Regularized Multivariate Functional Principal Component Analysis

Mehdi Maadooliat
(Joint work with Yue Zhao and Dr. Hossein Haghbin)

Associate Professor
Department of Mathematical and Statistical Sciences
Marquette University
mehdi.maadooliat@mu.edu

5th Seminar on Spatial Statistics and Its Applications
IKIU, Qazvin, IR
(October 25, 2023)

Background: From FPCA to Regularized FPCA

- Performance of FPCA is often enhanced by regularization techniques.

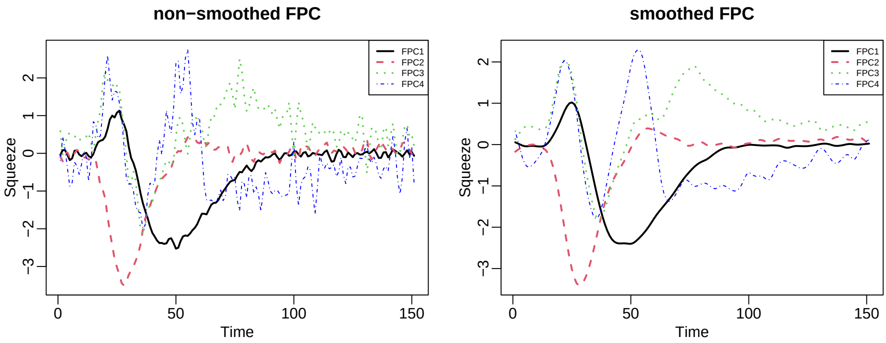
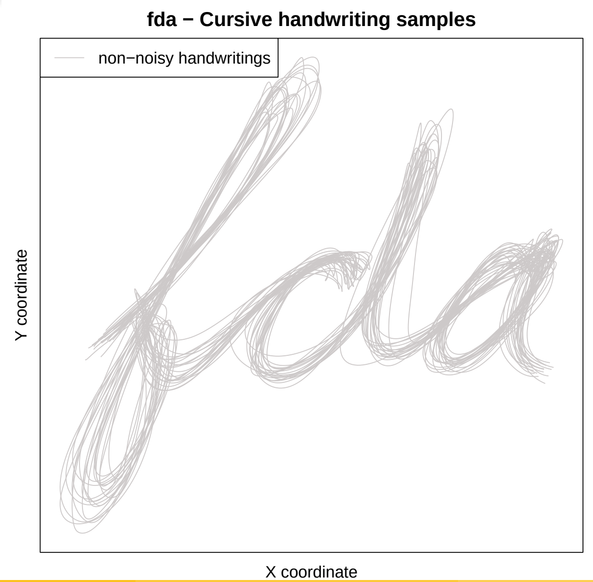
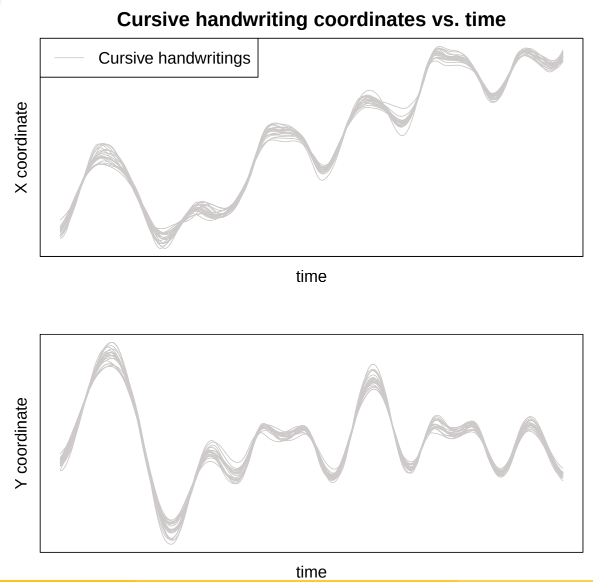


Figure: Estimated first four principal components for the pinch force data.
Left: non-smoothed FPCs.
Right: smoothed FPCs

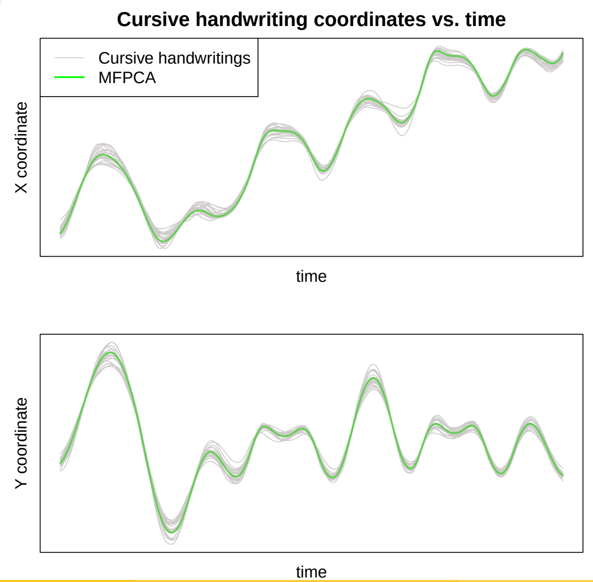
Motivating Example: Cursive Handwriting



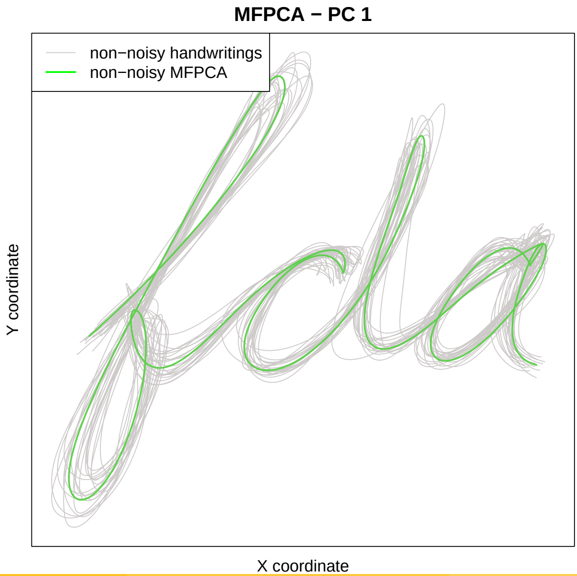
Motivating Example: Cursive Handwriting



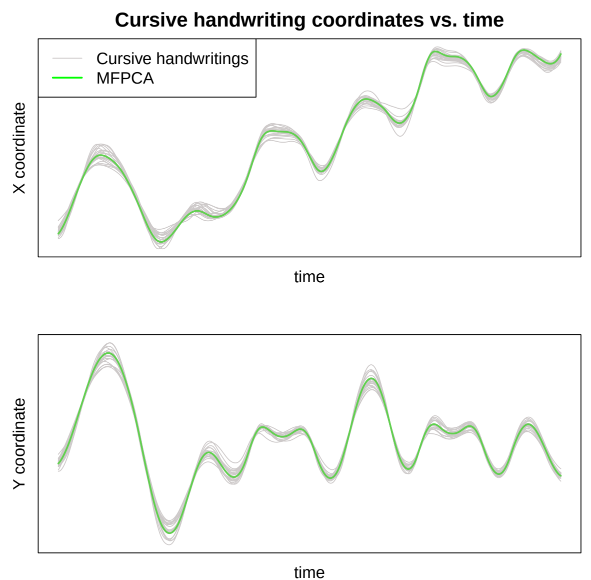
Motivating Example: Cursive Handwriting



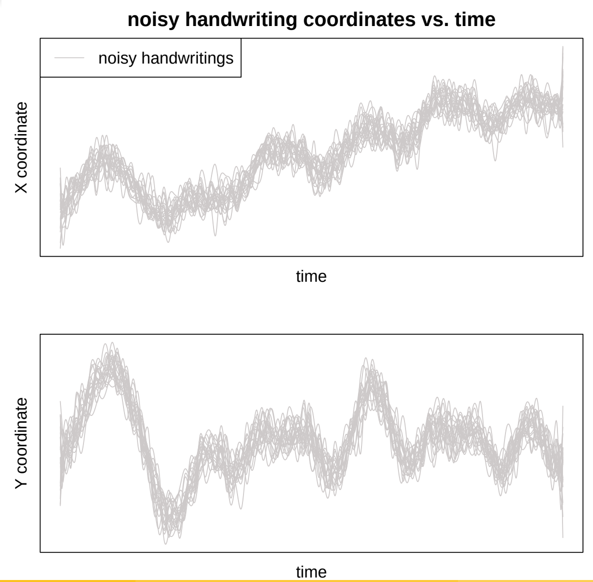
Motivating Example: Cursive Handwriting



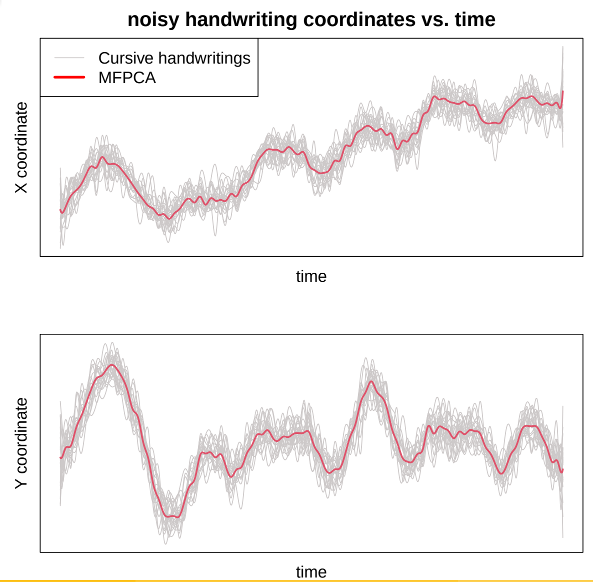
Motivating Example: Cursive Handwriting



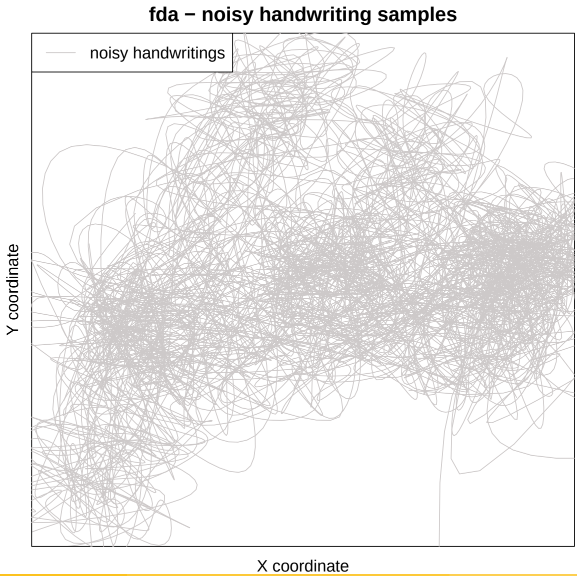
Motivating Example: Cursive Handwriting



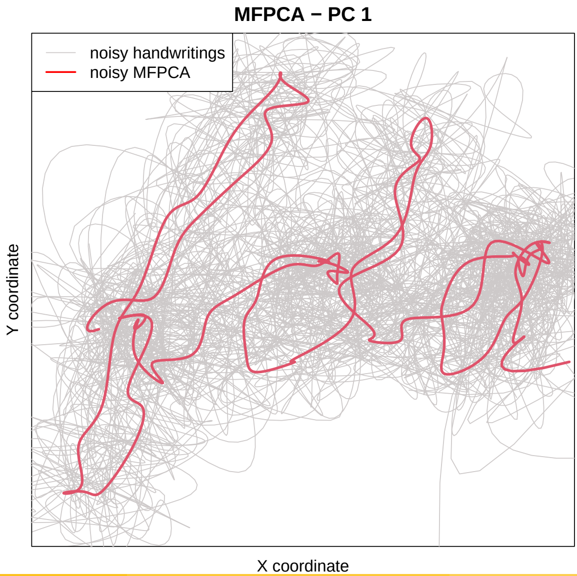
Motivating Example: Cursive Handwriting



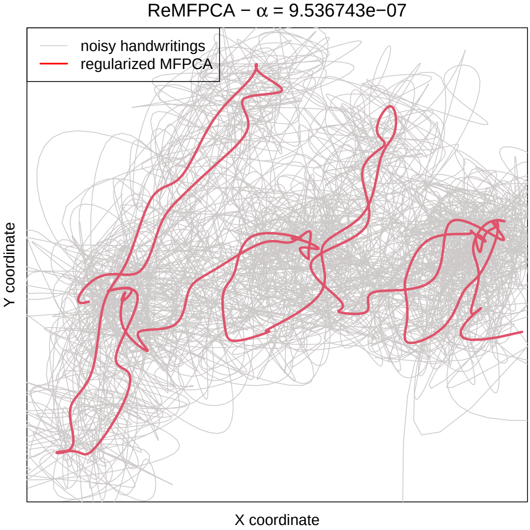
Motivating Example: Cursive Handwriting



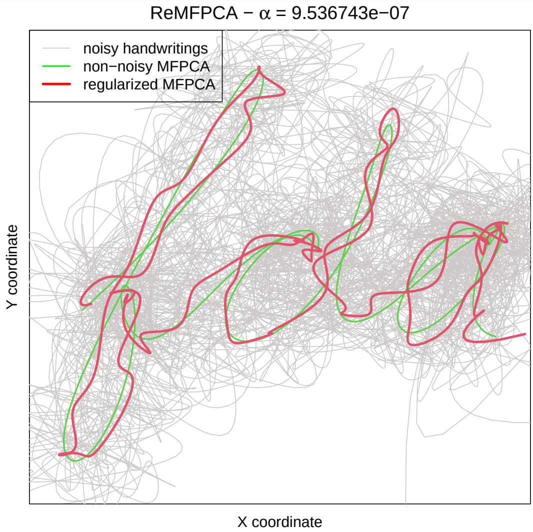
Motivating Example: Cursive Handwriting



Motivating Example: Cursive Handwriting



Motivating Example: Cursive Handwriting



Literature Review

- Rice and Silverman (1991) and Silverman (1996) are pioneer works in regularized FPCA (ReFPCA).
  - Studied functions in Hilbert space and developed roughness penalty based on derivative operators.
  - Mathematical foundation of the ReFPCA is developed in Sobolev spaces.
- Huang et al. (2008) proposed an alternative approach from the penalized SVD point of view.
  - Some nice computational properties.
  - A closed form of CV (GCV) criteria can be derived.
- Chiou et al. (2014) and Happ and Greven (2018) extend FPCA methods to multivariate version: Multivariate FPCA (MFPCA).
  - Enables exploring the relationships between multivariate functions.

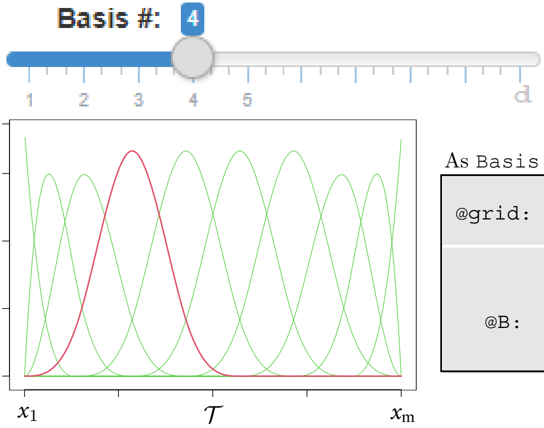
Literature Review

- Rice and Silverman (1991) and Silverman (1996) are pioneer works in regularized FPCA (ReFPCA).
  - Studied functions in Hilbert space and developed roughness penalty based on derivative operators.
  - Mathematical foundation of the ReFPCA is developed in Sobolev spaces.
- Huang et al. (2008) proposed an alternative approach from the penalized SVD point of view.
  - Some nice computational properties.
  - A closed form of CV (GCV) criteria can be derived.
- Chiou et al. (2014) and Happ and Greven (2018) extend FPCA methods to multivariate version: Multivariate FPCA (MFPCA).
  - Enables exploring the relationships between multivariate functions.

Literature Review

- Rice and Silverman (1991) and Silverman (1996) are pioneer works in regularized FPCA (ReFPCA).
  - Studied functions in Hilbert space and developed roughness penalty based on derivative operators.
  - Mathematical foundation of the ReFPCA is developed in Sobolev spaces.
- Huang et al. (2008) proposed an alternative approach from the penalized SVD point of view.
  - Some nice computational properties.
  - A closed form of CV (GCV) criteria can be derived.
- Chiou et al. (2014) and Happ and Greven (2018) extend FPCA methods to multivariate version: Multivariate FPCA (MFPCA).
  - Enables exploring the relationships between multivariate functions.

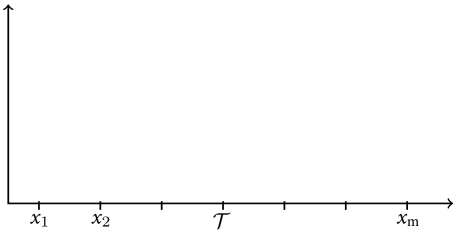
Representation of the functional data (FD) in ReMFPCA



As Basis object:

@grid:	vector	$(x_1, x_2, \ldots, x_{\mathrm{m}})$
@B:	matrix	$\begin{pmatrix} \mathbf{b_{11}} & \cdots & b_{\mathrm{d}1} \\ \vdots & \cdots & \vdots \\ \mathbf{b_{1m}} & \cdots & b_{\mathrm{dm}} \end{pmatrix}$

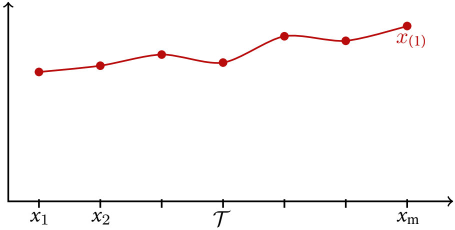
Representation of the functional data (FD) in ReMFPCA



As fd object:

@grid:	vector	$(x_1, x_2, \ldots, x_m)$
@B:	matrix	$\begin{pmatrix} b_{11} & \cdots & b_{d1} \\ \vdots & \cdots & \vdots \\ b_{1m} & \cdots & b_{dm} \end{pmatrix}$
@C:	matrix	$\begin{pmatrix} \\ \end{pmatrix}$

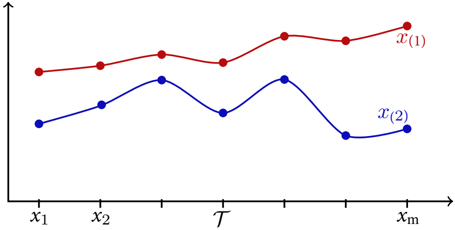
Representation of the functional data (FD) in ReMFPCA



As fd object:

@grid:	vector	$(x_1, x_2, \ldots, x_m)$
@B:	matrix	$\begin{pmatrix} b_{11} & \cdots & b_{d1} \\ \vdots & \cdots & \vdots \\ b_{1m} & \cdots & b_{dm} \end{pmatrix}$
@C:	matrix	$\begin{pmatrix} c_{11} \\ \vdots \\ c_{1d} \end{pmatrix}$

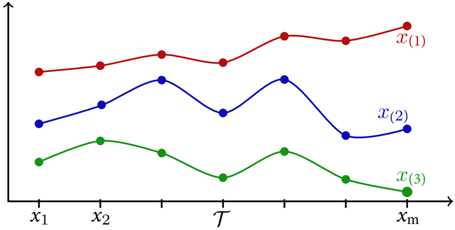
Representation of the functional data (FD) in ReMFPCA



As `fd` object:

`@grid:`	`vector`	$(x_1, x_2, \ldots, x_m)$
`@B:`	`matrix`	$\begin{pmatrix} b_{11} & \cdots & b_{d1} \\ \vdots & \cdots & \vdots \\ b_{1m} & \cdots & b_{dm} \end{pmatrix}$
`@C:`	`matrix`	$\begin{pmatrix} c_{11} & c_{21} & \\ \vdots & \vdots & \\ c_{1d} & c_{2d} & \end{pmatrix}$

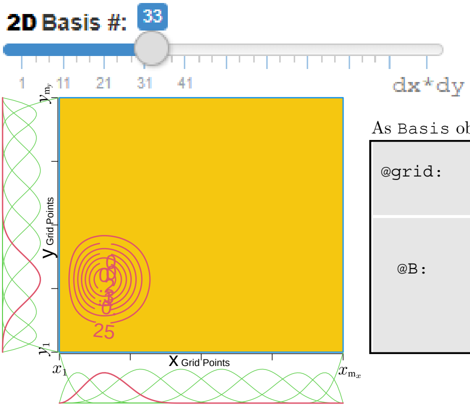
Representation of the functional data (FD) in ReMFPCA



As `fd` object:

`@grid:`	`vector`	$(x_1, x_2, \dots, x_m)$
`@B:`	`matrix`	$\begin{pmatrix} b_{11} & \dots & b_{d1} \\ \vdots & \dots & \vdots \\ b_{1m} & \dots & b_{dm} \end{pmatrix}$
`@C:`	`matrix`	$\begin{pmatrix} c_{11} & c_{21} & c_{31} & \dots \\ \vdots & \vdots & \vdots & \vdots \\ c_{1d} & c_{2d} & c_{3d} & \dots \end{pmatrix}$

Representation of the functional data (FD) in ReMFPCA



As `Basis` object:

`@grid:`	`matrix`	$\begin{pmatrix} x_1, x_2, \ldots, x_{m_x}, x_1, x_2, \ldots, x_{m_x}, x_1, x_2, \ldots, x_{m_x} \\ y_1, y_1, \ldots, y_1\ , y_2, y_2, \ldots, y_2\ , \quad \ldots \quad , y_{m_y} \end{pmatrix}$ $\tilde{m} = m_x * m_y$
`@B:`	`matrix`	$\begin{pmatrix} b_{11} & \ldots & b_{dx,1} & \ldots & b_{đ1} \\ \vdots & \ldots & \vdots & \ldots & \vdots \\ b_{1\tilde{m}} & \ldots & b_{dx,\tilde{m}} & \ldots & b_{đ\tilde{m}} \end{pmatrix}$ $đ = d_x * d_y$

Representation of the functional data (FD) in ReMFPCA

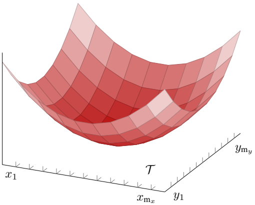
As `fd` object:

`@grid:`	`matrix`	$\begin{pmatrix} x_1, x_2, \ldots, x_{m_x}, x_1, x_2, \ldots, x_{m_x}, x_1, x_2, \ldots, x_{m_x} \\ y_1, y_1, \ldots, y_1\ , y_2, y_2, \ldots, y_2\ , \quad \ldots \quad , y_{m_y} \end{pmatrix}$
`@B:`	`matrix`	$\begin{pmatrix} b_{11} & \ldots & b_{\mathrm{dx},1} & \ldots & b_{đ1} \\ \vdots & \ldots & \vdots & \ldots & \vdots \\ b_{1\tilde{m}} & \ldots & b_{\mathrm{dx},\tilde{m}} & \ldots & b_{đ\tilde{m}} \end{pmatrix}$
`@C:`	`matrix`	$\begin{pmatrix} \quad \end{pmatrix}$

Representation of the functional data (FD) in ReMFPCA

As fd object:

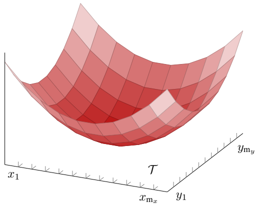
@grid:	matrix	$\begin{pmatrix} x_1, x_2, \dots, x_{m_x}, x_1, x_2, \dots, x_{m_x}, x_1, x_2, \dots, x_{m_x} \\ y_1, y_1, \dots, y_1 \ , y_2, y_2, \dots, y_2 \ , \quad \dots \quad , y_{m_y} \end{pmatrix}$
@B:	matrix	$\begin{pmatrix} b_{11} & \dots & b_{dx,1} & \dots & b_{\text{đ}1} \\ \vdots & \dots & \vdots & \dots & \vdots \\ b_{1\tilde{m}} & \dots & b_{dx,\tilde{m}} & \dots & b_{\text{đ}\tilde{m}} \end{pmatrix}$
@C:	matrix	$\begin{pmatrix} c_{11} \\ \vdots \\ c_{1\text{đ}} \end{pmatrix}$

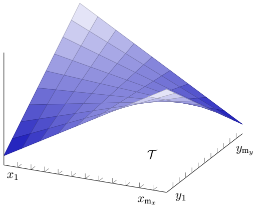


Representation of the functional data (FD) in ReMFPCA

As `fd` object:

`@grid:`	`matrix`	$\begin{pmatrix} x_1, x_2, \ldots, x_{m_x}, x_1, x_2, \ldots, x_{m_x}, x_1, x_2, \ldots, x_{m_x} \\ y_1, y_1, \ldots, y_1\ , y_2, y_2, \ldots, y_2\ , \quad \ldots \quad , y_{m_y} \end{pmatrix}$
`@B:`	`matrix`	$\begin{pmatrix} b_{11} & \ldots & b_{\mathrm{dx},1} & \ldots & b_{\text{đ}1} \\ \vdots & \ldots & \vdots & \ldots & \vdots \\ b_{1\tilde{m}} & \ldots & b_{\mathrm{dx},\tilde{m}} & \ldots & b_{\text{đ}\tilde{m}} \end{pmatrix}$
`@C:`	`matrix`	$\begin{pmatrix} c_{11} & c_{21} \\ \vdots & \vdots \\ c_{1\text{đ}} & c_{2\text{đ}} \end{pmatrix}$

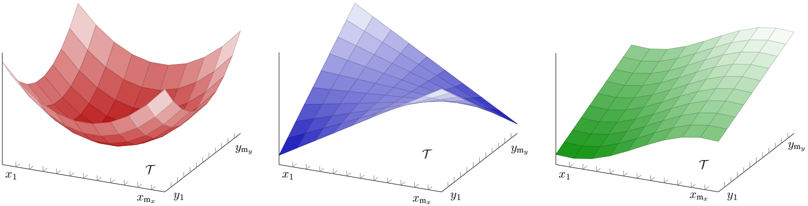




Representation of the functional data (FD) in ReMFPCA

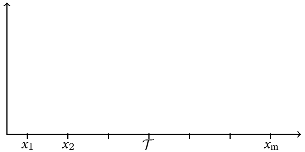
As `fd` object:

`@grid:`	`matrix`	$\begin{pmatrix} x_1, x_2, \dots, x_{m_x}, x_1, x_2, \dots, x_{m_x}, x_1, x_2, \dots, x_{m_x} \\ y_1, y_1, \dots, y_1, y_2, y_2, \dots, y_2, \dots, y_{m_y} \end{pmatrix}$
`@B:`	`matrix`	$\begin{pmatrix} b_{11} & \dots & b_{dx,1} & \dots & b_{\text{đ}1} \\ \vdots & \dots & \vdots & \dots & \vdots \\ b_{1\tilde{m}} & \dots & b_{dx,\tilde{m}} & \dots & b_{\text{đ}\tilde{m}} \end{pmatrix}$
`@C:`	`matrix`	$\begin{pmatrix} c_{11} & c_{21} & c_{31} & \dots \\ \vdots & \vdots & \vdots & \vdots \\ c_{1\text{đ}} & c_{2\text{đ}} & c_{3\text{đ}} & \dots \end{pmatrix}$

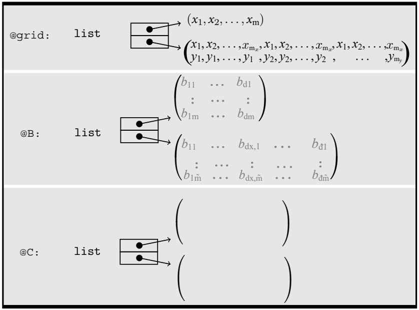


Representation of the functional data (FD) in ReMFPCA

- How about Multivariate Functional Data (MFD) observed over different dimensional domains?

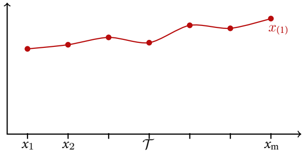


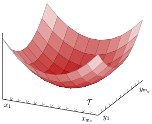
As fd object:



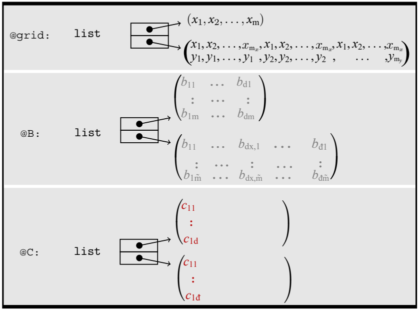
Representation of the functional data (FD) in ReMFPCA

- How about Multivariate Functional Data (MFD) observed over different dimensional domains?



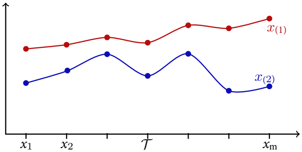


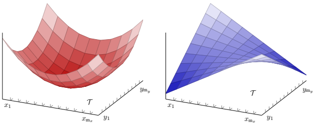
As fd object:



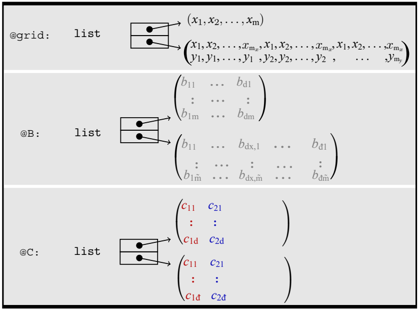
Representation of the functional data (FD) in ReMFPCA

- How about Multivariate Functional Data (MFD) observed over different dimensional domains?



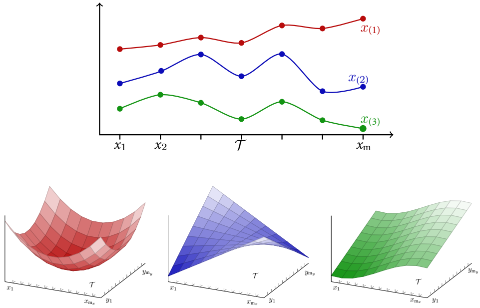


As fd object:

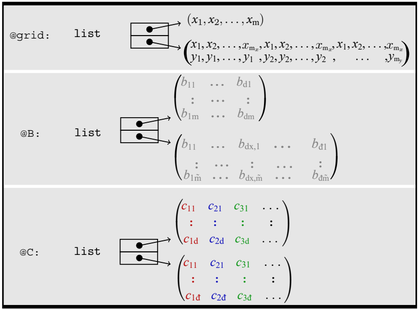


Representation of the functional data (FD) in ReMFPCA

- How about Multivariate Functional Data (MFD) observed over different dimensional domains?



As fd object:



Regularized MFPCA

- Regularized MFPCA (ReMFPCA) seems an intuitive next step to enhance the performance of MFPCA.

- Two ReMFPCA approaches are proposed by our research group.
  - **Regularized Eigen Decomposition of the Covariance Operator**: By extending Silverman (1996) approach into a multivariate framework. (Submitted: `https://doi.org/10.48550/arXiv.2306.13980`)
  - **Penalized Functional SVD (fSVD) of the Data Operator**: We study theoretical foundations and implementation of fSVD for MFPCA. Specifically we extend Huang et al. (2008) approach to the multivariate setup in Sobolev space with
    - Flexibility in tuning parameters selection, and
    - Computation efficiency.

  (Ongoing Project: focus of today's talk)

Regularized MFPCA

- Regularized MFPCA (ReMFPCA) seems an intuitive next step to enhance the performance of MFPCA.
- Two ReMFPCA approaches are proposed by our research group.
  - **Regularized Eigen Decomposition of the Covariance Operator**: By extending Silverman (1996) approach into a multivariate framework. (Submitted: `https://doi.org/10.48550/arXiv.2306.13980`)
  - **Penalized Functional SVD (fSVD) of the Data Operator**: We study theoretical foundations and implementation of fSVD for MFPCA. Specifically we extend Huang et al. (2008) approach to the multivariate setup in Sobolev space with
    - Flexibility in tuning parameters selection, and
    - Computation efficiency.

    (Ongoing Project: focus of today's talk)

Regularized MFPCA

- Regularized MFPCA (ReMFPCA) seems an intuitive next step to enhance the performance of MFPCA.

- Two ReMFPCA approaches are proposed by our research group.
  - **Regularized Eigen Decomposition of the Covariance Operator**: By extending Silverman (1996) approach into a multivariate framework. (Submitted: `https://doi.org/10.48550/arXiv.2306.13980`)
  - **Penalized Functional SVD (fSVD) of the Data Operator**: We study theoretical foundations and implementation of fSVD for MFPCA. Specifically we extend Huang et al. (2008) approach to the multivariate setup in Sobolev space with
    - Flexibility in tuning parameters selection, and
    - Computation efficiency.

  (Ongoing Project: focus of today's talk)

Preliminary Notations

- Let H_j to be a Hilbert space equipped with the inner product

$$\langle x, y\rangle_{H_j} = \int_{\mathcal{T}_j} x(t)y(t)dt, \quad \text{where } x, y \in H_j \text{ and } j = 1, \cdots, p.$$

- The Sobolev space W_j^2 is defined as

$$W_j^2 := \{x(\cdot) : x \text{ and } x' \text{ are absolutely continuous on } \mathcal{T}_j \text{ and } x'' \in H_j\}.$$

- Given a smoothing parameter $\alpha_j > 0$, we can define the inner product

$$\langle x, y\rangle_{\alpha_j} := \langle x, y\rangle_{H_j} + \alpha_j \langle x'', y''\rangle_{H_j}.$$

The α_j-orthogonality in Sobolev space W_j^2 is

$$\langle x_j, y_j\rangle_{\alpha_j} = 0.$$

Preliminary Notations

- Let H_j to be a Hilbert space equipped with the inner product

$$\langle x, y\rangle_{H_j} = \int_{\mathcal{T}_j} x(t)y(t)dt, \quad \text{where } x, y \in H_j \text{ and } j = 1, \cdots, p.$$

- The Sobolev space W_j^2 is defined as

$$W_j^2 := \{x(\cdot) : x \textit{ and } x' \textit{ are absolutely continuous on } \mathcal{T}_j \textit{ and } x'' \in H_j\}.$$

- Given a smoothing parameter $\alpha_j > 0$, we can define the inner product

$$\langle x, y\rangle_{\alpha_j} := \langle x, y\rangle_{H_j} + \alpha_j \langle x'', y''\rangle_{H_j}.$$

The α_j-orthogonality in Sobolev space W_j^2 is

$$\langle x_j, y_j\rangle_{\alpha_j} = 0.$$

Preliminary Notations cont.

- Define the cartesian Hilbert product space $\mathbb{H} := H_1 \times \cdots \times H_p$, where each H_j is a Hilbert space.
  For $\boldsymbol{x} = (x_1, \cdots, x_p)$ and $\boldsymbol{y} = (y_1, \cdots, y_p) \in \mathbb{H}$,

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle_{\mathbb{H}} = \sum_{j=1}^{p} \langle x_j, y_j \rangle_{H_j}.$$

- Define the cartesian Sobolev product spaces $\mathbb{W}^2 := W_1^2 \times \cdots \times W_p^2$.
  For $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{W}^2$ and smoothing parameter $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \cdots, \alpha_p) \in \mathbb{R}^p$,

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle_{\boldsymbol{\alpha}} = \sum_{j=1}^{p} \langle x_j, y_j \rangle_{\alpha_j}.$$

The $\boldsymbol{\alpha}$-orthogonality in Sobolev space $\mathbb{W}^2$ is $\langle \boldsymbol{x}, \boldsymbol{y} \rangle_{\boldsymbol{\alpha}} = 0$.

Preliminary Notations cont.

- Define the cartesian Hilbert product space $\mathbb{H} := H_1 \times \cdots \times H_p$, where each H_j is a Hilbert space.
  For $\boldsymbol{x} = (x_1, \cdots, x_p)$ and $\boldsymbol{y} = (y_1, \cdots, y_p) \in \mathbb{H}$,

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle_{\mathbb{H}} = \sum_{j=1}^{p} \langle x_j, y_j \rangle_{H_j}.$$

- Define the cartesian Sobolev product spaces $\mathbb{W}^2 := W_1^2 \times \cdots \times W_p^2$.
  For $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{W}^2$ and smoothing parameter $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \cdots, \alpha_p) \in \mathbb{R}^p$,

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle_{\boldsymbol{\alpha}} = \sum_{j=1}^{p} \langle x_j, y_j \rangle_{\alpha_j}.$$

  The $\boldsymbol{\alpha}$-orthogonality in Sobolev space $\mathbb{W}^2$ is $\langle \boldsymbol{x}, \boldsymbol{y} \rangle_{\boldsymbol{\alpha}} = 0$.

Functional SVD and penalized functional SVD

Theorem

Denote $\boldsymbol{x}_i := [x_{i,j}]_{j=1}^{p} \in \mathbb{H}$ and the data operator $\boldsymbol{\mathcal{X}} := [\boldsymbol{x}_i]_{i=1}^{n} \in \mathbb{F}^{p\times n}$ with rank $m \leq n$. There exist linearly independent elements $\boldsymbol{\varphi}_1, \cdots, \boldsymbol{\varphi}_m$ from $\mathbb{H}$ and $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_m$ from $\mathbb{R}^n$ that are orthonormal and

$$\boldsymbol{\mathcal{X}} = \sum_{i=1}^{m} \sqrt{\lambda_i}\, \boldsymbol{v}_i \otimes \boldsymbol{\varphi}_i, \tag{1}$$

where λ_i's are non-ascending positive scalars.

The goal is to obtain regularized FPCs, which is equivalent to solve the following penalized functional SVD problem:

$$\min_{\boldsymbol{\varphi}:\|\boldsymbol{\varphi}\|_{\boldsymbol{\alpha}}=1,\ \boldsymbol{v}\in\mathbb{R}^n} \|\boldsymbol{\mathcal{X}} - \boldsymbol{v}\otimes\boldsymbol{\varphi}\|_{\mathbb{F}}^2 + \boldsymbol{v}^\top\boldsymbol{v}\sum_{j=1}^{p}\alpha_j\langle\varphi_j'', \varphi_j''\rangle_{H_j} \tag{2}$$

Functional SVD and penalized functional SVD

Theorem

Denote $\boldsymbol{x}_i := [x_{i,j}]_{j=1}^{p} \in \mathbb{H}$ and the data operator $\boldsymbol{\mathcal{X}} := [\boldsymbol{x}_i]_{i=1}^{n} \in \mathbb{F}^{p\times n}$ with rank $m \leq n$. There exist linearly independent elements $\boldsymbol{\varphi}_1, \cdots, \boldsymbol{\varphi}_m$ from $\mathbb{H}$ and $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_m$ from $\mathbb{R}^n$ that are orthonormal and

$$\boldsymbol{\mathcal{X}} = \sum_{i=1}^{m} \sqrt{\lambda_i}\, \boldsymbol{v}_i \otimes \boldsymbol{\varphi}_i, \tag{1}$$

where λ_i's are non-ascending positive scalars.

The goal is to obtain regularized FPCs, which is equivalent to solve the following penalized functional SVD problem:

$$\min_{\boldsymbol{\varphi}:\|\boldsymbol{\varphi}\|_{\boldsymbol{\alpha}}=1,\ \boldsymbol{v}\in\mathbb{R}^n} \|\boldsymbol{\mathcal{X}} - \boldsymbol{v}\otimes\boldsymbol{\varphi}\|_{\mathbb{F}}^2 + \boldsymbol{v}^\top\boldsymbol{v}\sum_{j=1}^{p} \alpha_j \langle \varphi_j'', \varphi_j''\rangle_{H_j} \tag{2}$$

Finite dimensional representation of functional data

- In implementation, each functional observations are considered as projection on a finite dimensional subspace $H_j^{d_j} = sp\{v_j^k\}_{k=1}^{d_j} \subseteq H_j$. And we define $\mathbb{H}^{\boldsymbol{d}} := H_1^{d_1} \times \cdots \times H_p^{d_p}$.
- The minimization problem given in (2) becomes

$$\min_{\underset{\sim}{\boldsymbol{b}},\boldsymbol{v}} \|\underset{\sim}{\boldsymbol{B}} - \boldsymbol{v}\underset{\sim}{\boldsymbol{b}}^\top\|_F^2 + \boldsymbol{v}^\top \boldsymbol{v}\underset{\sim}{\boldsymbol{b}}^\top \boldsymbol{\Omega_\alpha} \underset{\sim}{\boldsymbol{b}}, \quad (3)$$

where $\underset{\sim}{\boldsymbol{B}} = \boldsymbol{B}\boldsymbol{G}^{\frac{1}{2}}$, $\underset{\sim}{\boldsymbol{b}} = \boldsymbol{G}^{\frac{1}{2}}\boldsymbol{b}$, $\boldsymbol{\Omega_\alpha} = \boldsymbol{G}^{-\frac{1}{2}}\boldsymbol{D_\alpha}\boldsymbol{G}^{-\frac{1}{2}}$, and

- $\boldsymbol{B}$ is the matrix associated to the projection coefficients of $\boldsymbol{\mathcal{X}}$ on $\mathbb{H}^{\boldsymbol{d}}$,
- $\boldsymbol{b}$ is the vector corresponding to the projection coefficients of φ on $\mathbb{H}^{\boldsymbol{d}}$,
- $\boldsymbol{G} := \text{diag}\,\{\boldsymbol{G}_1, \cdots, \boldsymbol{G}_p\}$, where $\boldsymbol{G}_j = \left[\langle v_j^l, v_j^k\rangle_{H_j}\right]_{l,k=1}^{d_j}$,
- $\boldsymbol{D_\alpha} := \text{diag}\{\alpha_1\boldsymbol{D}_1, \cdots, \alpha_p\boldsymbol{D}_p\}$, where $\boldsymbol{D}_j = \left[\langle v_j^{l\prime\prime}, v_j^{k\prime\prime}\rangle_{H_j}\right]_{l,k=1}^{d_j}$.

Finite dimensional representation of functional data

- In implementation, each functional observations are considered as projection on a finite dimensional subspace $H_j^{d_j} = sp\{v_j^k\}_{k=1}^{d_j} \subseteq H_j$. And we define $\mathbb{H}^{\boldsymbol{d}} := H_1^{d_1} \times \cdots \times H_p^{d_p}$.
- The minimization problem given in (2) becomes

$$\min_{\underset{\sim}{\boldsymbol{b}},\boldsymbol{v}} \|\underset{\sim}{\boldsymbol{B}} - \boldsymbol{v}\underset{\sim}{\boldsymbol{b}}^\top\|_F^2 + \boldsymbol{v}^\top\boldsymbol{v}\underset{\sim}{\boldsymbol{b}}^\top\boldsymbol{\Omega_\alpha}\underset{\sim}{\boldsymbol{b}}, \tag{3}$$

where $\underset{\sim}{\boldsymbol{B}} = \boldsymbol{B}\boldsymbol{G}^{\frac{1}{2}}$, $\underset{\sim}{\boldsymbol{b}} = \boldsymbol{G}^{\frac{1}{2}}\boldsymbol{b}$, $\boldsymbol{\Omega_\alpha} = \boldsymbol{G}^{-\frac{1}{2}}\boldsymbol{D_\alpha}\boldsymbol{G}^{-\frac{1}{2}}$, and

- $\boldsymbol{B}$ is the matrix associated to the projection coefficients of $\boldsymbol{\mathcal{X}}$ on $\mathbb{H}^{\boldsymbol{d}}$,
- $\boldsymbol{G} := \text{diag}\ \{\boldsymbol{G}_1, \cdots, \boldsymbol{G}_p\}$, where $\boldsymbol{G}_j = \left[\langle v_j^l, v_j^k\rangle_{H_j}\right]_{l,k=1}^{d_j}$,
- $\boldsymbol{b}$ is the vector corresponding to the projection coefficients of $\boldsymbol{\varphi}$ on $\mathbb{H}^{\boldsymbol{d}}$,
- $\boldsymbol{D_\alpha} := \text{diag}\{\alpha_1\boldsymbol{D}_1, \cdots, \alpha_p\boldsymbol{D}_p\}$, where $\boldsymbol{D}_j = \left[\langle v_j^{l\prime\prime}, v_j^{k\prime\prime}\rangle_{H_j}\right]_{l,k=1}^{d_j}$.

Implementation Strategy: Power algorithm

- To optimize (3), one may use the following iterative power algorithm:
  1. Initialize $\boldsymbol{b}$.
  2. Repeat until convergence:
     a. $\boldsymbol{v} \leftarrow \boldsymbol{BGb}$,
     b. $\boldsymbol{b} \leftarrow \mathrm{S}_\alpha^2 \boldsymbol{G} \boldsymbol{B}^\top \boldsymbol{v}$,
     c. normalize $\boldsymbol{b}$.

  Here $\mathrm{S}_\alpha = (\boldsymbol{G} + \boldsymbol{D}_\alpha)^{-\frac{1}{2}}$ is referred to a half-smoothing matrix.

- For a fixed $\boldsymbol{v}$, the penalized SVD in (3) becomes a penalized regression problem:

$$\|\bar{\boldsymbol{y}} - \bar{\boldsymbol{X}}\underset{\sim}{\boldsymbol{b}}\|^2 + \underset{\sim}{\boldsymbol{b}}^\top(\boldsymbol{v}^\top\boldsymbol{v}\boldsymbol{\Omega}_\alpha)\underset{\sim}{\boldsymbol{b}}, \tag{4}$$

where

$$\bar{\boldsymbol{y}} := \left[\underset{\sim}{\boldsymbol{B}}_{.1}^\top, \underset{\sim}{\boldsymbol{B}}_{.2}^\top, \ldots, \underset{\sim}{\boldsymbol{B}}_{.d}^\top\right]^\top \in \mathbb{R}^{nd}, \quad \bar{\boldsymbol{X}} := \begin{bmatrix} \boldsymbol{v} & & \\ & \ddots & \\ & & \boldsymbol{v} \end{bmatrix} \in \mathbb{R}^{nd \times d}.$$

Implementation Strategy: Power algorithm

- To optimize (3), one may use the following iterative power algorithm:
  1. Initialize $\boldsymbol{b}$.
  2. Repeat until convergence:
     a. $\boldsymbol{v} \leftarrow \boldsymbol{B}\boldsymbol{G}\boldsymbol{b}$,
     b. $\boldsymbol{b} \leftarrow \mathbf{S}_{\boldsymbol{\alpha}}^{2}\boldsymbol{G}\boldsymbol{B}^{\top}\boldsymbol{v}$,
     c. normalize $\boldsymbol{b}$.

  Here $\mathbf{S}_{\boldsymbol{\alpha}} = (\boldsymbol{G} + \boldsymbol{D}_{\boldsymbol{\alpha}})^{-\frac{1}{2}}$ is referred to a half-smoothing matrix.

- For a fixed v, the penalized SVD in (3) becomes a penalized regression problem:

$$\|\bar{y} - \bar{X}\underset{\sim}{b}\|^2 + \underset{\sim}{b}^\top(v^\top v\Omega_\alpha)\underset{\sim}{b}, \quad (4)$$

where

$$\bar{y} := \left[\underset{\sim}{B}_{.1}^\top, \underset{\sim}{B}_{.2}^\top, \ldots, \underset{\sim}{B}_{.d}^\top\right]^\top \in \mathbb{R}^{nd}, \quad \bar{X} := \begin{bmatrix} v & & \\ & \ddots & \\ & & v \end{bmatrix} \in \mathbb{R}^{nd\times d}.$$

Implementation Strategy: Power algorithm

- To optimize (3), one may use the following iterative power algorithm:
  1. Initialize $\boldsymbol{b}$.
  2. Repeat until convergence:
     - (a) $\boldsymbol{v} \leftarrow \boldsymbol{B}\boldsymbol{G}\boldsymbol{b}$,
     - (b) $\boldsymbol{b} \leftarrow \mathbf{S}_{\boldsymbol{\alpha}}^{2}\boldsymbol{G}\boldsymbol{B}^{\top}\boldsymbol{v}$,
     - (c) normalize $\boldsymbol{b}$.

  Here $\mathbf{S}_{\boldsymbol{\alpha}} = (\boldsymbol{G} + \boldsymbol{D}_{\boldsymbol{\alpha}})^{-\frac{1}{2}}$ is referred to a half-smoothing matrix.

- For a fixed $\boldsymbol{v}$, the penalized SVD in (3) becomes a penalized regression problem:

$$\|\bar{\boldsymbol{y}} - \bar{\boldsymbol{X}}\underset{\sim}{\boldsymbol{b}}\|^2 + \underset{\sim}{\boldsymbol{b}}^{\top}(\boldsymbol{v}^{\top}\boldsymbol{v}\boldsymbol{\Omega}_{\boldsymbol{\alpha}})\underset{\sim}{\boldsymbol{b}}, \quad (4)$$

  where

$$\bar{\boldsymbol{y}} := \left[\underset{\sim}{\boldsymbol{B}}_{.1}^{\top}, \underset{\sim}{\boldsymbol{B}}_{.2}^{\top}, \dots, \underset{\sim}{\boldsymbol{B}}_{.d}^{\top}\right]^{\top} \in \mathbb{R}^{nd}, \ \bar{\boldsymbol{X}} := \begin{bmatrix} \boldsymbol{v} & & \\ & \ddots & \\ & & \boldsymbol{v} \end{bmatrix} \in \mathbb{R}^{nd \times d}.$$

Tuning parameters selection based on GCV

The GCV criteria can be simply nested within the power algorithm

$$GCV_\alpha = \frac{1}{d}\sum_{k=1}^{p} \frac{\|(\boldsymbol{I}_k - \tilde{\boldsymbol{S}}_{\alpha_k})(\tilde{\boldsymbol{B}}_k^\top \boldsymbol{v})\|^2}{(1 - \frac{1}{d}tr\{\tilde{\boldsymbol{S}}_{\alpha_k}\})^2},$$

where $\tilde{S}_{\alpha_k}$ is k^{th} diagonal block of $\tilde{\boldsymbol{S}}_\alpha := \boldsymbol{G}^{\frac{1}{2}}\boldsymbol{S}_\alpha^2\boldsymbol{G}^{\frac{1}{2}}$

a) $\boldsymbol{v} \leftarrow \boldsymbol{BGb}$

b) $\boldsymbol{b} \leftarrow \boldsymbol{S}_\alpha^2\boldsymbol{GB}^T\boldsymbol{v}$

c) Normalize b

Simply nest GCV selection of $\boldsymbol{\alpha}$ inside step (b)

Two flexible choices in power algorithm

- Simultaneous power algorithm:
  - Obtaining FPCs jointly where all FPCs share the same tuning parameter.
  - Preserves the α-orthogonality in Sobolev space.
  - Since we compute the $(p > 1)$-dimensional subspace simultaneously, a QR factorization is needed in step 2(b).
- Sequential power algorithm:
  - Obtaining FPCs sequentially where different tuning parameter is allowed for each FPC.
  - The flexibility of having different level of smoothness for FPCs.
  - Losing the α-orthogonality property in Sobolev space.

Two flexible choices in power algorithm

- Simultaneous power algorithm:
    - Obtaining FPCs jointly where all FPCs share the same tuning parameter.
    - Preserves the α-orthogonality in Sobolev space.
    - Since we compute the $(p > 1)$-dimensional subspace simultaneously, a QR factorization is needed in step 2(b).

- Sequential power algorithm:
    - Obtaining FPCs sequentially where different tuning parameter is allowed for each FPC.
    - The flexibility of having different level of smoothness for FPCs.
    - Losing the α-orthogonality property in Sobolev space.

Simulation setup

- Let $\mathbf{X}(\boldsymbol{t})$ be a bivariate functional observation. We define a bivariate orthonormal basis system $\boldsymbol{\psi}_m(\boldsymbol{t})$, where

$$\psi_m^{(1)}(t) = \sin\left((2m-1)\pi t\right) \quad \text{and} \quad \psi_m^{(2)}(t) = \sin\left(\frac{(4m-3)\pi}{2} t\right).$$

We adopt the following functional data generating model:

$$\boldsymbol{X}_i(\boldsymbol{t}) = \sum_{m=1}^{M} \rho_{i,m} \boldsymbol{\psi}_m(\boldsymbol{t}), \qquad \rho_{i,m} \sim \mathbb{N}(0, \lambda_m), \quad i = 1, \dots, n. \tag{5}$$

- The goal is to examine scenarios where varying levels of noise are added to each $\boldsymbol{\psi}_m(\boldsymbol{t})$, where $\tilde{\boldsymbol{\psi}}_m(\boldsymbol{t}) = \boldsymbol{\psi}_m(\boldsymbol{t}) + \boldsymbol{\epsilon}_m(\boldsymbol{t})$.

- We simulate our observations, using

$$\boldsymbol{Y}_i(\boldsymbol{t}) = \sum_{m=1}^{M} \rho_{i,m}(\tilde{\boldsymbol{\psi}}_m(\boldsymbol{t})), \tag{6}$$

Simulation setup

- Let $\mathbf{X}(\boldsymbol{t})$ be a bivariate functional observation. We define a bivariate orthonormal basis system $\boldsymbol{\psi}_m(\boldsymbol{t})$, where

$$\psi_m^{(1)}(t) = \sin\left((2m-1)\pi t\right) \quad \text{and} \quad \psi_m^{(2)}(t) = \sin\left(\frac{(4m-3)\pi}{2} t\right).$$

We adopt the following functional data generating model:

$$\boldsymbol{X}_i(\boldsymbol{t}) = \sum_{m=1}^{M} \rho_{i,m} \boldsymbol{\psi}_m(\boldsymbol{t}), \qquad \rho_{i,m} \sim \mathbb{N}(0, \lambda_m), \quad i = 1, \ldots, n. \quad (5)$$

- The goal is to examine scenarios where varying levels of noise are added to each $\boldsymbol{\psi}_m(\boldsymbol{t})$, where $\tilde{\boldsymbol{\psi}}_m(\boldsymbol{t}) = \boldsymbol{\psi}_m(\boldsymbol{t}) + \boldsymbol{\epsilon}_m(\boldsymbol{t})$.

- We simulate our observations, using

$$\boldsymbol{Y}_i(\boldsymbol{t}) = \sum_{m=1}^{M} \rho_{i,m}(\tilde{\boldsymbol{\psi}}_m(\boldsymbol{t})), \quad (6)$$

Simulation setup

- Let $\mathbf{X}(\boldsymbol{t})$ be a bivariate functional observation. We define a bivariate orthonormal basis system $\boldsymbol{\psi}_m(\boldsymbol{t})$, where

$$\psi_m^{(1)}(t) = \sin\left((2m-1)\pi t\right) \quad \text{and} \quad \psi_m^{(2)}(t) = \sin\left(\frac{(4m-3)\pi}{2}t\right).$$

We adopt the following functional data generating model:

$$\boldsymbol{X}_i(\boldsymbol{t}) = \sum_{m=1}^{M} \rho_{i,m}\boldsymbol{\psi}_m(\boldsymbol{t}), \qquad \rho_{i,m} \sim \mathbb{N}(0, \lambda_m), \quad i = 1, \ldots, n. \quad (5)$$

- The goal is to examine scenarios where varying levels of noise are added to each $\boldsymbol{\psi}_m(\boldsymbol{t})$, where $\tilde{\boldsymbol{\psi}}_m(\boldsymbol{t}) = \boldsymbol{\psi}_m(\boldsymbol{t}) + \boldsymbol{\epsilon}_m(\boldsymbol{t})$.

- We simulate our observations, using

$$\boldsymbol{Y}_i(\boldsymbol{t}) = \sum_{m=1}^{M} \rho_{i,m}(\tilde{\boldsymbol{\psi}}_m(\boldsymbol{t})), \quad (6)$$

Simulation: Comparison

- To assess the performance of our sequential and joint approach, we compare them with two other methods: non-regularized MFPCA and Happ's approach (Happ and Greven, 2018).

- The accuracy of the estimated eigenvalue and eigenfunction pairs, denoted as $\hat{\lambda}_m$ and $\hat{\psi}_m$ respectively, was evaluated by comparing them to their original counterparts:

$$Err(\hat{\lambda}_m) = |\hat{\lambda}_m - \lambda_m| / |\lambda_m| \qquad \text{and} \qquad Err(\hat{\psi}_m) = \|\hat{\psi}_m - \psi\|_{\mathbb{H}}.$$

- Furthermore, the accuracy of the estimates for each replication is assessed using the mean relative absolute error (MRAE), defined as

$$\text{MRAE} = \frac{1}{n} \sum_{i=1}^{n} (\|\hat{x}_i - x_i\|_{\mathbb{H}}) / \|x_i\|_{\mathbb{H}},$$

where $\hat{x}_i = \sum_{m=1}^{J} \langle y_i, \hat{\psi}_m \rangle_{\mathbb{H}} \hat{\psi}_m$.

Simulation: Comparison

- To assess the performance of our sequential and joint approach, we compare them with two other methods: non-regularized MFPCA and Happ's approach (Happ and Greven, 2018).

- The accuracy of the estimated eigenvalue and eigenfunction pairs, denoted as $\hat{\lambda}_m$ and $\hat{\boldsymbol{\psi}}_m$ respectively, was evaluated by comparing them to their original counterparts:

$$Err(\hat{\lambda}_m) = |\hat{\lambda}_m - \lambda_m|/|\lambda_m| \qquad \text{and} \qquad Err(\hat{\boldsymbol{\psi}}_m) = \|\hat{\boldsymbol{\psi}}_m - \boldsymbol{\psi}\|_{\mathbb{H}}.$$

- Furthermore, the accuracy of the estimates for each replication is assessed using the mean relative absolute error (MRAE), defined as

$$\text{MRAE} = \frac{1}{n}\sum_{i=1}^{n}(\|\hat{\boldsymbol{x}}_i - \boldsymbol{x}_i\|_{\mathbb{H}})/\|\boldsymbol{x}_i\|_{\mathbb{H}},$$

where $\hat{\boldsymbol{x}}_i = \sum_{m=1}^{J}\langle \boldsymbol{y}_i, \hat{\boldsymbol{\psi}}_m\rangle_{\mathbb{H}}\hat{\boldsymbol{\psi}}_m$.

Simulation: Comparison

- To assess the performance of our sequential and joint approach, we compare them with two other methods: non-regularized MFPCA and Happ's approach (Happ and Greven, 2018).

- The accuracy of the estimated eigenvalue and eigenfunction pairs, denoted as $\hat{\lambda}_m$ and $\hat{\boldsymbol{\psi}}_m$ respectively, was evaluated by comparing them to their original counterparts:

$$Err(\hat{\lambda}_m) = |\hat{\lambda}_m - \lambda_m| / |\lambda_m| \qquad \text{and} \qquad Err(\hat{\boldsymbol{\psi}}_m) = \|\hat{\boldsymbol{\psi}}_m - \boldsymbol{\psi}\|_{\mathbb{H}}.$$

- Furthermore, the accuracy of the estimates for each replication is assessed using the mean relative absolute error (MRAE), defined as

$$\text{MRAE} = \frac{1}{n} \sum_{i=1}^{n} (\|\hat{\boldsymbol{x}}_i - \boldsymbol{x}_i\|_{\mathbb{H}}) / \|\boldsymbol{x}_i\|_{\mathbb{H}},$$

where $\hat{\boldsymbol{x}}_i = \sum_{m=1}^{J} \langle \boldsymbol{y}_i, \hat{\boldsymbol{\psi}}_m \rangle_{\mathbb{H}} \hat{\boldsymbol{\psi}}_m$.

Comparison result (for different trend patterns in eigenvalues)

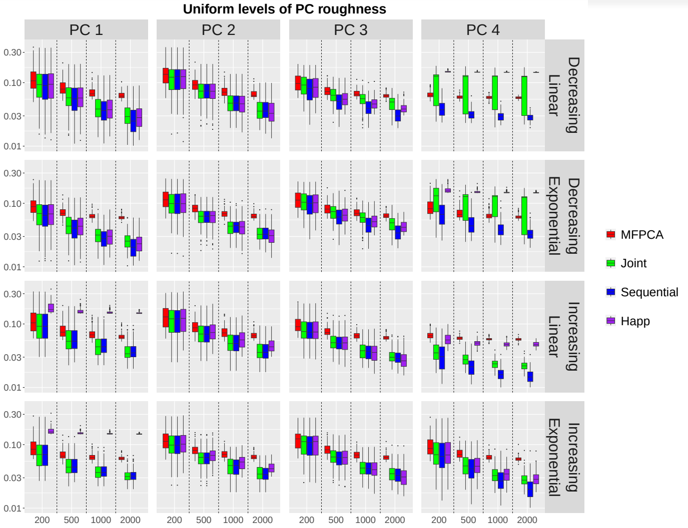


Figure: $Err(\hat{\psi}_m)$

Comparison result (for different trend patterns in eigenvalues)

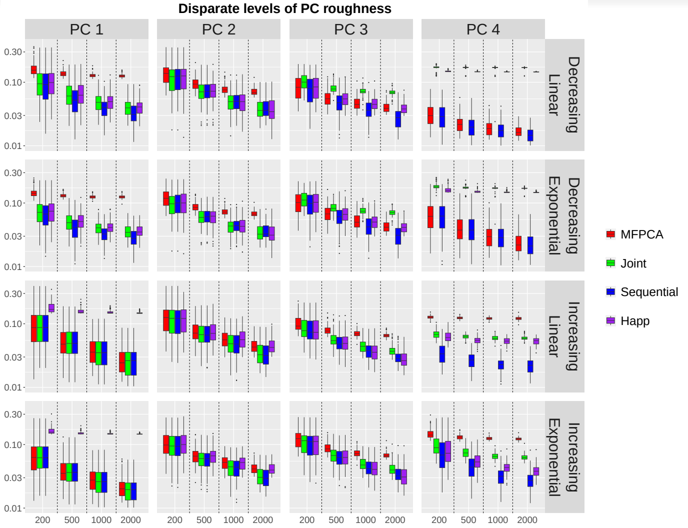


Figure: $Err(\hat{\boldsymbol{\psi}}_m)$

Comparison result (for different trend patterns in eigenvalues)

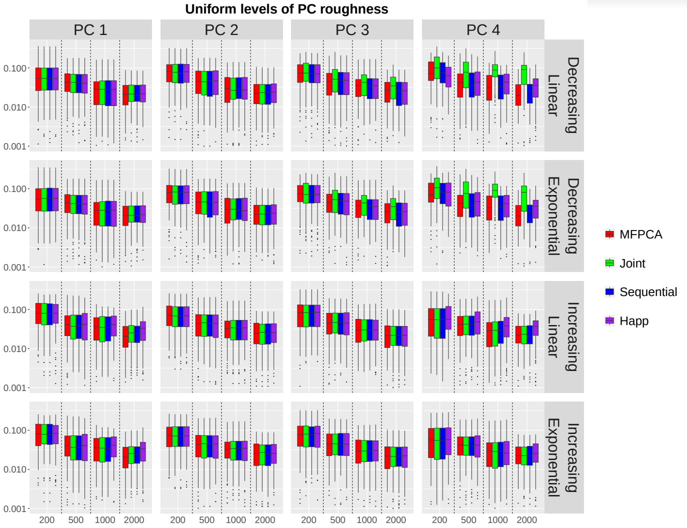


Figure: $Err(\hat{\lambda}_m)$

Comparison result (for different trend patterns in eigenvalues)

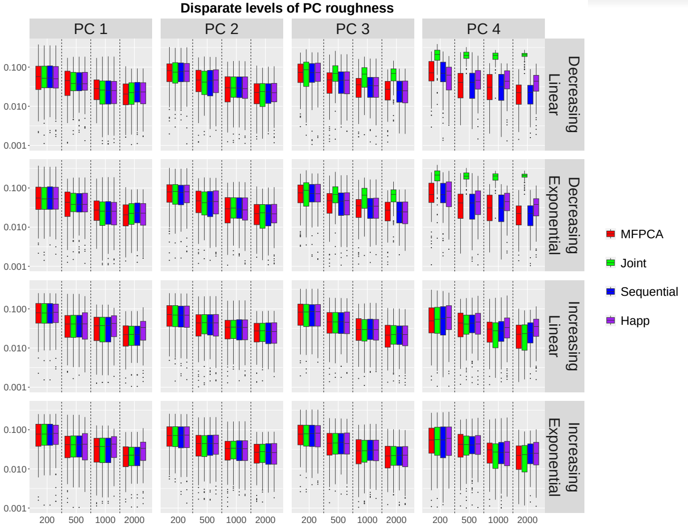


Figure: $Err(\hat{\lambda}_m)$

Comparison result (for different trend patterns in eigenvalues)

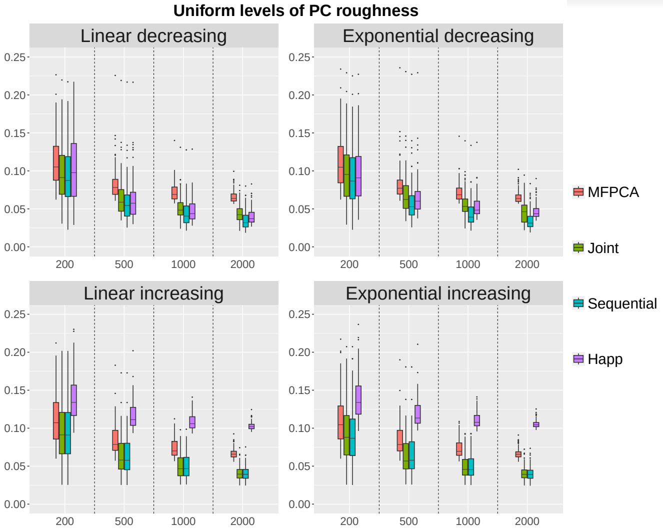


Figure: MRAE

Comparison result (for different trend patterns in eigenvalues)

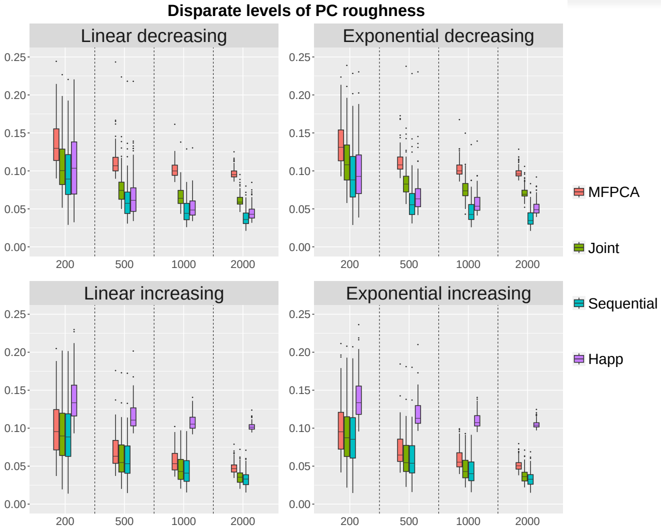
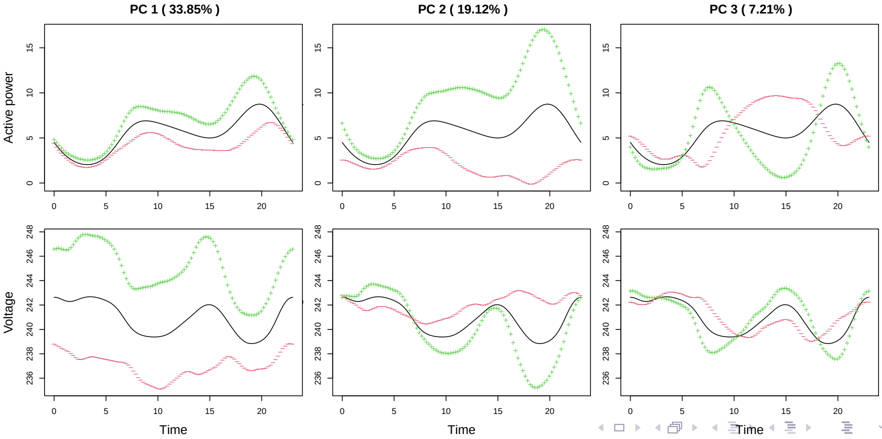


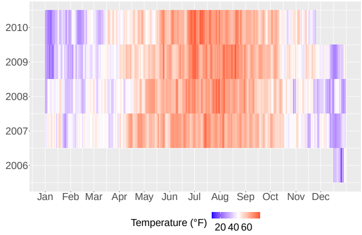
Figure: MRAE

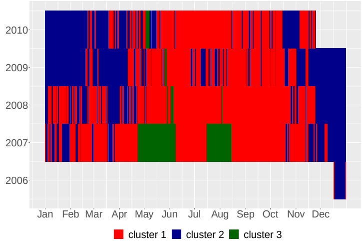
Household electric power consumption data

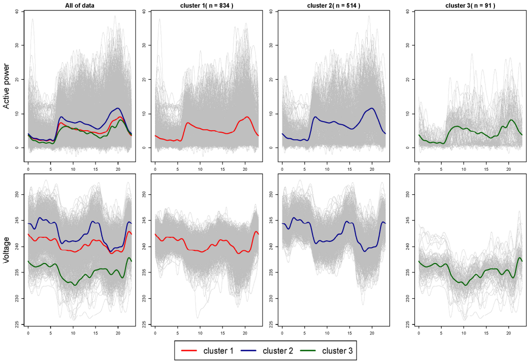
- Consider a bivariate functional data that include active power and voltage consumption of one household in Sceaux (7km of Paris, France) between December 2006 and November 2010.



Interpretation of PC scores: FPC1







Left top: Average temperature heatmap;
Left bottom: Clustering based on PC1 scores;
Right: Clustering details on original data.

Interpretation of PC scores: FPC3

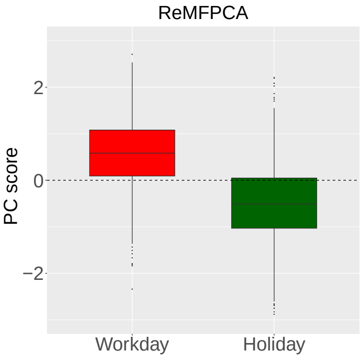


Figure: Boxplot of PC3 scores

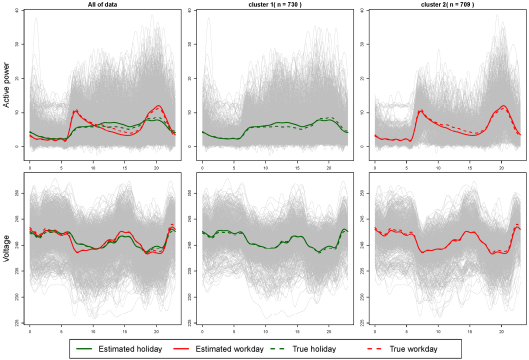


Figure: Clustering details on original data.

Conclusion

- We developed ReMFPCA based on regularized functional SVD approach.
- An efficient power algorithm is proposed with two flexible choices:
  - Simultaneous power method: Jointly estimates all FPCs with a common smoothing parameter (FPCs will have the α-orthogonality in Sobolev space).
  - Sequential power method: Estimating each FPC sequentially, where different smoothing parameters are allowed for each FPC (we will lose the α-orthogonal property).
- A closed form GCV is derived from the regularized functional SVD approach, where it can significantly improve computational efficiency.
  - Proposed GCV criteria can be embedded within the power algorithm.

Thank you!

- Collaborators
  - Yue Zhao, Ph.D. Candidate, Marquette University
  - Hossein Haghbin, Assistant Professor, Persian Gulf University

Questions?

References

Chiou, J.-M., Chen, Y.-T., and Yang, Y.-F. (2014). Multivariate functional principal component analysis: A normalization approach. *Statistica Sinica*, pages 1571–1596.

Happ, C. and Greven, S. (2018). Multivariate functional principal component analysis for data observed on different (dimensional) domains. *Journal of the American Statistical Association*, 113(522):649–659.

Huang, J. Z., Shen, H., Buja, A., et al. (2008). Functional principal components analysis via penalized rank one approximation. *Electronic Journal of Statistics*, 2:678–695.

Rice, J. A. and Silverman, B. W. (1991). Estimating the mean and covariance structure nonparametrically when the data are curves. *Journal of the Royal Statistical Society: Series B (Methodological)*, 53(1):233–243.

Silverman, B. W. (1996). Smoothed functional principal components analysis by choice of norm. *The Annals of Statistics*, 24(1):1–24.

Zhao, Y., Haghbin, H., and Maadooliat, M. (tobeSubmitted). Regularized multivariate functional principal component analysis. *arXiv*.