Simulation Study of Automated Thickness Analyzing Machine (ATAM)

Chen Yue & Aaron Fisher

Johns Hopkins Bloomberg School of Public Health
Department of Biostatistics

December 20, 2012

Background and Motivation

Background and Motivation

1. **Several diseases have been found to be related to white matter loss in brain, such as Autism (Vidal 2006) and ADHD (Luders 2009). Here, we measure white matter loss based the thickness of the mid-sagittal slice of the corpus callosum (CC), the largest white matter structure in brain.**

Background and Motivation

1. Several diseases have been found to be related to white matter loss in brain, such as Autism (Vidal 2006) and ADHD (Luders 2009). Here, we measure white matter loss based the thickness of the mid-sagittal slice of the corpus callosum (CC), the largest white matter structure in brain.

2. **Our objective: To analyze the link between white matter thickness along the CC, and Expanded Disability Status Scale (EDSS) score in patients with multiple sclerosis (MS). Specifically, if there are regions of the CC that are predictive of EDSS score, we want to find out where they are.**

Background and Motivation

1. Several diseases have been found to be related to white matter loss in brain, such as Autism (Vidal 2006) and ADHD (Luders 2009). Here, we measure white matter loss based the thickness of the mid-sagittal slice of the corpus callosum (CC), the largest white matter structure in brain.
2. Our objective: To analyze the link between white matter thickness along the CC, and Expanded Disability Status Scale (EDSS) score in patients with multiple sclerosis (MS). Specifically, if there are regions of the CC that are predictive of EDSS score, we want to find out where they are.
3. **Today we'll present a simulation study on our proposed method.**

Background and Motivation

Background and Motivation

1. **176 patients with Multiple Sclerosis (MS), 466 scans.**

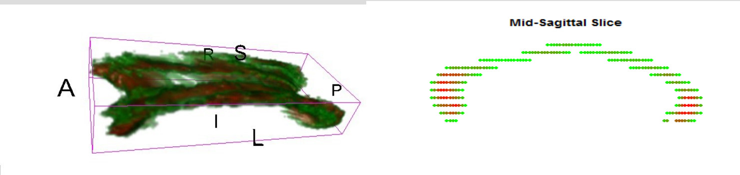
Background and Motivation

1. 176 patients with Multiple Sclerosis (MS), 466 scans.
2. **EDSS score is ranging from 0 to 8.5, higher value suggests worse situation of disease.**

Background and Motivation

1. 176 patients with Multiple Sclerosis (MS), 466 scans.
2. EDSS score is ranging from 0 to 8.5, higher value suggests worse situation of disease.

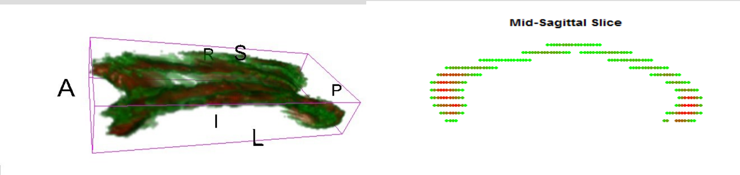
Figure: CC 3D rendering and mid-sagittal slice



Background and Motivation

1. 176 patients with Multiple Sclerosis (MS), 466 scans.
2. EDSS score is ranging from 0 to 8.5, higher value suggests worse situation of disease.

Figure: CC 3D rendering and mid-sagittal slice



Color indicates corresponding Fractional Anisotropy (FA) value.

Pipeline

Pipeline

1. **Obtain the center curve (principal curve) of the target data cloud.**

Pipeline

1. Obtain the center curve (principal curve) of the target data cloud.
2. **Measure the thickness.**

Pipeline

1. Obtain the center curve (principal curve) of the target data cloud.
2. Measure the thickness.
3. **Regress the thickness function on a scalar outcome.**

Pipeline

1. Obtain the center curve (principal curve) of the target data cloud.
2. Measure the thickness.
3. Regress the thickness function on a scalar outcome.
4. **Hypothesis testing.**

Simulation Settings

Simulation Settings

For ease of computation, we first generate Y_i, $(i = 1, 2, ...I)$ and then generate I shapes with features of the i^{th} shape determined by the value of Y_i.

Simulation Settings

For ease of computation, we first generate Y_i, $(i = 1, 2, ...I)$ and then generate I shapes with features of the i^{th} shape determined by the value of Y_i.

- **Case 1: I=300, response $Y_i \sim$ Possion$(\lambda = 3)$. Thickness of image$_i$ within $(\frac{\pi}{3}, \frac{3\pi}{4})$ and $(\frac{5\pi}{4}, \frac{7\pi}{4})$ is related to Y_i.**

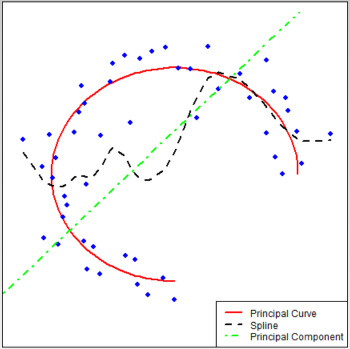
Simulation Settings

For ease of computation, we first generate Y_i, $(i = 1, 2, ...I)$ and then generate I shapes with features of the i^{th} shape determined by the value of Y_i.

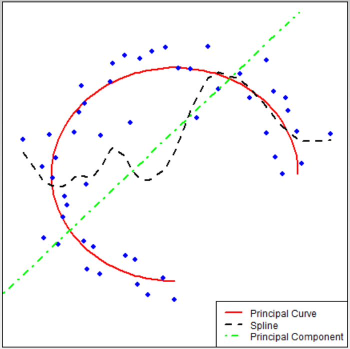
- Case 1: I=300, response $Y_i \sim$ Possion$(\lambda = 3)$. Thickness of image$_i$ within $(\frac{\pi}{3}, \frac{3\pi}{4})$ and $(\frac{5\pi}{4}, \frac{7\pi}{4})$ is related to Y_i.

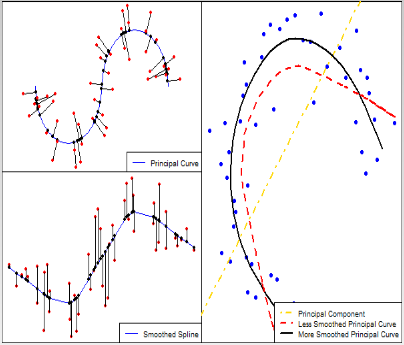
- **Case 2: I=300, response $Y_i \sim$ Possion$(\lambda = 3)$. Thickness of image$_i$ in the vertical bar is related to Y_i.**

Principal curve



Principal curve





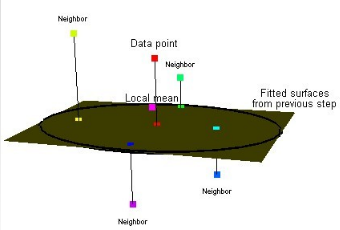
Principal Curve Algorithm

Principal Curve Algorithm

1. **The Projection Step. The points are projected onto the curve of the previous iteration.**

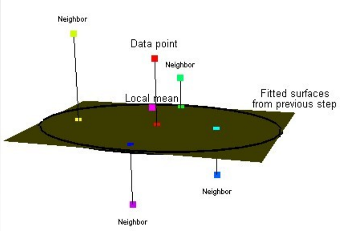
Principal Curve Algorithm

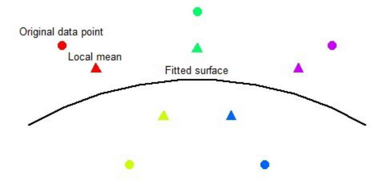
1. **The Projection Step.** The points are projected onto the curve of the previous iteration.
2. **The Conditional Expectation Step. For each data point $\mathrm{x}^{(i)}$, we calculate a locally average $\bar{\mathrm{x}}^{(i)}$.**



Principal Curve Algorithm

1. **The Projection Step.** The points are projected onto the curve of the previous iteration.
2. **The Conditional Expectation Step.** For each data point $\mathbf{x}^{(i)}$, we calculate a locally average $\bar{\mathbf{x}}^{(i)}$.
3. **The Smoothing Step. Fitting a fast TPS using all the local average data point $\bar{\mathbf{x}}^{(i)}$, obtain f^{new}.**





Principal curve fitting result

Obtaining Thickness along the principal curve

Obtaining Thickness along the principal curve

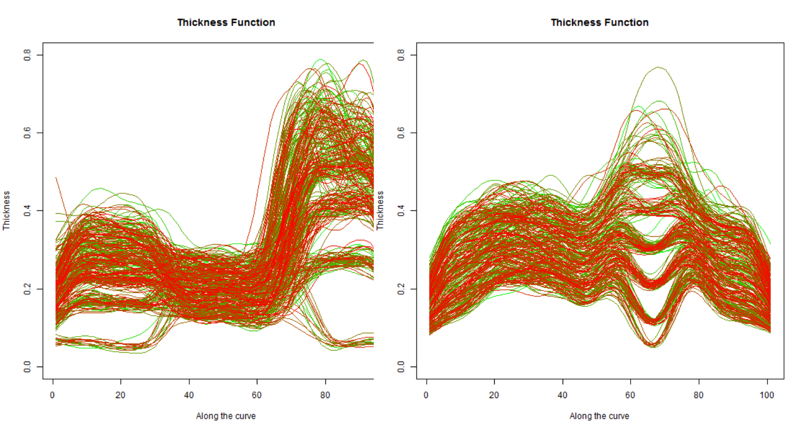
- **Thick**$^*(t) =$ **Quantile**$\big(\{2 \times$ **dist**$_{t_j}, \big||t_j - t| \leq .02\},\ 0.95\big)$**.**

Obtaining Thickness along the principal curve

- Thick$^*(t) = \text{Quantile}\big(\{2 \times \text{dist}_{t_j}, \big||t_j - t| \leq .02\},\ 0.95\big)$.
- **Fit a smooth spline across all Thick$^*(t)$.**

Obtaining Thickness along the principal curve

- Thick$^*(t)$ = Quantile$\left(\{2 \times \text{dist}_{t_j}, \big| |t_j - t| \le .02\},\ 0.95\right)$.
- Fit a smooth spline across all Thick$^*(t)$.



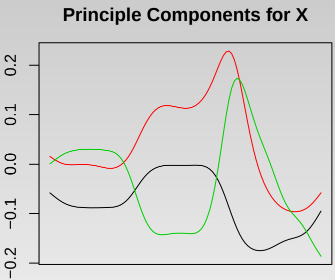
Regression Notation

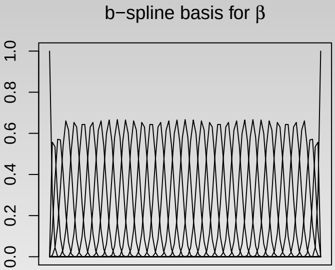
Y_i = Outcome (simulated EDSS score)
$X_i(t)$= The thickness function along the track of the structure
$\beta(t)$ = Function representing effect of thickness on outcome, at different points.
$t \in [0, 1]$.

$$Y_i \sim Poisson(\theta_i)$$
$$\theta_i = \beta_0 + \int_0^1 X_i(t)\beta(t)dt$$

Regression Notation

Estimate the functions $X_i(t)$ and $\beta(t)$ using basis functions.
For X, we'll use it's principle components. For β, we'll use a b-spline basis.





$$X_i(t) \approx \sum_{k=1}^{K_x} \xi_{ik} \Psi_k(t)$$

$$\beta(t) \approx \sum_{l=1}^{K_\beta} \beta_l \varphi_l(t)$$

Regression Notation

$$
\begin{aligned}
\int_0^1 X_i(t)\beta(t)dt &\approx \int_0^1 \left(\sum_{k\in\{A\}} \xi_{ik}\Psi_k(t)\right)\left(\sum_{j\in\{B\}} \beta_j\varphi_j(t)\right)dt \\
&= \int_0^1 \left(\sum_{(k,j)\in\{A\}\times\{B\}} \xi_{ik}\Psi_k(t)\varphi_j(t)\beta_j\right)dt \\
&= \sum_{(k,j)\in\{A\}\times\{B\}} \xi_{ik}\left(\int_0^1 \Psi_k(t)\varphi_j(t)dt\right)\beta_j \\
&=: \xi_i\mathbf{J}\beta^T
\end{aligned}
$$

- Where $\xi_i = [\xi_{i1}, ..\xi_{iK_X}]^T$, $\beta = [\beta_1, ...\beta_{K_\beta}]^T$, and $\mathbf{J}$ is a $K_x \times K_\beta$ matrix with $(i,j)^{th}$ element equal to $\int_0^1 \Psi_i(t)\varphi_j(t)dt$.

Regression Notation

$$
\begin{aligned}
\int_0^1 X_i(t)\beta(t)dt &\approx \int_0^1 \left(\sum_{k\in\{A\}} \xi_{ik}\Psi_k(t)\right)\left(\sum_{j\in\{B\}} \beta_j\varphi_j(t)\right)dt \\
&= \int_0^1 \left(\sum_{(k,j)\in\{A\}\times\{B\}} \xi_{ik}\Psi_k(t)\varphi_j(t)\beta_j\right)dt \\
&= \sum_{(k,j)\in\{A\}\times\{B\}} \xi_{ik}\left(\int_0^1 \Psi_k(t)\varphi_j(t)dt\right)\beta_j \\
&=: \xi_i\mathbf{J}\beta^T
\end{aligned}
$$

- Where $\xi_i = [\xi_{i1}, ..\xi_{iK_X}]^T$, $\beta = [\beta_1, ...\beta_{K_\beta}]^T$, and $\mathbf{J}$ is a $K_x \times K_\beta$ matrix with $(i,j)^{th}$ element equal to $\int_0^1 \Psi_i(t)\varphi_j(t)dt$.
- For this project we'll assume $\xi_i\mathbf{J}$ is fixed, and will put a penalty prior on β.

Full Model

$$
\begin{aligned}
Y_i &\sim Poisson(\theta_i) \\
\theta_i &= \beta_0 + \xi_i \mathbf{J} \beta^T \\
\beta_0, \beta_1 &\sim N(0, 100) \\
\beta_j &\sim N(\beta_{j-1}, 1/\tau_\beta) \\
\tau_\beta &\sim \Gamma(.001, .001)
\end{aligned}
$$

[1] (Crainiceanu and Goldsmith 2010; Brezger, Kneib, and Lang 2005; Lang and Brezger 2004; Goldsmith, Feder, Crainiceanu, Caffo, and Reich, 2010)

Full Model

$$
\begin{aligned}
Y_i &\sim Poisson(\theta_i) \\
\theta_i &= \beta_0 + \xi_i \mathbf{J} \beta^T \\
\beta_0, \beta_1 &\sim N(0, 100) \\
\beta_j &\sim N(\beta_{j-1}, 1/\tau_\beta) \\
\tau_\beta &\sim \Gamma(.001, .001)
\end{aligned}
$$

This could be made to account for measurement error by letting:
$W_i(t) = \sum_{k=1}^{K_x} \xi_{ik} \Psi_k(t) + \epsilon_i(t)$, and assigning priors for $\epsilon_i(t)$ and ξ_{ik}.

[1] (Crainiceanu and Goldsmith 2010; Brezger, Kneib, and Lang 2005; Lang and Brezger 2004; Goldsmith, Feder, Crainiceanu, Caffo, and Reich, 2010)

Full Model

$$
\begin{aligned}
Y_i &\sim Poisson(\theta_i) \\
\theta_i &= \beta_0 + \xi_i \mathbf{J} \beta^T \\
\beta_0, \beta_1 &\sim N(0, 100) \\
\beta_j &\sim N(\beta_{j-1}, 1/\tau_\beta) \\
\tau_\beta &\sim \Gamma(.001, .001)
\end{aligned}
$$

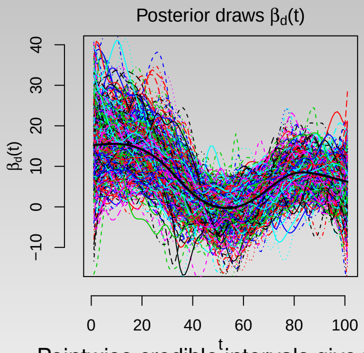
This could be made to account for measurement error by letting: $W_i(t) = \sum_{k=1}^{K_x} \xi_{ik} \Psi_k(t) + \epsilon_i(t)$, and assigning priors for $\epsilon_i(t)$ and ξ_{ik}.

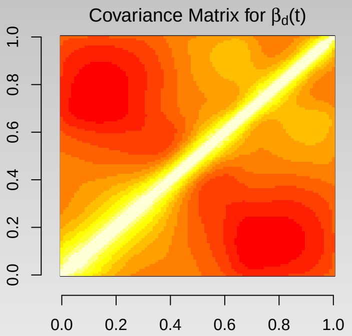
Fit this in WinBUGS. All parameters mixed well except τ_β.[1]

[1] (Crainiceanu and Goldsmith 2010; Brezger, Kneib, and Lang 2005; Lang and Brezger 2004; Goldsmith, Feder, Crainiceanu, Caffo, and Reich, 2010)

Hypothesis Testing

Given draws $(\beta_d(t))$ from the posterior:



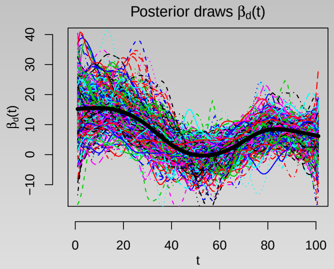


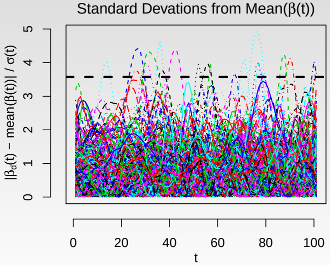
- Pointwise credible intervals give the range along the curve where 95% of the points will fall
  - To estimate this, we need the quantiles for $\beta_d(t)$.
- Joint credible intervals give the area within which 95% of the curves will be completely contained
  - Roughly speaking, we need the quantiles for each curve's largest deviation from the mean curve.

Hypothesis Testing

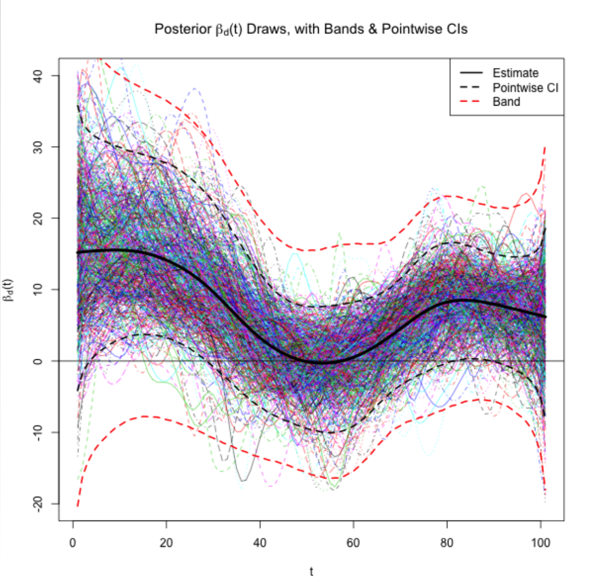
Let:

- $\hat{\beta}(t) = mean(\beta_d(t))$
- $\sigma(t)^2 = var(\beta_d(t))$
- $S_d(t) = \frac{|\beta_d(t) - \hat{\beta}(t)|}{\sigma(t)}$
    - Assume $P(S_d(t) > s)$ is constant over t, given s
- $m_d = max_t \{S_d(t)\}$
- $q(m_d, .95)$ is the 95% quantile of m_d across d.
- Reconstruct joint CIs using $\hat{\beta}(t) \pm q(m_d, .95)\sigma(t)$.





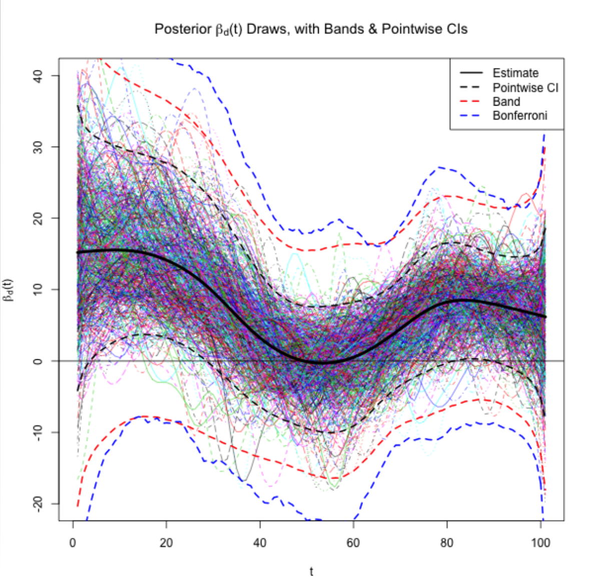
Pointwise & Join Credible Intervals - C Shape



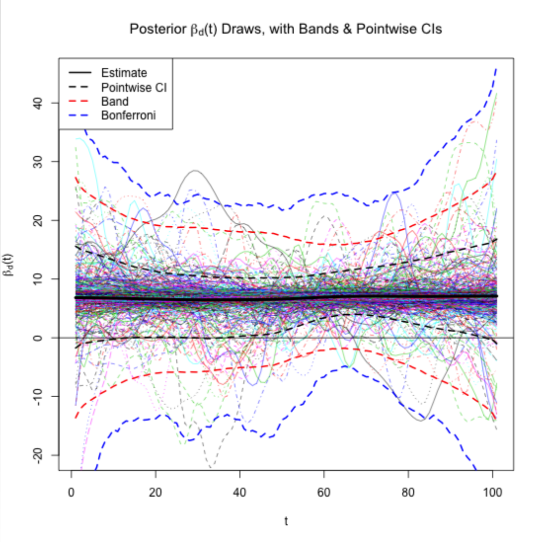
Comparison to a Bonferroni

- Pointwise intervals above are calculated above using the (.5α) & (1-.5α) quantiles at each point t (with $\alpha = .05$).
- We could also create a confidence band using a Bonferroni style adjustment, using the ($.5\alpha/100$) and $(1 - \frac{.5\alpha}{100})$

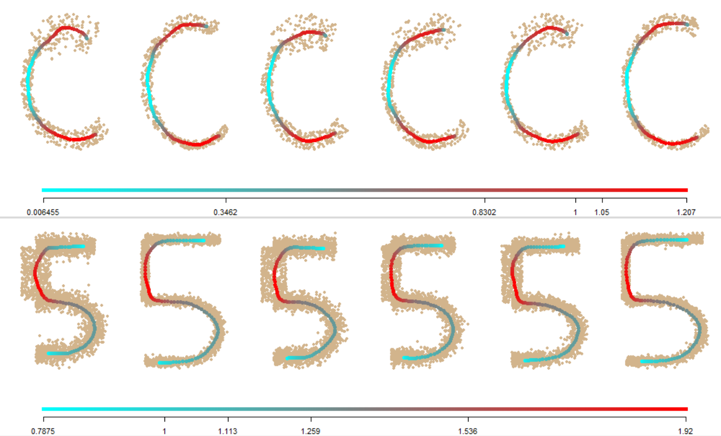
Pointwise & Joint Credible Intervals - C Shape



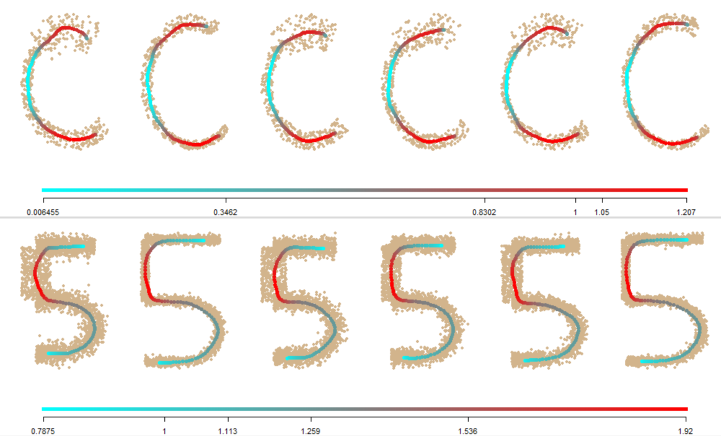
Pointwise & Joint Credible Intervals - 5 Shape



Back Mapping



Back Mapping



- **Color indicates value** $\frac{\hat{\beta}(t)}{\frac{1}{2}\text{Width}_{CI}(t)}$

Discussion

- **The joint CI never rises above zero, but the pointwise CI does at several points. These points could be the subject of future tests, if these were real data.**

Discussion

- The joint CI never rises above zero, but the pointwise CI does at several points. These points could be the subject of future tests, if these were real data.

- **With respect to the simulation:**
  - **In our favor, not much non-noise signal in $X(t)$, so PCA works well.**
  - **Against us, the true underlying $\beta(t)$ is not smooth.**

Discussion

- The joint CI never rises above zero, but the pointwise CI does at several points. These points could be the subject of future tests, if these were real data.
- With respect to the simulation:
  - In our favor, not much non-noise signal in $X(t)$, so PCA works well.
  - Against us, the true underlying $\beta(t)$ is not smooth.
- **Principle curve procedure sometimes doesn't work so well at the end points, it they are especially thick.**
- **Results are slightly sensitive to number of splines chosen.**

Discussion

- The joint CI never rises above zero, but the pointwise CI does at several points. These points could be the subject of future tests, if these were real data.
- With respect to the simulation:
  - In our favor, not much non-noise signal in $X(t)$, so PCA works well.
  - Against us, the true underlying $\beta(t)$ is not smooth.
- Principle curve procedure sometimes doesn't work so well at the end points, it they are especially thick.
- Results are slightly sensitive to number of splines chosen.
- **Pipeline is good,**

Discussion

- The joint CI never rises above zero, but the pointwise CI does at several points. These points could be the subject of future tests, if these were real data.
- With respect to the simulation:
  - In our favor, not much non-noise signal in $X(t)$, so PCA works well.
  - Against us, the true underlying $\beta(t)$ is not smooth.
- Principle curve procedure sometimes doesn't work so well at the end points, it they are especially thick.
- Results are slightly sensitive to number of splines chosen.
- Pipeline is good,
- **but hard.**
  - **Make sure one step are not throwing garbage into the next one.**

Acknowledgement

Thanks Dr. Crainiceanu, C., Gellar, J and Huang, L. for their help.

Thanks for everyone's patience!