Methods to Assess an Exercise Intervention Trial based on Three-Level Functional Data

Ordered Data Analysis, Models and Health Research Methods

Houssein Assaad

Department of Statistics, Texas A&M University, College Station, TX, USA.

Mar 7 – 9, 2014

1 Joint work with Haocheng Li, Jianhua Huang and Raymond Carroll.
Outline

Introduction

Data and Model Setup
  Motivation
  Summary I
  Estimation Algorithm

Simulation Results
  Model Fit
  Wald Test Performance

Application: Revisiting the Physical activity data

Summary
What is longitudinal/multi-level functional data?

The basic observational unit is a function.

<table>
<thead>
<tr>
<th>Longitudinal Functional Data</th>
<th>Visit 1 At time $s_{i1}$</th>
<th>Visit 2 At time $s_{i2}$</th>
<th>...</th>
<th>Visit $J_i$ At time $s_{ij_i}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject 1</td>
<td><img src="image1.png" alt="Graph" /></td>
<td><img src="image2.png" alt="Graph" /></td>
<td></td>
<td><img src="image3.png" alt="Graph" /></td>
</tr>
<tr>
<td>Subject 2</td>
<td><img src="image4.png" alt="Graph" /></td>
<td><img src="image5.png" alt="Graph" /></td>
<td></td>
<td><img src="image6.png" alt="Graph" /></td>
</tr>
<tr>
<td></td>
<td><img src="image7.png" alt="Graph" /></td>
<td><img src="image8.png" alt="Graph" /></td>
<td></td>
<td><img src="image9.png" alt="Graph" /></td>
</tr>
<tr>
<td></td>
<td><img src="image10.png" alt="Graph" /></td>
<td><img src="image11.png" alt="Graph" /></td>
<td></td>
<td><img src="image12.png" alt="Graph" /></td>
</tr>
<tr>
<td>Subject $I$</td>
<td><img src="image13.png" alt="Graph" /></td>
<td><img src="image14.png" alt="Graph" /></td>
<td></td>
<td><img src="image15.png" alt="Graph" /></td>
</tr>
</tbody>
</table>
Mercer theorem and K-L expansion

Let \( Y(t) \in L^2[0, 1] \), with \( \mu(t) = E\{Y(t)\} \) and 
\( K(s, t) = \text{cov}\{Y(s), Y(t)\} \).

Spectral decomposition of \( K(s, t) \) (Mercer’s Theorem):

\[
K(s, t) = \sum_{l=1}^{\infty} \lambda_l f_l(s)f_l(t)
\]

where \( \lambda_1 \geq \lambda_2 \geq \cdots \geq 0 \) are eigenvalues and \( f_l \)'s are the corresponding orthonormal eigenfunctions.

Karhunen-Loève (K-L) expansion:

\[
Y(t) = \mu(t) + \sum_{l=1}^{\infty} f_l(t)\alpha_l
\]

where \( \alpha_l = \int_{0}^{1} \{Y(t) - \mu(t)\} f_l(t)dt \) are \( \overset{\text{unc}}{\sim} (0, \lambda_l) \) called PC scores or loadings. \( \{\lambda_l, f_l\}_l \) are as defined in Mercer’s Thm.
Outline

Introduction

Data and Model Setup
  Motivation
  Summary I
  Estimation Algorithm

Simulation Results
  Model Fit
  Wald Test Performance

Application: Revisiting the Physical activity data

Summary
Exercise Intervention Trial Data

- Data are from (Kozey-Keadle et al., 2013).
- Consist of estimates of relative energy expenditure (metabolic units, or METs) on \( n = 63 \) individuals every 5 minutes, 5 days a week for 5 separate weeks (study weeks 0, 3, 6, 9, and 12).
- The treatment assignment was done after the baseline week and consisted of assignment to either a control arm or an exercise intervention where subjects completed a standardized aerobic exercise program.
- Interested in the patterns of physical activity for individuals and how these patterns vary across the treatment groups.
- Methodology for the analysis of three-level functional data is very limited and non-existent for the analysis of physical activity data.
Figure: This figure shows an example of the estimates of METs for a subject in the exercise group.
Model Setup

Let $Y_{ijk}(t)$ be the response at time $t$ for subject $i = 1, \ldots, n$, in week $j$ ($j = 1, \ldots, J$) on day $k$ ($k = 1, \ldots, K$)

\[
Y_{ijk}(t) = \underbrace{\mu_{..}(t) + \mu_{j}.(t) + \mu_{.k}(t) + \mu_{jk}(t)}_{\text{fixed-effect curves}} + \underbrace{\xi_i(t) + \eta_{ij}(t) + \zeta_{ik}(t) + \gamma_{ijk}(t) + \epsilon_{ijk}(t)}_{\text{random-effect curves}}
\]

Where

- $\mu_{..}(t)$ are the population mean curves, $\mu_{j}.(t)$, $\mu_{.k}(t)$ and $\mu_{jk}(t)$ are week-specific, day-specific and week×day interaction mean curves.
- $\xi_i(t)$, $\eta_{ij}(t)$, $\zeta_{ik}(t)$, and $\gamma_{ijk}(t)$ are mutually independent.
- They represent the subject-specific, week-within-subject, day-within-subject and week×day interaction-within-subject random effects curves, respectively.
- $\epsilon_{ijk}(t) \overset{\text{unc}}{\sim} \mathcal{WN}(0, \sigma^2)$
- For identifiability, we constrain $\mu_{1.}(t) = \mu_{.1}(t) = \mu_{1k}(t) = \mu_{j1}(t) = 0$ for all $k, j$. 

Covariance Structures

Same subject, week, day

$$\text{cov} \{ Y_{ijk}(t), Y_{ijk}(s) \} = \text{cov} \{ \xi_i(t), \xi_i(s) \} + \text{cov} \{ \eta_{ij}(t), \eta_{ij}(s) \} + \text{cov} \{ \zeta_{ik}(t), \zeta_{ik}(s) \} + \text{cov} \{ \gamma_{ijk}(t), \gamma_{ijk}(s) \}$$

Same subject, week, different day

$$k \neq k', \text{cov} \{ Y_{ijk}(t), Y_{ijk'}(s) \} = \text{cov} \{ \xi_i(t), \xi_i(s) \} + \text{cov} \{ \eta_{ij}(t), \eta_{ij}(s) \}$$

Same subject, day, different week

$$j \neq j', \text{cov} \{ Y_{ijk}(t), Y_{ij'k}(s) \} = \text{cov} \{ \xi_i(t), \xi_i(s) \} + \text{cov} \{ \zeta_{ik}(t), \zeta_{ik}(s) \}$$

And, same subject only

$$j \neq j', k \neq k', \text{cov} \{ Y_{ijk}(t), Y_{ij'k'}(s) \} = \text{cov} \{ \xi_i(t), \xi_i(s) \}$$
The Karhunen-Loève (K-L) expansion

Applying (K-L) expansion on random effect curves,

\[ \xi_i(t) = \sum_{l=1}^{P_\xi} f_{\xi,l}(t) \alpha_{\xi,i,l} = f_{\xi}^T \alpha_{\xi,i} ; \eta_{ij}(t) = \sum_{l=1}^{P_\eta} f_{\eta,l}(t) \alpha_{\eta,ij,l} = f_{\eta}^T \alpha_{\eta,ij} \]

\[ \zeta_{ik}(t) = \sum_{l=1}^{P_\zeta} f_{\zeta,l}(t) \alpha_{\zeta,ik,l} = f_{\zeta}^T \alpha_{\zeta,ik} ; \gamma_{ijk}(t) = \sum_{l=1}^{P_\gamma} f_{\gamma,l}(t) \alpha_{\gamma,ijk,l} = f_{\gamma}^T \alpha_{\gamma,ijk} \]

where

\[ \int f_{\dagger,l}(t)f_{\dagger,l'}(t)dt = \delta_{ll'}, \text{ where } \dagger \text{ is one of } \xi, \eta, \zeta, \gamma \]
Modeling with B-splines

Let $b(t) = \{b_1(t), ..., b_q(t)\}^T$ be the $q \times 1$ vector of B-splines basis fcns. evaluated at $t$. We model the fixed effects as

$$
\mu_{..}(t) = b^T(t)\beta_{..}; \quad \mu_{j.}(t) = b^T(t)\beta_{j.}; \quad \mu_{.k}(t) = b^T(t)\beta_{.k}; \quad \mu_{jk}(t) = b^T(t)\beta_{jk}
$$

and the FPC eigenfunctions as

$$
f_{\xi,l}(t) = b^T(t)g_{\xi,l}; \quad f_{\eta,l}(t) = b^T(t)g_{\eta,l}; \quad f_{\zeta,l}(t) = b^T(t)g_{\zeta,l}; \quad f_{\gamma,l}(t) = b^T(t)g_{\gamma,l}
$$

To maintain the orthogonality restrictions on the eigenfunctions, $b(t)$, $g_{\xi,l}$, $g_{\eta,l}$, $g_{\zeta,l}$ and $g_{\gamma,l}$ are restricted to be orthogonal.

For identifiability, we constrain $\beta_{1.}(t) = \beta_{.1}(t) = \beta_{1k}(t) = \beta_{j1}(t) = 0$ for all $k, j$. 
Outline

Introduction

Data and Model Setup
  Motivation
  Summary I
  Estimation Algorithm

Simulation Results
  Model Fit
  Wald Test Performance

Application: Revisiting the Physical activity data

Summary
Summary 1

We started with

\[ Y_{ijk}(t) = \mu..(t) + \mu.j.(t) + \mu..k(t) + \mu.jk(t) \\
+ \xi_i(t) + \eta_{ij}(t) + \zeta_{ik}(t) + \gamma_{ijk}(t) + \epsilon_{ijk}(t) \]

- **Apply K-L expansion** on \( \xi_i(t), \eta_{ij}(t), \zeta_{ik}(t), \gamma_{ijk}(t) \)
- **Represent the fixed effects** \( \mu..(t), \mu.j.(t), \mu..k, \mu.jk(t) \) and the eigenfunctions \( (f_{\dagger,l}(t) = b^T(t)g_{\dagger,l}) \) using **B-splines**.

The resulting model is

\[ Y_{ijk}(t) = b^T(t)\beta.. + b^T(t)\beta.j + b^T(t)\beta.k + b^T(t)\beta.jk \\
+ b^T(t)G_{\xi} \alpha_{\xi,i} + b^T(t)G_{\eta} \alpha_{\eta,ij} + b^T(t)G_{\zeta} \alpha_{\zeta,ik} + b^T(t)G_{\gamma} \alpha_{\gamma,ijk} \\
+ \epsilon_{ijk}(t) \]

where \( G_{\dagger} = (g_{\dagger,1}, ..., g_{\dagger,P_{\dagger}}) \) and \( \dagger \) is one of \( \xi, \eta, \zeta, \gamma \).
Outline

Introduction

Data and Model Setup
  Motivation
  Summary I
  Estimation Algorithm

Simulation Results
  Model Fit
  Wald Test Performance

Application: Revisiting the Physical activity data

Summary
Let \( \Delta_\xi = \text{cov}(\alpha_{\xi,i}) \), \( \Delta_\eta = \text{cov}(\alpha_{\eta,ij}) \), \( \Delta_\zeta = \text{cov}(\alpha_{\zeta,ik}) \), and \( \Delta_\gamma = \text{cov}(\alpha_{\gamma,ijk}) \). \( \Delta_\dagger \) is a \( P_\dagger \times P_\dagger \) (\( \dagger \) is one of \( \xi, \eta, \zeta, \gamma \)) diagonal matrix with decreasing positive diagonal elements.

Define new \( q \)-dim r.v.

\[
\begin{align*}
    u_{\xi,i} &= G_\xi \alpha_{\xi,i}; \\
    u_{\eta,ij} &= G_\eta \alpha_{\eta,ij}; \\
    u_{\zeta,ik} &= G_\zeta \alpha_{\zeta,ik}; \\
    u_{\gamma,ijk} &= G_\gamma \alpha_{\gamma,ijk}
\end{align*}
\]

The resulting model is

\[
Y_{ijk}(t) = b^T(t)\beta + b^T(t)\beta_j + b^T(t)\beta_k + b^T(t)\beta_{jk} + b^T(t)u_{\xi,i} + b^T(t)u_{\eta,ij} + b^T(t)u_{\zeta,ik} + b^T(t)u_{\gamma,ijk} + \epsilon_{ijk}(t)
\]  

(1) with

\[
\begin{align*}
    \text{cov}(u_{\xi,i}) &= \psi_\xi = G_\xi \Delta_\xi G_\xi^T; \\
    \text{cov}(u_{\eta,ij}) &= \psi_\eta = G_\eta \Delta_\eta G_\eta^T; \\
    \text{cov}(u_{\zeta,ik}) &= \psi_\zeta = G_\zeta \Delta_\zeta G_\zeta^T; \\
    \text{cov}(u_{\gamma,ijk}) &= \psi_\gamma = G_\gamma \Delta_\gamma G_\gamma^T
\end{align*}
\]

If \( P_\xi = P_\zeta = P_\eta = P_\gamma = q \), then \( \psi_\xi, \psi_\zeta, \psi_\eta \) and \( \psi_\gamma \) would be full rank and (1) is equivalent to a 3-level LMM with unstructured random effect covariance matrices.
Estimation Procedure

\[ Y_{ijk}(t) = b^T(t)\beta_i + b^T(t)\beta_j + b^T(t)\beta_k + b^T(t)\beta_{jk} + b^T(t)u_{\xi,i} + b^T(t)u_{\eta,ij} + b^T(t)u_{\zeta,ik} + b^T(t)u_{\gamma,ijk} + \epsilon_{ijk}(t) \]

Let \( \beta \) be a \( JKq \)-vector containing all the \( \beta_i \) and \( U_i \) be a \( \{J + 1)(K + 1)q\} \)-r.v. containing all the \( u_i \). Then, (1) can be expressed as

\[ Y_i = B_i^\mu \beta + B_i^U U_i + \epsilon_i \]

with

\[ \Psi \overset{\text{def}}{=} \text{cov} (U_i) = \text{diag} (\psi_{\xi}, I_J \otimes \psi_{\eta}, I_K \otimes \psi_{\zeta}, I_{JK} \otimes \psi_{\gamma}) \]

\[ V_i \overset{\text{def}}{=} \text{cov} (Y_i) / \sigma^2 = I_{N_i} + B_i^U \Psi (B_i^U)^T / \sigma^2 \]

\[ S_i \overset{\text{def}}{=} \text{cov} (U_i | Y_i) / \sigma^2 = \Psi / \sigma^2 - \Psi (B_i^U)^T V_i^{-1} B_i^U \Psi / \sigma^4 \]

\[
\begin{align*}
(\beta_{\text{curr}}, \sigma^2_{\text{curr}}, \Psi_{\text{curr}}) \overset{(3),(4)}{\rightarrow} (S_{i,\text{curr}}, V_{i,\text{curr}}) \overset{\text{ECME}}{\rightarrow} (\beta_{\text{new}}, \sigma^2_{\text{new}}, \Psi_{\text{full}}) \overset{\text{RR model}}{\rightarrow} \Psi_{\text{new}}
\end{align*}
\]

The ECME step and the Reduced-Rank (RR) model will be described in the next slide.
ECME (Schafer, 1998) and the Reduced-Rank (RR) Model

ECME:

\[
\begin{align*}
\beta_{\text{new}} &= \left( \sum_{i=1}^{n} (B_i^\mu)^T V_{i,\text{curr}}^{-1} B_i^\mu \right)^{-1} \left( \sum_{i=1}^{n} (B_i^\mu)^T V_{i,\text{curr}}^{-1} Y_i \right) \\
\sigma^2_{\text{new}} &= N^{-1} \sum_{i=1}^{n} (Y_i - B_i^\mu \beta_{\text{new}})^T V_{i,\text{curr}}^{-1} (Y_i - B_i^\mu \beta_{\text{new}}) \\
\Psi_{\text{full}}^{\text{new}} &= n^{-1} \sum_{i=1}^{n} \left( \hat{U}_i \hat{U}_i^T + \sigma^2_{\text{new}} S_{i,\text{curr}} \right) \quad \text{and} \quad \hat{U}_i = S_{i,\text{curr}} (B_i U_i)^T (Y_i - B_i^\mu \beta_{\text{new}})
\end{align*}
\]

How to obtain \( \Psi_{\text{RR}, \text{curr}}^{\text{RR}} \)? (Recall: \( \dagger \) is one of \( \xi, \eta, \zeta \) or \( \gamma \))

- Extract all \( \psi_{\dagger, \text{curr}} \) from the diagonal of \( \Psi_{\text{full}}^{\text{new}} \) and perform eigenvalue decompostion

\[
\psi_{\dagger, \text{curr}} = \tilde{G}_{\dagger, \text{new}} \tilde{\Delta}_{\dagger, \text{new}} \tilde{G}_{\dagger, \text{new}}^T_{q \times q}
\]

- Select the number of principal components \( (P_\xi, P_\eta, P_\zeta, P_\gamma) \) to keep, by requiring the ratio of variance explained to be \( \geq P = 0.85 \).
- Retain 1st \( P_\dagger \) eigenvectors and eigenvalues from \( \tilde{G}_{\dagger, \text{new}} \) and \( \tilde{\Delta}_{\dagger, \text{new}} \) to form \( G_{\dagger, \text{new}}_{q \times P_\dagger} \)

and \( \Delta_{\dagger, \text{new}}_{P_\dagger \times P_\dagger} \).

- Obtain the updated RR matrices \( \psi_{\text{RR}, \text{new}}^{\text{RR}} = G_{\dagger, \text{new}} \Delta_{\dagger, \text{new}} G_{\dagger, \text{new}}^T_{P_\dagger \times P_\dagger} \) and combine these to get \( \Psi_{\text{new}}^{\text{RR}} \) according to (2).
Outline

Introduction

Data and Model Setup
  Motivation
  Summary I
  Estimation Algorithm

Simulation Results
  Model Fit
  Wald Test Performance

Application: Revisiting the Physical activity data

Summary
Simulation: Model Fit

A measurement at time $t$ on day $k$ in week $j$ for subject $i$ is generated according to

$$Y_{ijk}(t) = \mu_i(t) + \mu_j(t) + \mu_k(t) + \sum_{l=1}^{2} f_{\xi,l}(t)\alpha_{\xi,i,l} + \sum_{l=1}^{2} f_{\eta,l}(t)\alpha_{\eta,ij,l}$$

$$+ \sum_{l=1}^{2} f_{\zeta,l}(t)\alpha_{\zeta,ik,l} + \sum_{l=1}^{2} f_{\gamma,l}(t)\alpha_{\gamma,ijk,l} + \epsilon_{ijk}(t)$$

- $n = 60$ subjects, $J = 5$ weeks and $K = 5$ days and each day has 36 measurement times.
- B-spline basis functions with 10 equispaced knots to fit the data.

Table: The average estimates and mean squared errors (MSE) of the parameters in the joint model. The number marked with an asterisk is the actual number multiplied by 1000.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>$\sigma^2$</th>
<th>$\Delta_{\xi,1}$</th>
<th>$\Delta_{\xi,2}$</th>
<th>$\Delta_{\eta,1}$</th>
<th>$\Delta_{\eta,2}$</th>
<th>$\Delta_{\zeta,1}$</th>
<th>$\Delta_{\zeta,2}$</th>
<th>$\Delta_{\gamma,1}$</th>
<th>$\Delta_{\gamma,2}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>True</td>
<td>1.00</td>
<td>8.00</td>
<td>4.00</td>
<td>6.00</td>
<td>3.00</td>
<td>4.00</td>
<td>2.00</td>
<td>2.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Mean</td>
<td>1.00</td>
<td>8.33</td>
<td>3.86</td>
<td>6.00</td>
<td>2.86</td>
<td>3.89</td>
<td>2.00</td>
<td>1.95</td>
<td>0.93</td>
</tr>
<tr>
<td>MSE</td>
<td>0.08$^*$</td>
<td>2.11</td>
<td>0.55</td>
<td>0.26</td>
<td>0.11</td>
<td>0.18</td>
<td>0.04</td>
<td>0.03</td>
<td>0.02</td>
</tr>
</tbody>
</table>
Simulation: Fixed-Effect Curves Fit

Figure 2: Fitted fixed effects curves for 500 simulated data sets: (a) fixed effects curve $\mu_1^{(1)}(t)$, (b)(c)(d)(e) fixed effects curves $\mu_2^{(1)}(t)$, $\mu_3^{(1)}(t)$, $\mu_4^{(1)}(t)$, $\mu_5^{(1)}(t)$, (f)(g)(h)(i) fixed effects curves $\mu_2^{(1)}(t)$, $\mu_3^{(1)}(t)$, $\mu_4^{(1)}(t)$, $\mu_5^{(1)}(t)$. Dotted lines denote true curves. Solid lines represent the averaged values of fitted curves. The upper and lower dashed lines are the 10% and 90% quantiles of the fitted values in 500 simulation studies.
Figure 3: Fitted principal components curves for 500 simulated data sets: (a)(b) principal components curves $f_{c,1}(t)$ and $f_{c,2}(t)$, (c)(d) principal components curves $f_{u,1}(t)$ and $f_{u,2}(t)$, (e)(f) principal components curves $f_{c,1}(t)$ and $f_{c,2}(t)$, (g)(h) principal components curves $f_{u,1}(t)$ and $f_{u,2}(t)$. Dotted lines denote true curves. Solid lines represent the averaged values of fitted curves. The upper and lower dashed lines are the 10% and 90% quantiles of the fitted values in 500 simulation studies.
Outline

Introduction

Data and Model Setup
  Motivation
  Summary I
  Estimation Algorithm

Simulation Results
  Model Fit
  Wald Test Performance

Application: Revisiting the Physical activity data

Summary
Performance of the Wald statistics to test $H_0 : \mu_{jk}(t) = 0$ for all $j, k$

- Complete and incomplete data scenarios are considered.
- 500 replicates for $n = 60$ and 200 replicates for $n = 250$.
- The probability that a day’s record is observed is 50%.

**Table:** Rejection rate of true null hypothesis for four methods with $\alpha = 0.05$. PEN10 represents the penalized likelihood with 10 knots. UN10, UN24 and UN30 are the unpenalized likelihood with 10, 24 and 30 knots, respectively.

<table>
<thead>
<tr>
<th></th>
<th>Complete Data $n = 60$</th>
<th>Missing Data $n = 60$</th>
<th>Complete Data $n = 250$</th>
<th>Missing Data $n = 250$</th>
</tr>
</thead>
<tbody>
<tr>
<td>PEN10</td>
<td>0.37</td>
<td>0.32</td>
<td>0.62</td>
<td>0.46</td>
</tr>
<tr>
<td>UN10</td>
<td>0.06</td>
<td>0.08</td>
<td>0.07</td>
<td>0.06</td>
</tr>
<tr>
<td>UN24</td>
<td>0.06</td>
<td>0.15</td>
<td>0.05</td>
<td>0.07</td>
</tr>
<tr>
<td>UN30</td>
<td>0.10</td>
<td>0.23</td>
<td>0.04</td>
<td>0.07</td>
</tr>
</tbody>
</table>
The activPAL™ is an electronic device used to estimate energy expenditure, which is expressed in units of metabolic equivalents (METs).

A person’s MET value for an activity is defined as the ratio of her energy expenditure during the activity to her resting energy expenditure.

A value of METs $\geq 3$ defines moderate to vigorous physical activity (MVPA).

In the study, 63 individuals wore the activPAL™ for five weeks (denoted weeks 0, 3, 6, 9, 12), five days in a week (Monday to Friday) and measurements were recorded every 5 minutes during each day.

We limited the data for each day to one hour before and two hours after in the day’s first instance of MVPA, if it occurs.
Figure: Fixed effects structures for the physical activity data. The upper and lower dashed lines are the 5% and 95% quantiles of the fitted values over 200 bootstrap estimates.
Figure: Random effects structures for the physical activity data. The upper and lower dashed lines are the 5% and 95% quantiles of the fitted values over 200 bootstrap estimates.
Summary

- Proposed a general modeling and estimation strategy in the presence of three-level correlated functional data.
- Our algorithm can handle incomplete data and employs a data-based method to select the number of principal components for the random curves.
- Simulation results for the model fitting strategy are encouraging and indicate little bias.
- We also discussed the performance of the Wald test to assess the evidence that linear combinations of mean structure parameters are zero.
- We applied our model to analyze data from a physical activity intervention trial. The detailed functional data analysis revealed how the patterns of activity differed in terms of timing and intensity.