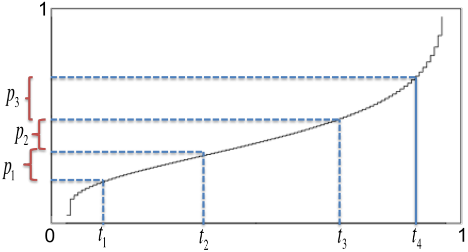
The algorithm can be salvaged . . .



There are two main ingredients in the algorithm:

- Construction of partition $0 < t_1 < \cdots < t_n < 1$ of $[0, 1]$;
- sampling of increments from a distribution on the unit simplex.

The algorithm can be salvaged . . .



There are two main ingredients in the algorithm:

- Construction of partition $0 < t_1 < \cdots < t_n < 1$ of $[0, 1]$;
- sampling of increments from a distribution on the unit simplex.

Choosing a random partition does the trick.

Modified algorithm with random partition

1. Choose $\theta > 0$;
2. discretize $[0, 1]$ with order statistics $0 =: t_0 < t_1 < t_2 < \cdots < t_{n-1} < t_n := 1$ of a random sample from distribution function H on $[0, 1]$;

Modified algorithm with random partition

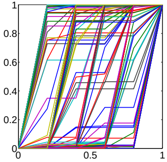
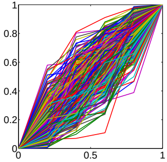
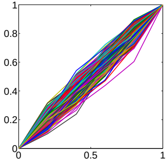
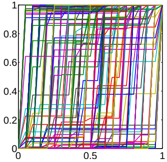
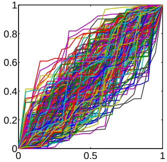
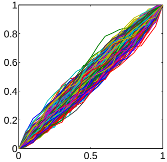
1. Choose $\theta > 0$;
2. discretize $[0, 1]$ with order statistics $0 =: t_0 < t_1 < t_2 < \cdots < t_{n-1} < t_n := 1$ of a random sample from distribution function H on $[0, 1]$;
3. sample an n-dimensional Dirichlet distributed random vector $(p_1, \ldots, p_n)$ with parameters set to $\theta\,(t_{1:n} - t_{0:n}, \ldots, t_{n:n} - t_{n-1:n})$;

Modified algorithm with random partition

1. Choose $\theta > 0$;
2. discretize $[0,1]$ with order statistics $0 =: t_0 < t_1 < t_2 < \cdots < t_{n-1} < t_n := 1$ of a random sample from distribution function H on $[0,1]$;
3. sample an n-dimensional Dirichlet distributed random vector $(p_1, \ldots, p_n)$ with parameters set to $\theta\,(t_{1:n} - t_{0:n}, \ldots, t_{n:n} - t_{n-1:n})$;
4. construct a warp map on $[0,1]$ by linear interpolation.

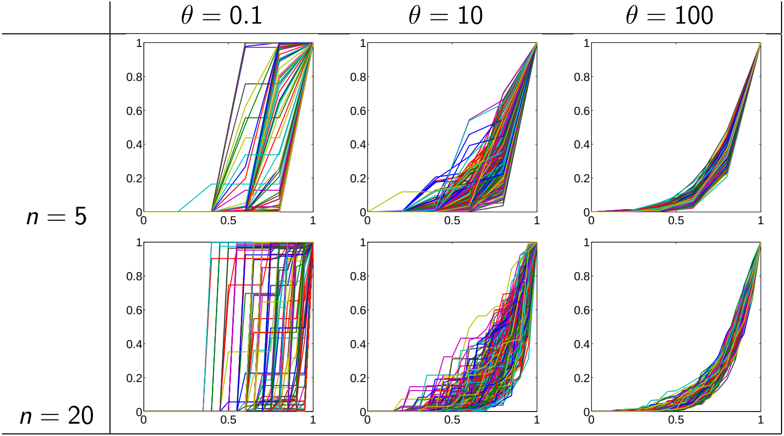
Example

Samples of warps with partition based on $H(t) = t$

	$\theta = 0.1$	$\theta = 10$	$\theta = 100$
$n = 5$			
$n = 20$			

Example

Samples of warps with partition based Beta (5,1) CDF as H



Justification for modified algorithm

- Consider partition based on H.
- Independently, let $(p_1, \ldots, p_n)$ obtained as spacings of an i.i.d sequence x_i from a density f (not necessarily Dirichlet distributed).

Justification for modified algorithm

- Consider partition based on H.
- Independently, let $(p_1, \ldots, p_n)$ obtained as spacings of an i.i.d sequence x_i from a density f (not necessarily Dirichlet distributed).
- Let $\lambda(x) = \theta \int_x^\infty \frac{e^{-y}}{y} dy$ for $\theta > 0$.
- Set $v_1 = p_1$ and $v_i = p_1 + \cdots + p_i, i = 2, \ldots, n$ and consider the transformed random variables $z_{i,n} = \lambda^{-1}(nf(F^{-1}(\zeta_{i,n}))v_i)$ where $0 \le \zeta_{i,n} \le 1$ is a deterministic sequence such that $\max_{1 \le i \le n} |\frac{i}{n} - \zeta_{i,n}| = O(1/n)$.

Process based on random partition

Theorem

- *Under some conditions on f, the linearly interpolated version of the process*

$$G_n(t) := \sum_i \lambda^{-1}(z_{i,n})\mathbb{I}_{t_i \le t}, \quad t \in [0,1],$$

converges weakly to the time-changed pure jump Gamma process $\mathcal{G} \circ H$ in the Skorohod M_1 topology (J_1 topology is too strong).

Process based on random partition

Theorem

- *Under some conditions on f, the linearly interpolated version of the process*

$$G_n(t) := \sum_i \lambda^{-1}(z_{i,n})\mathbb{I}_{t_i \le t}, \quad t \in [0,1],$$

  converges weakly to the time-changed pure jump Gamma process $\mathcal{G} \circ H$ in the Skorohod M_1 topology (J_1 topology is too strong).
- *$\mathcal{D}^\theta \circ H = \mathcal{G}(H(t))/\mathcal{G}(H(\theta))$ is referred to as the Dirichlet process with sample paths as warp maps of $[0,1]$ with*

$$\mathbb{E}\mathcal{D}^\theta \circ H = H, \quad \forall \theta.$$

Interpretation of $\mathbb{P}_{H,\theta}$, the law of $\mathcal{D}^\theta \circ H$

Conditioned on a partition $0 =: t_{0:n} < t_{1:n} < \cdots < t_{n-1:n} < t_{n:n} := 1$ from H,

$$\mathbb{P}_{H,\theta}(\gamma(t_{1:n}) \in dx_1, \ldots, \gamma(t_{n-1:n}) \in dx_{n-1})$$
$$= \frac{\Gamma(\theta)}{\prod_{i=1}^n \Gamma(\theta(t_{i:n} - t_{i-1:n}))} \prod_{i=1}^n (x_i - x_{i-1})^{(t_{i:n} - t_{i-\mathbf{1}:n})} dx_1 \ldots dx_{n-1}.$$

[4] **T. Ferguson (1973). Annals of Statistics. 209–230**

Interpretation of $\mathbb{P}_{H,\theta}$, the law of $\mathcal{D}^\theta \circ H$

Conditioned on a partition $0 =: t_{0:n} < t_{1:n} < \cdots < t_{n-1:n} < t_{n:n} := 1$ from H,

$$\mathbb{P}_{H,\theta}(\gamma(t_{1:n}) \in dx_1, \ldots, \gamma(t_{n-1:n}) \in dx_{n-1})$$
$$= \frac{\Gamma(\theta)}{\prod_{i=1}^{n} \Gamma(\theta(t_{i:n} - t_{i-1:n}))} \prod_{i=1}^{n} (x_i - x_{i-1})^{(t_{i:n} - t_{i-\mathbf{1}:n})} dx_1 \ldots dx_{n-1}.$$

- $\mathbb{P}_{H,\theta}$ is push-forward of the well-known Dirichlet process[4] on $\mathcal{P}[0,1]$, the set of prob measures on $[0,1]$, under the map that takes a probability measure to its quantile function.

[4] **T. Ferguson (1973). Annals of Statistics. 209–230**

Comments

- $\mathbb{P}_{H,\theta}$ can be centred at desired warp map: choice of H when constructing the random partition.
- θ acts like a variance parameter.

Some properties of $\mathbb{P}_{H,\theta}$

- It has full support on set of warp maps Γ for every θ (several topologies coincide).

Some properties of $\mathbb{P}_{H,\theta}$

- It has full support on set of warp maps Γ for every θ (several topologies coincide).
- For every $\gamma \in \Gamma$, $\gamma_{\#}\mathbb{P}_{H,\theta}$ is absolutely continuous w.r.t $\mathbb{P}_{H,\theta}$ (quasi-invariance under composition).

Some properties of $\mathbb{P}_{H,\theta}$

- It has full support on set of warp maps Γ for every θ (several topologies coincide).
- For every $\gamma \in \Gamma$, $\gamma_{\#}\mathbb{P}_{H,\theta}$ is absolutely continuous w.r.t $\mathbb{P}_{H,\theta}$ (quasi-invariance under composition).
- When restricted to $[a, b]$ and rescaled, the resulting distribution is $\mathbb{D}_{\theta(H(b)-H(a))}$ (Self-similarity).
- Independent of $\mathbb{D}_{H,\theta}$ on $[0, 1]\backslash[a, b]$ except at a and b ('Markov').

Some properties of $\mathbb{P}_{H,\theta}$

- It has full support on set of warp maps Γ for every θ (several topologies coincide).
- For every $\gamma \in \Gamma$, $\gamma_{\#}\mathbb{P}_{H,\theta}$ is absolutely continuous w.r.t $\mathbb{P}_{H,\theta}$ (quasi-invariance under composition).
- When restricted to $[a, b]$ and rescaled, the resulting distribution is $\mathbb{D}_{\theta(H(b)-H(a))}$ (Self-similarity).
- Independent of $\mathbb{D}_{H,\theta}$ on $[0, 1]\backslash[a, b]$ except at a and b ('Markov').
- As $\theta \to 0$, $\mathbb{D}_{H,\theta}$ converges to uniform distribution on the set $\{\mathbb{I}_{[H(t),1]} : 0 \leq t \leq 1\}$;
- as $\theta \to \infty$, $\mathbb{D}_{H,\theta}$ converges to point mass at $H(t)$.

$\mathbb{P}_{H,\theta}$ generates warps with jumps...

Jumps correspond to large deformations of subsets of $[0, 1]$. How large? How often do they occur?

$\mathbb{P}_{H,\theta}$ generates warps with jumps...

Jumps correspond to large deformations of subsets of $[0,1]$. How large? How often do they occur? Let $\xi_{i,n} := np_i - \log n$ be normalised increments.

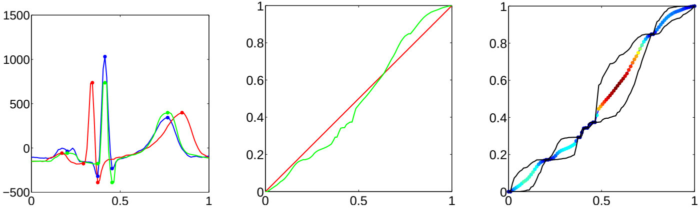
Consider the process $Y_n(t) := \sum_{i=1}^{\lfloor nt \rfloor}[\xi_i - E(\xi_i \mathbb{I}_{\xi_i \leq 1})],\ 0 \leq t \leq 1$.

Theorem

The sequence Y_n converges weakly in the J_1 topology to a real-valued Levy jump process with Levy measure $\nu(dy) = e^{-y}dy$ in $D[0,1]$.

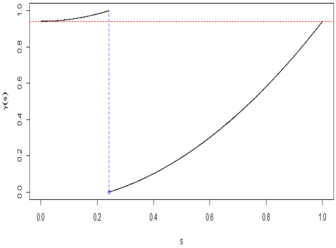
ECG Data: Bayesian model on Ambient/Top space

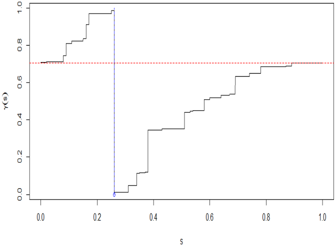
- Gaussian likelihood on $q_1 - q_2(\gamma)|\gamma$, where $q_i, i = 1, 2$ are Square Root Velocity transform (SRV) of $f_i, i = 1, 2$.
- $\mathbb{D}_{H,\theta}$ prior on Γ with $H(t) = t$.



Green: Aligned function; estimated warp map. (Pointwise) Posterior credible intervals. (Blue: less uncertainty; Red: more uncertainty).

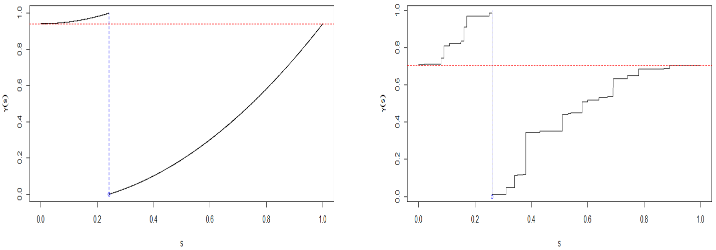
Sampling when $D = \mathbb{S}^1$ by unwrapping





$$\left\{ \text{AC warp maps of } \mathbb{S}^1 \right\} \overset{\textit{Bijection}}{\longrightarrow} \left\{ \text{AC warp maps of } [0,1] \right\} \times \mathbb{S}^1$$

Sampling when $D = \mathbb{S}^1$ by unwrapping

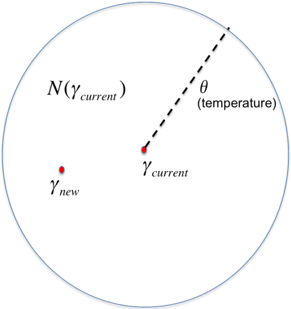


$$\left\{ \text{AC warp maps of } \mathbb{S}^1 \right\} \overset{Bijection}{\rightarrow} \left\{ \text{AC warp maps of } [0,1] \right\} \times \mathbb{S}^1$$

1. Choose c from μ on $[0,1]$ (by identifying $\mathbb{S}^1$ with $\mathbb{R}/\mathbb{Z}$);
2. Sample γ from $\mathbb{D}_\theta \circ H$; and
3. Set $\gamma_s(t) := (\gamma(t) + c) \mod 1$.

Simulated Annealing with $\mathbb{P}_{H,\theta}$

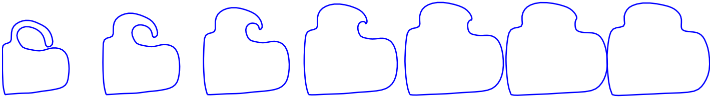
Solve: $\arg\min_{\gamma \in \Gamma} d(f_1(t), f_2(\gamma(t)))$

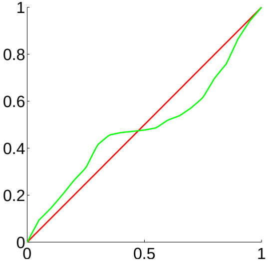


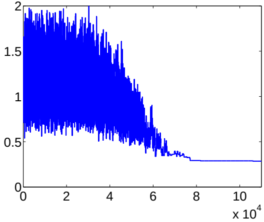
- Choose $\gamma_{new} \in N(\gamma_{current})$ by sampling $\mathbb{P}_{\gamma_{current},\theta}$ centered at $\gamma_{current}$;
- compute $\Delta = d(f_1(t), f_2(\gamma_{new}(t)))$;
- if $\Delta \leq 0$, $\gamma_{current} \leftarrow \gamma_{new}$, else $\gamma_{current} \leftarrow \gamma_{new}$ with prob $e^{-\frac{\Delta}{\theta}}$;
- repeat.

Aligning shapes of closed curves in $\mathbb{R}^2$: Simulated annealing

Geodesic path between leftmost and rightmost figures under SRVF.







Summary

- Landmark constraints under curves impose constraints on warp maps.
- For alignment with landmarks $\mathbb{P}_{H,\theta}$ is compatible with Focus invariance and Decomposability.
- These characterize $\mathbb{P}_{H,\theta}$.
- Algorithm is trivial to implement, and can generate maps from distribution centered at derived at any warp map.

Details and more examples in paper:
Partition-based sampling of warp maps for curve alignment.
arXiv:1708.04891.