Functional Data Analysis: Part III
Inverse Problem in FDA

Jane-Ling Wang and Hans-Georg Müller

University of California Davis

August 26-27, 2016

Outline

Outline

Inverse Problem with the Covariance Operator

- A major challenge in FDA is the inverse problem, which stems from the inversion of the covariance operator.

  \- The covariance operator is compact, hence its inverse is not a bounded operator, whence the complication.

Inverse Problem with the Covariance Operator

- A major challenge in FDA is the inverse problem, which stems from the inversion of the covariance operator.

  - The covariance operator is compact, hence its inverse is not a bounded operator, whence the complication.

- We illustrate this through two examples where such an inverse problem occurs.

Inverse Problem with the Covariance Operator

- A major challenge in FDA is the inverse problem, which stems from the inversion of the covariance operator.

  \- The covariance operator is compact, hence its inverse is not a bounded operator, whence the complication.

- We illustrate this through two examples where such an inverse problem occurs.

  (i) Functional canonical correlation analysis (FCCA)

  (ii) Functional linear models.

Outline

Functional Correlation

Let $(X(t), Y(t))$ be a pair of functional data.

How do we extend the concept of correlation to functional data?

Functional Correlation

Let $(X(t), Y(t))$ be a pair of functional data.

How do we extend the concept of correlation to functional data?

- The first attempt by Leurgans et al. (1993) is to extend the canonical correlation for multivariate data to functional data.

Functional Correlation

Let $(X(t), Y(t))$ be a pair of functional data.

How do we extend the concept of correlation to functional data?

- The first attempt by Leurgans et al. (1993) is to extend the canonical correlation for multivariate data to functional data.
- For p-dimensional multivariate data (X, Y), there are p canonical correlations.

Review of Multivariate Canonical Correlation

- The first canonical correlation ρ_1 is the maximal Pearson correlation between any two linear directions $\langle \alpha, X \rangle$ and $\langle \beta, Y \rangle$,

Review of Multivariate Canonical Correlation

- The first canonical correlation ρ_1 is the maximal Pearson correlation between any two linear directions $\langle \alpha, X \rangle$ and $\langle \beta, Y \rangle$,

  $\Longrightarrow$ The first canonical correlation ρ_1 and its associated weights (α_1, β_1) are defined as follows:

$$\rho_1 = \sup_{\alpha, \beta \in \mathcal{R}^p} \mathsf{cov}(\langle \alpha, X \rangle, \langle \beta, Y \rangle) = \mathsf{cov}(\langle \alpha_1, X \rangle, \langle \beta_1, Y \rangle),$$

  subject to $\mathsf{var}(\langle \alpha, X \rangle) = 1$ and $\mathsf{var}(\langle \beta, Y \rangle) = 1$.

Review of Multivariate Canonical Correlation

- The k-th $(k > 1)$ canonical correlation ρ_k and its associated weights can be defined similarly as:

$$\rho_k = \sup_{\alpha, \beta \in \mathcal{R}^p} \mathsf{cov}(\langle \alpha, X \rangle, \langle \beta, Y \rangle) = \mathsf{cov}(\langle \alpha_k, X \rangle, \langle \beta_k, Y \rangle),$$

subject to $\mathsf{var}(\langle \alpha, X \rangle) = 1$, $\mathsf{var}(\langle \beta, Y \rangle) = 1$, and

$$(U_k, V_k) = (\langle \alpha_k, X \rangle, \langle \beta_k, Y \rangle)$$

is uncorrelated to all previous pairs

$$(U_j, V_j) = (\langle \alpha_j, X \rangle, \langle \beta_j, Y \rangle), \text{ for } j = 1, \ldots, k-1.$$

Definition of Functional Canonical Correlation

- Replacing the inner product in Euclidean space with the L^2-inner product, $\langle f, g \rangle = \int_I f(t)g(t)dt$, we arrive at functional canonical correlations with a series of functional canonical components

  $(\rho_k, \alpha_k, \beta_k, U_k, V_k)$, $k \geq 1$,

  where U_k and V_k are maximally correlated and (U_k, V_k) are uncorrelated across k.

Definition of Functional Canonical Correlation

- Replacing the inner product in Euclidean space with the L^2-inner product, $\langle f, g \rangle = \int_I f(t)g(t)dt$, we arrive at functional canonical correlations with a series of functional canonical components

  $(\rho_k, \alpha_k, \beta_k, U_k, V_k)$, $k \geq 1$,

  where U_k and V_k are maximally correlated and (U_k, V_k) are uncorrelated across k.

- It can be shown (as for multivariate data) that functional canonical correlation analysis (FCCA) corresponds to eigenanalysis of the operator $R = \Sigma_{XX}^{-1/2}\Sigma_{XY}\Sigma_{YY}^{-1/2}$,

  where Σ_{XX} and Σ_{YY} are the corresponding covariance operator for X and Y respectively, and Σ_{XY} is the cross-covariance operator.

Inverse Problem of FCCA

- Existence of the canonical components is guaranteed if R is a bounded operator.

Inverse Problem of FCCA

- Existence of the canonical components is guaranteed if R is a bounded operator.

  -While this is feasible for multivariate data, the inverse of a covariance operator is NEVER bounded because a covariance operator is a compact operator.

Inverse Problem of FCCA

- Existence of the canonical components is guaranteed if R is a bounded operator.

  -While this is feasible for multivariate data, the inverse of a covariance operator is NEVER bounded because a covariance operator is a compact operator.

  - A remedy was provided in He, M. and W. (2003), which defined a generalized inverse under strong conditions that requires the eigenvalues decay fast to zero.

More Problems with FCCA

- Although it is possible to resolve the inverse problem by imposing strong conditions, CCA often has an overfitting problem in that the first canonical correlation tends to be very large and hard to interpret.

More Problems with FCCA

- Although it is possible to resolve the inverse problem by imposing strong conditions, CCA often has an overfitting problem in that the first canonical correlation tends to be very large and hard to interpret.

  - This overfitting problem already exists for multivariate data when the dimension is relatively high (but still finite).

  - It is caused by the large degree of freedoms in choosing the p-dim canonical weights.

More Problems with FCCA

- This overfitting problem is magnified for functional data as theoretically the d.f. is ∞ because the weight functions are infinitely dimensional.

  In practice, the weight functions are obtained on a dense grid but it still involves a large d.f.

More Problems with FCCA

- This overfitting problem is magnified for functional data as theoretically the d.f. is ∞ because the weight functions are infinitely dimensional.

  In practice, the weight functions are obtained on a dense grid but it still involves a large d.f.

- Another challenge with FCCA is the theory. The ill-posed nature of FCCA triggers theoretical challenges.

  - e.g. $\sqrt{n}$-rate of convergence is not feasible for estimates of the canonical correlations ρ_k within the L^2 paradigm.

Cautionary Remarks with FCCA

- For these reasons, FCCA must be used with caution.

Cautionary Remarks with FCCA

- For these reasons, FCCA must be used with caution.

- Alternative functional correlations are called for.

  One of them is the dynamic correlation (DC) proposed by Dubin and M. (2005), which is a functional version of Pearson correlation.

Cautionary Remarks with FCCA

- For these reasons, FCCA must be used with caution.
- Alternative functional correlations are called for.

  One of them is the dynamic correlation (DC) proposed by Dubin and M. (2005), which is a functional version of Pearson correlation.
- DC avoids involving the entire covariance operator to overcome both the inverse and overfitting problem intrinsic to FCCA.

  - We may explore this in a later section with neuroimaging applications, if time permits.

End of FCCA

Outline

Inverse Problem in Functional Regression

- Intuitively, any regression method/model involves an inverse problem, but not all inverse problems are ill-posed.

Inverse Problem in Functional Regression

- Intuitively, any regression method/model involves an inverse problem, but not all inverse problems are ill-posed.

- Any procedure that involves inverting a covariance operator is ill-posed.

Inverse Problem in Functional Regression

- Intuitively, any regression method/model involves an inverse problem, but not all inverse problems are ill-posed.
- Any procedure that involves inverting a covariance operator is ill-posed.
- There are three scenarios for functional regressions:
  (i) scalar (or vector) response with functional and possibly additional vector covariates

  (ii) Functional response with scalar covariates

  (iii) Functional response with functional and possibly additional vector covariates.

Functional Linear Regression:

Scalar Response Y & Functional Covariate $X(t)$

- How to construct a functional linear model?

Functional Linear Regression:

Scalar Response Y & Functional Covariate $X(t)$

- How to construct a functional linear model?

  Answer: Always think of the vector case first, then replace the inner product !

  Linear model: $Y = \beta_0 + \langle \beta_1, X \rangle +$ error.

Functional Linear Regression:

Scalar Response Y & Functional Covariate $X(t)$

- How to construct a functional linear model?

  Answer: Always think of the vector case first, then replace the inner product !

  Linear model: $Y = \beta_0 + \langle \beta_1, X \rangle +$ error.

- Functional linear model: $Y = \beta_0 + \int_I \beta_1(t) X(t) dt +$ error.

Functional Linear Regression:

Scalar Response Y & Functional Covariate $X(t)$

- How to construct a functional linear model?

  Answer: Always think of the vector case first, then replace the inner product !

  Linear model: $Y = \beta_0 + \langle \beta_1, X \rangle +$ error.

- Functional linear model: $Y = \beta_0 + \int_I \beta_1(t) X(t) dt +$ error.

- Is this an ill-pose problem?

Functional Linear Regression:

Scalar Response Y & Functional Covariate $X(t)$

- How to construct a functional linear model?

  Answer: Always think of the vector case first, then replace the inner product !

  Linear model: $\quad Y = \beta_0 + \langle \beta_1, X \rangle +$ error.

- Functional linear model: $Y = \beta_0 + \int_I \beta_1(t) X(t) dt +$ error.

- Is this an ill-pose problem?

- YES! $\quad \beta_1 = \Sigma_{XX}^{-1}$ cov (Y, X).

Functional Linear Regression:

Functional Response Y(t) & p-dim Vector Covariate X

- How to construct a functional linear model?

Functional Linear Regression:

Functional Response Y(t) & p-dim Vector Covariate X

- How to construct a functional linear model?

  Answer: Replace the inner product in the vector (Y) case!

  Multivariate Linear model: $Y = \beta_0 + \beta_1 X +$ error.
  - Here Y and β_0 are q-dim vectors, but β_1 is a $q \times p$ matrix.

Functional Linear Regression:

Functional Response Y(t) & p-dim Vector Covariate X

- How to construct a functional linear model?

  Answer: Replace the inner product in the vector (Y) case!

  Multivariate Linear model: $Y = \beta_0 + \beta_1 X +$ error.
  - Here Y and β_0 are q-dim vectors, but β_1 is a $q \times p$ matrix.

- Functional linear model: $Y(t) = \beta_0(t) + \langle \beta_1(t), X \rangle + b(t) +$ error, where $\beta_1(t)$ consists of p columns of functions and $b(t)$ is a random function. $\longrightarrow$ "varying-coefficient" model.

Functional Linear Regression:

Functional Response Y(t) & p-dim Vector Covariate X

- How to construct a functional linear model?

  Answer: Replace the inner product in the vector (Y) case!

  Multivariate Linear model: $Y = \beta_0 + \beta_1 X +$ error.
  - Here Y and β_0 are q-dim vectors, but β_1 is a $q \times p$ matrix.

- Functional linear model: $Y(t) = \beta_0(t) + \langle \beta_1(t), X \rangle + b(t) +$ error, where $\beta_1(t)$ consists of p columns of functions and $b(t)$ is a random function. $\longrightarrow$ "varying-coefficient" model.

- Is this an ill-posed problem?

Functional Linear Regression:

Functional Response Y(t) & p-dim Vector Covariate X

- How to construct a functional linear model?

  Answer: Replace the inner product in the vector (Y) case!

  Multivariate Linear model: $Y = \beta_0 + \beta_1 X +$ error.
  - Here Y and β_0 are q-dim vectors, but β_1 is a $q \times p$ matrix.

- Functional linear model: $Y(t) = \beta_0(t) + \langle \beta_1(t), X \rangle + b(t) +$ error, where $\beta_1(t)$ consists of p columns of functions and $b(t)$ is a random function. $\longrightarrow$ "varying-coefficient" model.

- Is this an ill-posed problem?

- No! $\beta_1(t) = \Sigma_{XX}^{-1}$ cov $(Y(t), X)$.

Functional Linear Regression:

Functional Response & and Functional Covariate

- How to construct a functional linear model?

Functional Linear Regression:

Functional Response & and Functional Covariate

- How to construct a functional linear model?

  Answer: Replace the inner product in

  Multivariate Linear model : $\quad Y = \beta_0 + \langle \beta_1, X \rangle +$ error,
  where X and Y are vectors and β_1 is a matrix.

Functional Linear Regression:

Functional Response & and Functional Covariate

- How to construct a functional linear model?

  Answer: Replace the inner product in

  Multivariate Linear model : $\quad Y = \beta_0 + \langle \beta_1, X \rangle +$ error,
  where X and Y are vectors and β_1 is a matrix.

- Functional linear model:
  $Y(t) = \beta_0(t) + \int_{I_X} \beta_1(s,t) X(s) ds + b(t) +$ error.

Functional Linear Regression:

Functional Response & and Functional Covariate

- How to construct a functional linear model?

  Answer: Replace the inner product in

  Multivariate Linear model : $Y = \beta_0 + \langle \beta_1, X \rangle +$ error,
  where X and Y are vectors and β_1 is a matrix.

- Functional linear model:
  $Y(t) = \beta_0(t) + \int_{I_X} \beta_1(s,t)X(s)ds + b(t) +$ error.

- Is this an ill-posed problem?

Functional Linear Regression:

Functional Response & and Functional Covariate

- How to construct a functional linear model?

  Answer: Replace the inner product in

  Multivariate Linear model : $\quad Y = \beta_0 + \langle \beta_1, X \rangle +$ error,
  where X and Y are vectors and β_1 is a matrix.

- Functional linear model:
  $Y(t) = \beta_0(t) + \int_{I_X} \beta_1(s,t) X(s) ds + b(t) +$ error.

- Is this an ill-posed problem?

- YES! $\quad \beta_1(t) = \Sigma_{XX}^{-1}$ cov $(Y(t), X(t))$.

Functional Linear Regression when $X(t)$ and $Y(t)$ are measured together

- When X and Y are measured simultaneously, future values of X should not be used to predict $Y(t)$.
  $\Rightarrow Y(t) = \beta_0(t) + \int_0^t \beta_1(s,t)X(s)ds + b(t) +$ error.

  - This is called a "historical model" as only the history of X up to time t is used to predict Y at time t.

Functional Linear Regression when $X(t)$ and $Y(t)$ are measured together

- When X and Y are measured simultaneously, future values of X should not be used to predict $Y(t)$.
  $\Rightarrow Y(t) = \beta_0(t) + \int_0^t \beta_1(s,t) X(s) ds + b(t) +$ error.

  - This is called a "historical model" as only the history of X up to time t is used to predict Y at time t.

  - This model is ill-posed.

Functional Linear Regression when $X(t)$ and $Y(t)$ are measured together

- When X and Y are measured simultaneously, future values of X should not be used to predict $Y(t)$.
  $\Rightarrow Y(t) = \beta_0(t) + \int_0^t \beta_1(s,t)X(s)ds + b(t) +$ error.

  - This is called a "historical model" as only the history of X up to time t is used to predict Y at time t.

  - This model is ill-posed.

- There is another functional linear model:

  $Y(t) = \beta_0(t) + \beta_1(t)X(t) + b(t) +$ error.

  - This is a "concurrent" varying-coefficient model as only the current value of X is associated with the current Y value at time t.

Functional Linear Regression when $X(t)$ and $Y(t)$ are measured together

- When X and Y are measured simultaneously, future values of X should not be used to predict $Y(t)$.
  $\Rightarrow Y(t) = \beta_0(t) + \int_0^t \beta_1(s,t)X(s)ds + b(t) +$ error.

  - This is called a "historical model" as only the history of X up to time t is used to predict Y at time t.

  - This model is ill-posed.

- There is another functional linear model:

  $Y(t) = \beta_0(t) + \beta_1(t)X(t) + b(t) +$ error.

  - This is a "concurrent" varying-coefficient model as only the current value of X is associated with the current Y value at time t.

  - This model is NOT ill-posed.

Comparison of Functional Linear Models vs Linear Mixed-effects Model

- Functional varying-coefficient model:
  $Y(t) = \beta_0(t) + \beta_1(t)X(t) + b(t) +$ error.

- Historical Functional linear model:
  $Y(t) = \beta_0 + \int_0^t \beta_1(s,t)X(s)ds + b(t) +$ error

- Functional linear model using the entire X trajectory:
  $Y(t) = \beta_0 + \int_{I_X} \beta_1(s,t)X(s)ds + b(t) +$ error

  - Note here that X and Y need not be measured at the same time period.

Comparison of Functional Linear Models vs Linear Mixed-effects Model

- Functional varying-coefficient model:
  $Y(t) = \beta_0(t) + \beta_1(t)X(t) + b(t) +$ error.

- Historical Functional linear model:
  $Y(t) = \beta_0 + \int_0^t \beta_1(s,t)X(s)ds + b(t) +$ error

- Functional linear model using the entire X trajectory:
  $Y(t) = \beta_0 + \int_{I_X} \beta_1(s,t)X(s)ds + b(t) +$ error
  - Note here that X and Y need not be measured at the same time period.

- Linear Mixed-effects Model: $Y(t) = \beta_0 + \beta_1 X(t) + bZ(t) +$ error, where the coefficient β_1 is time-invariant, so is the random effect b.

More about Functional Regression Models

- All the aforementioned functional linear models can be extended to functional generalized linear model by adding a known link function and a prespecified variance function.

More about Functional Regression Models

- All the aforementioned functional linear models can be extended to functional generalized linear model by adding a known link function and a prespecified variance function.

- They can also be extended to an unknown link function, termed "functional single index model".

  e.g. $Y = \beta_0 + g(\int_I \beta_1(t)X(t)) +$ error,

  where g is an unknown function.
  (X needs to be a continuous process)

More about Functional Regression Models

- All the aforementioned functional linear models can be extended to functional generalized linear model by adding a known link function and a prespecified variance function.

- They can also be extended to an unknown link function, termed "functional single index model".

  e.g. $Y = \beta_0 + g(\int_I \beta_1(t)X(t)) +$ error,

  where g is an unknown function.
  (X needs to be a continuous process)

- Additional continuous time-dependent and time-independent covariates Z can be added to the model (How?)

More about Functional Regression Models

- If some of these covariates are discrete, a partial linear single-index model will be needed to model the discrete covariates linearly.

  e.g. $Y = \beta_0 + \theta_1 Z_1 + g(\theta_2 Z_2 + \int_I \beta_q(t)X(t)dt) +$ error.

More about Functional Regression Models

- If some of these covariates are discrete, a partial linear single-index model will be needed to model the discrete covariates linearly.

  e.g. $Y = \beta_0 + \theta_1 Z_1 + g(\theta_2 Z_2 + \int_I \beta_q(t)X(t)dt) +$ error.

- The single-index can be extended to multiple indices etc.

  e.g. $Y = \beta_0 + g(\int_I \beta_1(t)X(t)dt + \int_I \beta_2(t)X(t)dt) +$ error.

More about Functional Regression Models

- If some of these covariates are discrete, a partial linear single-index model will be needed to model the discrete covariates linearly.

  e.g. $Y = \beta_0 + \theta_1 Z_1 + g(\theta_2 Z_2 + \int_I \beta_q(t) X(t) dt) +$ error.

- The single-index can be extended to multiple indices etc.

  e.g. $Y = \beta_0 + g(\int_I \beta_1(t) X(t) dt + \int_I \beta_2(t) X(t) dt) +$ error.

- All these index based models are "dimension reduction" models. Another type of dimension reduction model is the "additive model",
  .e.g. $Y(t) = \beta_0(t) + \sum \phi_k(X_k) + b(t) +$ error.

Time-varying Additive Model: Zhang and W. (2015, Biometrika)

- The additive model, $Y(t) = \beta_0(t) + \sum \phi_k(X_k) + b(t)+$ error ,

  can be extended to allow for time-varying coefficients.

  $Y(t) = \beta_0(t) + \sum \beta_k(t)\phi_k(X_k) = b(t)+$ error.

  "Time-varying additive model"

Time-varying Additive Model: Zhang and W. (2015, Biometrika)

- The additive model, $Y(t) = \beta_0(t) + \sum \phi_k(X_k) + b(t) +$ error , can be extended to allow for time-varying coefficients.

  $Y(t) = \beta_0(t) + \sum \beta_k(t)\phi_k(X_k) = b(t) +$ error.

  "Time-varying additive model"

- There are lots of models that one can construct!

  - More later by Hans.

  - Review on functional regression by Morris (2015).

End of Inverse Problem

- Just hang in there! -

Outline

Next Generation FD

- So far, the functional data are 1D independent curve data.
- These curve data can be dependent

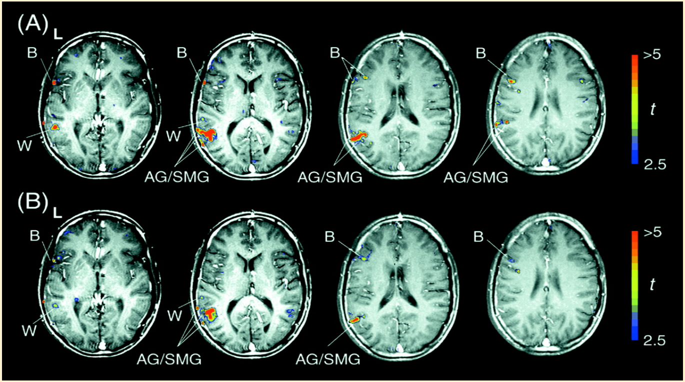
  e.g. Functional time series and spatio-temporal data.

Next Generation FD

- So far, the functional data are 1D independent curve data.
- These curve data can be dependent

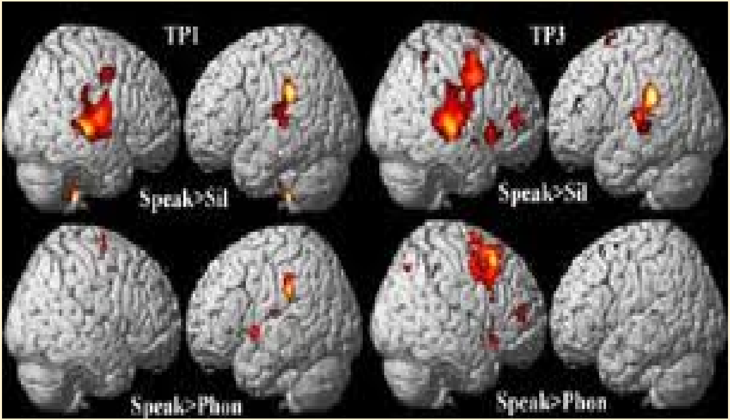
  e.g. Functional time series and spatio-temporal data.
- Functional data can be 2D, 3D images, or even 4D data.
- Functional data can also include objects, shapes, trees, networks, etc.

2D Functional Data: http://www.pnas.org/content/97/11/6150/F1.expansion.html



3D Functional Data:

http://www.musicianbrain.com/images/fmri.jpg



What is the Dimension of fMRI data?

- For a single subject:

  Spatially (3D) correlated temporal (1D) data

  Temporally correlated 3D data

  $\rightarrow$ Atom of Longitudinal 3D functional data

What is the Dimension of fMRI data?

- For a single subject:

  Spatially (3D) correlated temporal (1D) data

  Temporally correlated 3D data

  $\rightarrow$ Atom of Longitudinal 3D functional data

- For multi subjects:

  Independent Sample of 4D functional data

  multi-level data if there are multiple scans

FDA for Neuroimaging Data

- Neuroimaging data is intrinsically functional, often multifaceted.

FDA for Neuroimaging Data

- Neuroimaging data is intrinsically functional, often multifaceted.
- Neuroimaging data is a gold mine for FDA.

FDA for Neuroimaging Data

- Neuroimaging data is intrinsically functional, often multifaceted.
- Neuroimaging data is a gold mine for FDA.
- The most prevailing FDA method for NDA has been Functional PCA.

FDA for Neuroimaging Data

- Neuroimaging data is intrinsically functional, often multifaceted.
- Neuroimaging data is a gold mine for FDA.
- The most prevailing FDA method for NDA has been Functional PCA.
- Other functional approaches are emerging daily!

Summary of Next Generation Functional Data

- Correlated functional data

  e.g. functional time series, spatio-temporal data
- Independent k-D data
- Correlated k-D data
- Longitudinal functional data
- Multi-level functional data
- ...... any others?

Next Generation Functional Data

- *Object data*

  e.g. shapes, trees, networks, etc.

Next Generation Functional Data

- *Object data*

  e.g. shapes, trees, networks, etc.

- All seems complex - Some non-Euclidean!

Next Generation Functional Data

- *Object data*

  e.g. shapes, trees, networks, etc.

- All seems complex - Some non-Euclidean!

- Many are BIG!

  e.g. Neuroimaging data

Next Generation Functional Data

- *Object data*

  e.g. shapes, trees, networks, etc.

- All seems complex - Some non-Euclidean!

- Many are BIG!

  e.g. Neuroimaging data

- They are challenging, but represent opportunities!

End of Next Generation Functional Data

List of Review Papers

- A review paper (W. Chiou and M., 2016, Annual Review of Statistics and Its Application)
- Another review paper on functional regression by Morris (2015, Annual Review of Statistics and Its Application)

List of Review Papers

- A review paper (W. Chiou and M., 2016, Annual Review of Statistics and Its Application)
- Another review paper on functional regression by Morris (2015, Annual Review of Statistics and Its Application)
- Additional review papers:
  Müller (2005, Scan J. Stat.) - regression and classification

  Müller (2008, Handbook on LDA, ed. Davidian, Verbeke and Molenberghs) - functional modeling for LD

  Müller (2011, International Encyclopedia of Stat. Science) - FDA

List of Books on Functional Data

- Ramsay and Silverman (2005)
- Ferraty and Vieu (2006)
- Wu and Zhang (2006)
- Horváth and Kokoska (2012)
- Hsing and Eubank (2015)

End of Part III: Thanks You!