Multivariate Functional Data Analysis

ORI ROSEN

Department of Mathematical Sciences, University of Texas at El Paso,

El Paso, Texas, USA

and

WESLEY THOMPSON

Department of Statistics, University of Pittsburgh, Pittsburgh,

Pennsylvania, USA

Outline

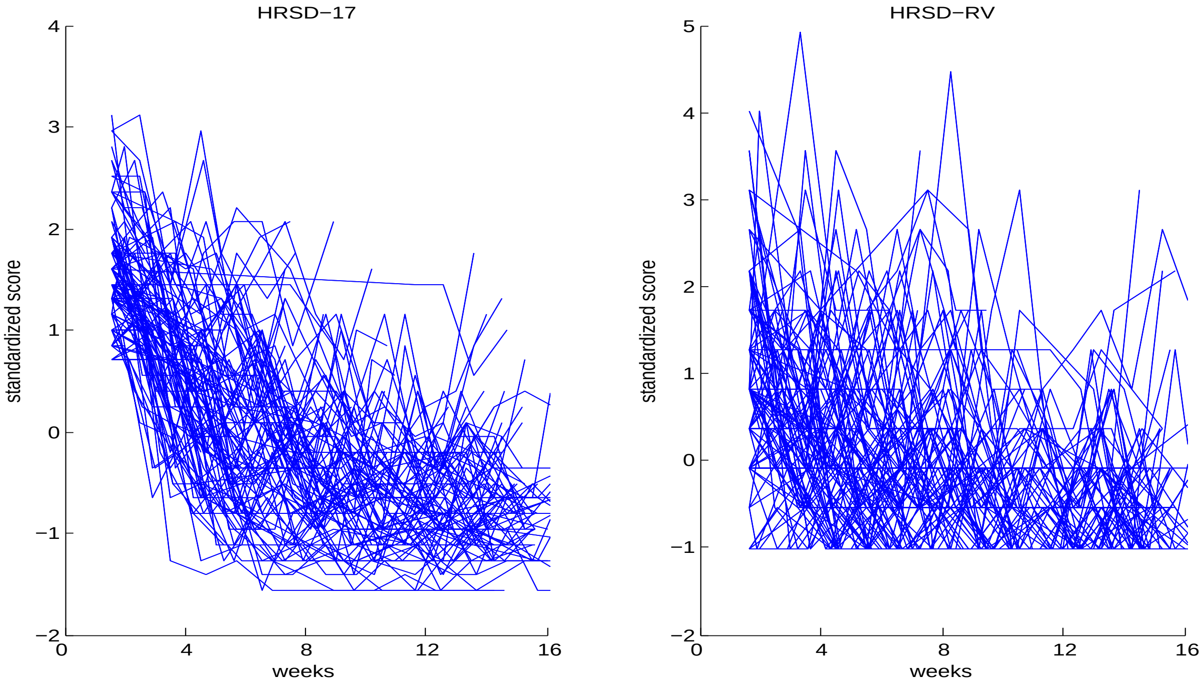
- Motivating example
- The model
- Estimation
- Example

Motivating Example

- Psychiatric study to compare psychotherapy to pharmacotherapy at the University of Pisa, Italy. Total of 103 acutely depressed subjects.

- Response variables: two subscales of the clinician-administered Hamilton Depression rating Scale.
  - HRSD-17 – measures anhedonia, sleep disturbance, agitation, anxiety, weight loss
  - HRSD-RV – measures reverse vegetative symptoms such as weight gain and hyperphagia

- Covariates: treatment group, lifetime depression spectrum


HRSD−17
standardized score
4
3
2
1
0
−1
−2
0
4
8
12
16
weeks
HRSD−RV
standardized score
5
4
3
2
1
0
−1
−2
0
4
8
12
16
weeks

The Model

$$\begin{aligned}
y_{i1}(t_{ij}) &= \boldsymbol{x}'_{ij}\boldsymbol{\mu_1}(t_{ij}) + \boldsymbol{z}'_{ij}\boldsymbol{g_{i1}}(t_{ij}) + \delta_{i1}(t_{ij}) \\
y_{i2}(t_{ij}) &= \boldsymbol{x}'_{ij}\boldsymbol{\mu_2}(t_{ij}) + \boldsymbol{z}'_{ij}\boldsymbol{g_{i2}}(t_{ij}) + \delta_{i2}(t_{ij}) \\
\vdots &= \vdots \\
y_{ip}(t_{ij}) &= \boldsymbol{x}'_{ij}\boldsymbol{\mu_p}(t_{ij}) + \boldsymbol{z}'_{ij}\boldsymbol{g_{ip}}(t_{ij}) + \delta_{ip}(t_{ij})
\end{aligned}$$

The Model

$$
\begin{aligned}
y_{i1}(t_{ij}) &= \boldsymbol{x}'_{ij}\boldsymbol{\mu_1}(t_{ij}) + \boldsymbol{z}'_{ij}\boldsymbol{g_{i1}}(t_{ij}) + \delta_{i1}(t_{ij}) \\
y_{i2}(t_{ij}) &= \boldsymbol{x}'_{ij}\boldsymbol{\mu_2}(t_{ij}) + \boldsymbol{z}'_{ij}\boldsymbol{g_{i2}}(t_{ij}) + \delta_{i2}(t_{ij}) \\
\vdots &= \vdots \\
y_{ip}(t_{ij}) &= \boldsymbol{x}'_{ij}\boldsymbol{\mu_p}(t_{ij}) + \boldsymbol{z}'_{ij}\boldsymbol{g_{ip}}(t_{ij}) + \delta_{ip}(t_{ij})
\end{aligned}
$$

- $\boldsymbol{\mu}_k(t) = (\mu_{k1}(t), \ldots, \mu_{kr}(t))'$ is an r-vector of fixed functions, $k = 1, \ldots, p$

The Model

$$\begin{aligned}
y_{i1}(t_{ij}) &= \boldsymbol{x}'_{ij}\boldsymbol{\mu_1}(t_{ij}) + \boldsymbol{z}'_{ij}\boldsymbol{g_{i1}}(t_{ij}) + \delta_{i1}(t_{ij}) \\
y_{i2}(t_{ij}) &= \boldsymbol{x}'_{ij}\boldsymbol{\mu_2}(t_{ij}) + \boldsymbol{z}'_{ij}\boldsymbol{g_{i2}}(t_{ij}) + \delta_{i2}(t_{ij}) \\
\vdots &= \vdots \\
y_{ip}(t_{ij}) &= \boldsymbol{x}'_{ij}\boldsymbol{\mu_p}(t_{ij}) + \boldsymbol{z}'_{ij}\boldsymbol{g_{ip}}(t_{ij}) + \delta_{ip}(t_{ij})
\end{aligned}$$

- $\boldsymbol{\mu}_k(t) = (\mu_{k1}(t), \ldots, \mu_{kr}(t))'$ is an r-vector of fixed functions, $k = 1, \ldots, p$
- $\boldsymbol{g}_{ik}(t) = (g_{ik1}(t), \ldots, g_{iks}(t))'$ is an s-vector of random functions, $k = 1, \ldots, p$

The Model (Cont.)

$$\boldsymbol{y}_i(t_{ij}) = X_{ij}\boldsymbol{\mu}(t_{ij}) + Z_{ij}\boldsymbol{g_i}(t_{ij}) + \boldsymbol{\delta}_i(t_{ij}) \ ,$$

- $\boldsymbol{y}_i(t_{ij}) = (y_{i1}(t_{ij}), \ldots, y_{ip}(t_{ij}))'$, $i = 1, \ldots, n$, $j = 1, \ldots, m_i$.

The Model (Cont.)

$$\boldsymbol{y}_i(t_{ij}) = X_{ij}\boldsymbol{\mu}(t_{ij}) + Z_{ij}\boldsymbol{g_i}(t_{ij}) + \boldsymbol{\delta}_i(t_{ij}) \ ,$$

- $\boldsymbol{y}_i(t_{ij}) = (y_{i1}(t_{ij}), \ldots, y_{ip}(t_{ij}))'$, $i = 1, \ldots, n$, $j = 1, \ldots, m_i$.
- $\boldsymbol{\mu}(t) = (\boldsymbol{\mu}_1'(t), \ldots, \boldsymbol{\mu}_p'(t))'$

The Model (Cont.)

$$\boldsymbol{y}_i(t_{ij}) = X_{ij}\boldsymbol{\mu}(t_{ij}) + Z_{ij}\boldsymbol{g_i}(t_{ij}) + \boldsymbol{\delta}_i(t_{ij})\ ,$$

- $\boldsymbol{y}_i(t_{ij}) = (y_{i1}(t_{ij}), \ldots, y_{ip}(t_{ij}))'$, $i = 1, \ldots, n$, $j = 1, \ldots, m_i$.
- $\boldsymbol{\mu}(t) = (\boldsymbol{\mu}_1'(t), \ldots, \boldsymbol{\mu}_p'(t))'$
- $\boldsymbol{g}_i(t) = (\boldsymbol{g}_{i1}'(t), \ldots, \boldsymbol{g}_{ip}'(t))'$

The Model (Cont.)

$$\boldsymbol{y}_i(t_{ij}) = X_{ij}\boldsymbol{\mu}(t_{ij}) + Z_{ij}\boldsymbol{g_i}(t_{ij}) + \boldsymbol{\delta}_i(t_{ij}) \ ,$$

- $\boldsymbol{y}_i(t_{ij}) = (y_{i1}(t_{ij}), \ldots, y_{ip}(t_{ij}))'$, $i = 1, \ldots, n$, $j = 1, \ldots, m_i$.
- $\boldsymbol{\mu}(t) = (\boldsymbol{\mu}_1'(t), \ldots, \boldsymbol{\mu}_p'(t))'$
- $\boldsymbol{g}_i(t) = (\boldsymbol{g}_{i1}'(t), \ldots, \boldsymbol{g}_{ip}'(t))'$
- $X_{ij} = I_p \otimes \boldsymbol{x}_{ij}'$, $Z_{ij} = I_p \otimes \boldsymbol{z}_{ij}'$.

The Model (Cont.)

$$\boldsymbol{y}_i(t_{ij}) = X_{ij}\boldsymbol{\mu}(t_{ij}) + Z_{ij}\boldsymbol{g_i}(t_{ij}) + \boldsymbol{\delta}_i(t_{ij}) \ ,$$

- $\boldsymbol{y}_i(t_{ij}) = (y_{i1}(t_{ij}), \ldots, y_{ip}(t_{ij}))'$, $i = 1, \ldots, n$, $j = 1, \ldots, m_i$.
- $\boldsymbol{\mu}(t) = (\boldsymbol{\mu}_1'(t), \ldots, \boldsymbol{\mu}_p'(t))'$
- $\boldsymbol{g}_i(t) = (\boldsymbol{g}_{i1}'(t), \ldots, \boldsymbol{g}_{ip}'(t))'$
- $X_{ij} = I_p \otimes \boldsymbol{x}_{ij}'$, $Z_{ij} = I_p \otimes \boldsymbol{z}_{ij}'$.
- $\boldsymbol{\delta}_i(t_{ij}) = (\delta_{i1}(t_{ij}), \ldots, \delta_{ip}(t_{ij}))'$

Example

Assume 2 groups of subjects and a bivariate response ($p = 2$):

$\boldsymbol{x}_{ij} = (x_{ij1}, x_{ij2})'$, where $x_{ij1} = 1, \quad x_{ij2} = \begin{cases} 1 & \text{if} \quad \text{treatment} \\ 0 & \text{if} \quad \text{control} \end{cases}$

Example

Assume 2 groups of subjects and a bivariate response ($p = 2$):

$\boldsymbol{x}_{ij} = (x_{ij1}, x_{ij2})'$, where $x_{ij1} = 1, \quad x_{ij2} = \begin{cases} 1 & \text{if} \quad \text{treatment} \\ 0 & \text{if} \quad \text{control} \end{cases}$

$$\begin{aligned} y_{i1}(t_{ij}) &= (1 \quad x_{ij2}) \begin{pmatrix} \mu_{11}(t_{ij}) \\ \mu_{12}(t_{ij}) \end{pmatrix} + g_{i1}(t_{ij}) + \delta_{i1}(t_{ij}) \\ y_{i2}(t_{ij}) &= (1 \quad x_{ij2}) \begin{pmatrix} \mu_{21}(t_{ij}) \\ \mu_{22}(t_{ij}) \end{pmatrix} + g_{i2}(t_{ij}) + \delta_{i2}(t_{ij}) \end{aligned}$$

Example

Assume 2 groups of subjects and a bivariate response ($p = 2$):

$\boldsymbol{x}_{ij} = (x_{ij1}, x_{ij2})'$, where $x_{ij1} = 1$, $\quad x_{ij2} = \begin{cases} 1 & \text{if} \quad \text{treatment} \\ 0 & \text{if} \quad \text{control} \end{cases}$

$$
\begin{aligned}
y_{i1}(t_{ij}) &= (1 \quad x_{ij2}) \begin{pmatrix} \mu_{11}(t_{ij}) \\ \mu_{12}(t_{ij}) \end{pmatrix} + g_{i1}(t_{ij}) + \delta_{i1}(t_{ij}) \\
y_{i2}(t_{ij}) &= (1 \quad x_{ij2}) \begin{pmatrix} \mu_{21}(t_{ij}) \\ \mu_{22}(t_{ij}) \end{pmatrix} + g_{i2}(t_{ij}) + \delta_{i2}(t_{ij})
\end{aligned}
$$

so

$$
\begin{aligned}
y_{i1}(t_{ij}) &= \mu_{11}(t_{ij}) + x_{ij2} \cdot \mu_{12}(t_{ij}) + g_{i1}(t_{ij}) + \delta_{i1}(t_{ij}) \\
y_{i2}(t_{ij}) &= \mu_{21}(t_{ij}) + x_{ij2} \cdot \mu_{22}(t_{ij}) + g_{i2}(t_{ij}) + \delta_{i2}(t_{ij})
\end{aligned}
$$

The Error Term

$\boldsymbol{\delta}_i(t)$ follows a multivariate Ornstein-Uhlenbeck (O-U) process, i.e.,

$$d\boldsymbol{\delta}_i(t) = -A\boldsymbol{\delta}_i(t) + Bd\boldsymbol{W}_i(t) \ .$$

- A and B are $p \times p$ matrices of full rank common to all $i = 1, \ldots, n$,
- $\boldsymbol{W}_i(t)$ is the p-dimensional Wiener process.
- Stationary if the EV of A have positive real parts.
- The stationary variance-covariance matrix Σ is the solution to

$$A\Sigma + \Sigma A' = BB'$$

The Error Term (Cont.)

In the stationary state, for $s < t$

$$\mathbb{C}\mathrm{ov}(\boldsymbol{\delta}_i(t), \boldsymbol{\delta}_i(s)) = \exp\{-A(t-s)\}\Sigma \ ,$$

so A controls how rapidly the influence of time s dies off by time $t > s$.

The Transition Density

Let $\Delta t_{ij} = t_{ij} - t_{i,j-1}\ j = 1, \ldots, m_i$, $t_{i0} = 0$.

The transition density of the O-U process is given by

$$p(\boldsymbol{\delta}_i(t_{ij}) | \boldsymbol{\delta}_i(t_{i,j-1}), \Delta t_{ij}) \propto |\Omega_{\Delta t_{ij}}|^{-1/2} \exp\left\{ -\frac{1}{2} \boldsymbol{\gamma}'_{t_{ij}} \Omega^{-1}_{\Delta t_{ij}} \boldsymbol{\gamma}_{t_{ij}} \right\} \ ,$$

where

$$\begin{aligned} \boldsymbol{\gamma}_{t_{ij}} &= \boldsymbol{\delta}_i(t_{ij}) - \exp(-A\Delta t_{ij}) \boldsymbol{\delta}_i(t_{i,j-1}) \\ \Omega_{\Delta t_{ij}} &= \Sigma - \exp(-A\Delta t_{ij}) \Sigma \exp(-A'\Delta t_{ij}). \end{aligned}$$

Nonparametric Function Estimation– Simple Example

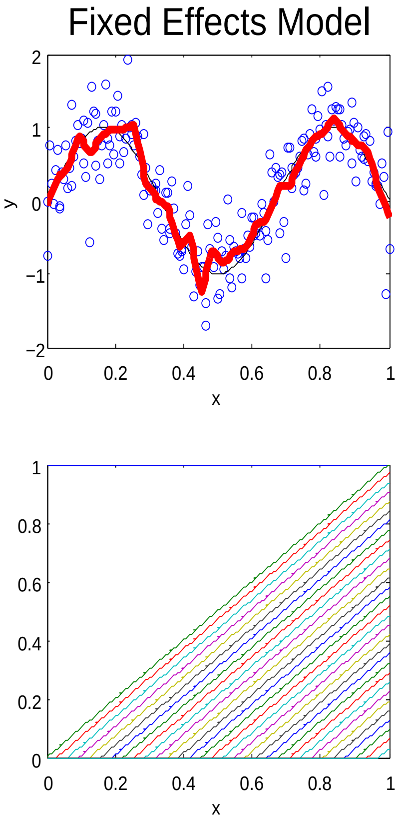
200 observations were generated from

$$y_i = \sin(3\pi x_i) + \epsilon_i, \quad , 0 \leq x_i \leq 1, \quad \epsilon_i \sim N(0, 0.4^2)$$

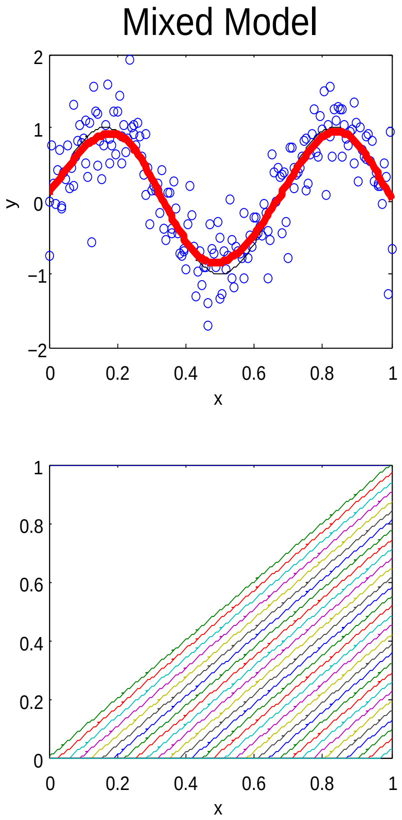
Model:

$$y_i = \beta_0 + \beta_1 x_i + \sum_{k=1}^{30} u_k (x_i - \kappa_k)_+ + \epsilon_i, \quad \epsilon_i \sim N(0, \sigma_\epsilon^2)$$

- Low rank truncated line basis functions.
- Fixed effects: βs and us fixed.
- Mixed effects: βs fixed, $u_1, \ldots, u_{30} \overset{iid}{\sim} N(0, \sigma_u^2)$.



Fixed Effects Model



Mixed Model

Nonparametric Function Estimation (Cont.)

We estimate

$\mu_{kl}(t)$ and $g_{ikm}(t), k = 1, \ldots, p, l = 1, \ldots, r, m = 1, \ldots, s, i = 1, \ldots, n$

by cubic splines using low-rank radial basis functions (French et al., 2001, Ruppert et al., 2003):

Write $\mu_{kl}(t_{ij}) = \boldsymbol{\phi}'_{ij}\boldsymbol{\beta}_{kl} + \boldsymbol{\psi}'_{ij}\boldsymbol{v}_{kl}$ and $g_{ikm}(t_{ij}) = \boldsymbol{\phi}'_{ij}\boldsymbol{w}_{ikm} + \boldsymbol{\psi}'_{ij}\boldsymbol{u}_{ikm}$.

Nonparametric Function Estimation (Cont.)

We estimate

$\mu_{kl}(t)$ and $g_{ikm}(t), k = 1, \ldots, p, l = 1, \ldots, r, m = 1, \ldots, s, i = 1, \ldots, n$

by cubic splines using low-rank radial basis functions (French et al., 2001, Ruppert et al., 2003):

Write $\mu_{kl}(t_{ij}) = \boldsymbol{\phi}'_{ij}\boldsymbol{\beta}_{kl} + \boldsymbol{\psi}'_{ij}\boldsymbol{v}_{kl}$ and $g_{ikm}(t_{ij}) = \boldsymbol{\phi}'_{ij}\boldsymbol{w}_{ikm} + \boldsymbol{\psi}'_{ij}\boldsymbol{u}_{ikm}$.

- Place K knots $\kappa_1, \ldots, \kappa_K$ at sample quantiles of t_{ij}.

Nonparametric Function Estimation (Cont.)

We estimate

$$\mu_{kl}(t) \quad \text{and} \quad g_{ikm}(t), k = 1, \ldots, p, l = 1, \ldots, r, m = 1, \ldots, s, i = 1, \ldots, n$$

by cubic splines using low-rank radial basis functions (French et al., 2001, Ruppert et al., 2003):

Write $\mu_{kl}(t_{ij}) = \boldsymbol{\phi}'_{ij}\boldsymbol{\beta}_{kl} + \boldsymbol{\psi}'_{ij}\boldsymbol{v}_{kl}$ and $g_{ikm}(t_{ij}) = \boldsymbol{\phi}'_{ij}\boldsymbol{w}_{ikm} + \boldsymbol{\psi}'_{ij}\boldsymbol{u}_{ikm}$.

- Place K knots $\kappa_1, \ldots, \kappa_K$ at sample quantiles of t_{ij}.
- Let $\boldsymbol{\phi}'_{ij} = (1 \quad t_{ij})$, $\boldsymbol{\xi}'_{ij} = (|t_{ij} - \kappa_1|^3, \ldots, |t_{ij} - \kappa_K|^3)$,
- Compute $\Lambda_K = [|\kappa_k - \kappa_{k'}|^3]_{1 \le k, k' \le K}$ and $\boldsymbol{\psi}'_{ij} = \boldsymbol{\xi}'_{ij}\Lambda_K^{-1/2}$ (via SVD)

Prior Distributions

Priors on the basis function coefficients:

1. Flat priors on $\boldsymbol{\beta}_{kl}$.
2. $\boldsymbol{v}_{kl} \overset{ind}{\sim} N(\mathbf{0}, \sigma^2_{v_{kl}} I_K)$
3. $\boldsymbol{w}_{ikm} \overset{ind}{\sim} N(\mathbf{0}, \text{diag}(\sigma^2_{w_{km0}}, \sigma^2_{w_{km1}}))$.
4. $\boldsymbol{u}_{ikm} \overset{ind}{\sim} N(\mathbf{0}, \sigma^2_{u_{km}} I_K)$.

Priors on the variance components:
independent $IG(a_i, b_i)$, e.g., $p(\sigma^2_{v_{kl}}) \propto (\sigma^2_{v_{kl}})^{-(a_1+1)} \exp(-b_1/\sigma^2_{v_{kl}})$

Prior Distributions (Cont.)

Prior on A:

- Let $A = S\Psi S^{-1}$, where

$$S = \begin{pmatrix} 1 & s_{12} & s_{13} & \dots & s_{1p} \\ s_{21} & 1 & s_{23} & \dots & s_{2p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ s_{p1} & s_{p2} & s_{p3} & \dots & 1 \end{pmatrix}$$

  and $\Psi = \text{diag}(\psi_1, \dots, \psi_p)$, $\psi_i > 0$, $i = 1, \dots, p$.

- $s_{ij} \sim N(0, \sigma_s^2)$, $\log \psi_i \sim N(0, \sigma_\psi^2)$

Prior Distributions (Cont.)

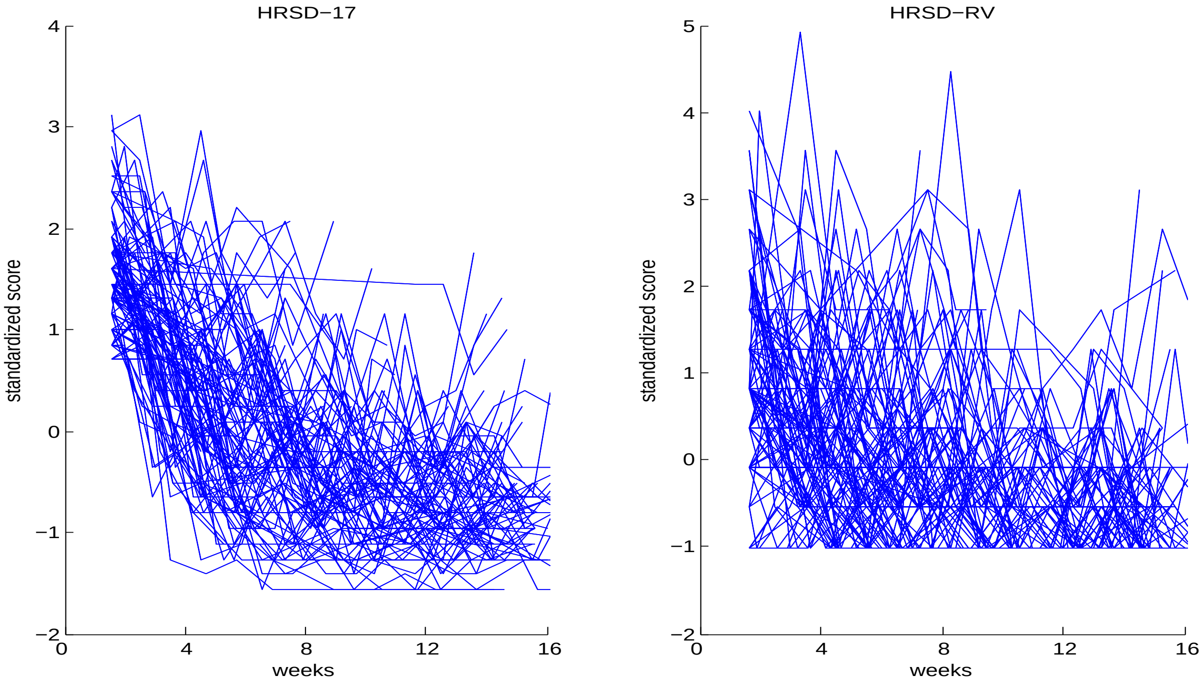
Prior on $C = BB'$:

- Let $C = LDL'$ (modified Cholesky factorization), where

$$L = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ l_{21} & 1 & 0 & \dots & 0 \\ l_{31} & l_{32} & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ l_{p1} & l_{p2} & l_{p3} & \dots & 1 \end{pmatrix}$$

  and $D = \text{diag}(d_1, \dots, d_p)$, $d_i > 0$, $i = 1, \dots, p$

- $l_{ij} \sim N(0, \sigma_L^2)$, $\log(d_i) \sim N(0, \sigma_D^2)$


HRSD−17
HRSD−RV
standardized score
standardized score
weeks
weeks
4
3
2
1
0
−1
−2
5
4
3
2
1
0
−1
−2
0
4
8
12
16
0
4
8
12
16

Example

- Psychiatric study to compare psychotherapy to pharmacotherapy at the University of Pisa, Italy. Total of 103 acutely depressed subjects.

- Response variables: 2 subscales of the clinician-administered Hamilton Depression rating Scale (measured at baseline and weekly):

  HRSD-17 (0–31)
  HRSD-RV (0–13).

  Higher values indicate more depressive symptoms.

- Covariates:

  treatment group
  lifetime depression spectrum (LDS), assessed at baseline.

Example (Cont.)

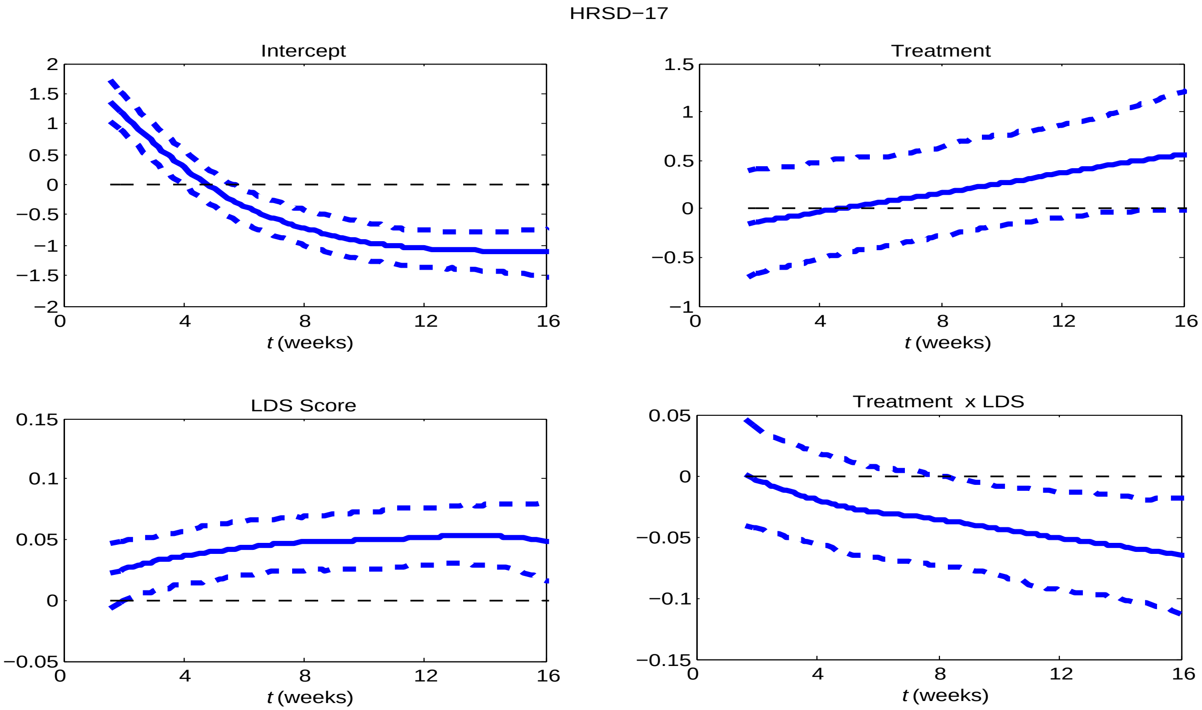
$$\boldsymbol{y}_i(t_{ij}) = X_{ij}\boldsymbol{\mu}(t_{ij}) + Z_{ij}\boldsymbol{g_i}(t_{ij}) + \boldsymbol{\delta}_i(t_{ij}) \ ,$$

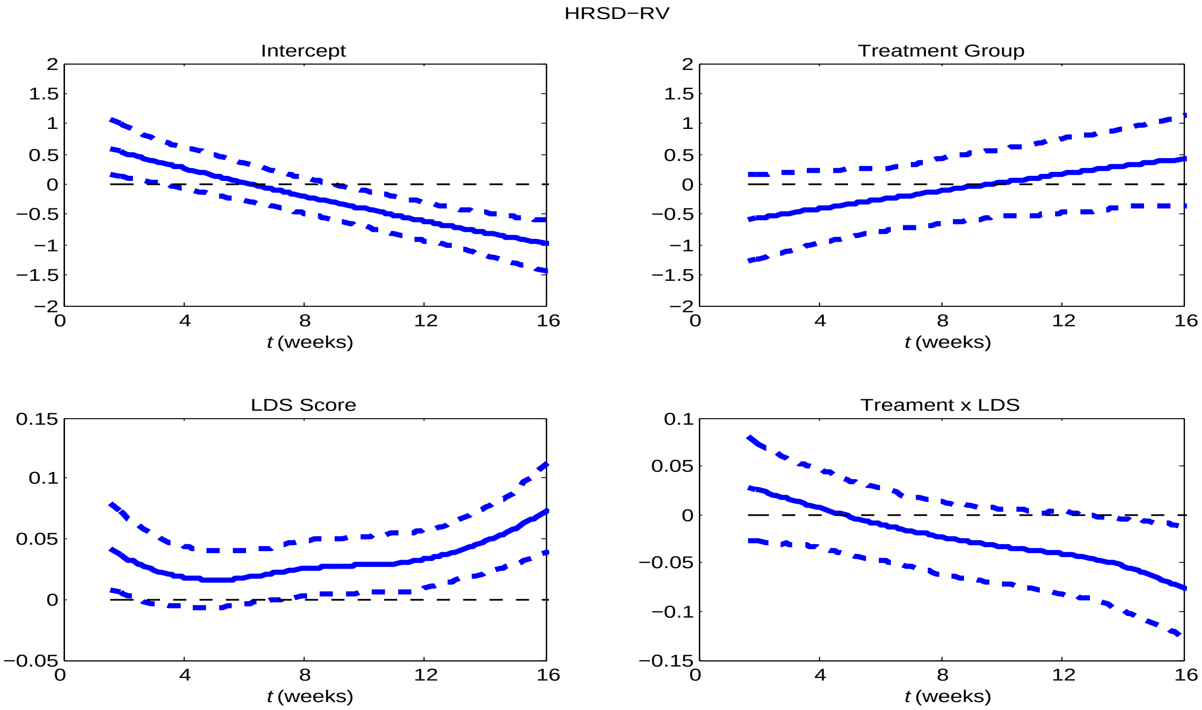
where $X_{ij} = I_p \otimes \boldsymbol{x}'_{ij}$, $Z_{ij} = I_p \otimes \boldsymbol{z}'_{ij}$.

In the example $\boldsymbol{y}_i(t_{ij}) = (\text{HRSD-17}, \text{HRSD-RV})'$
$\boldsymbol{z}_{ij} = 1$
$\boldsymbol{x}_{ij} = (x_{ij1}, x_{ij2}, x_{ij3}, x_{ij4})'$
where $x_{ij1} = 1$

$$x_{ij2} = \begin{cases} 1 & \text{if subject } i \text{ received psychotherapy} \\ 0 & \text{if subject } i \text{ received pharmacotherapy,} \end{cases}$$

$x_{ij3} = \text{LDS score}$
$x_{ij4} = x_{ij2}x_{ij3}$.


HRSD−17
Intercept
2
1.5
1
0.5
0
−0.5
−1
−1.5
−2
0
4
8
12
16
t (weeks)
Treatment
1.5
1
0.5
0
−0.5
−1
0
4
8
12
16
t (weeks)
LDS Score
0.15
0.1
0.05
0
−0.05
0
4
8
12
16
t (weeks)
Treatment x LDS
0.05
0
−0.05
−0.1
−0.15
0
4
8
12
16
t (weeks)


HRSD−RV

Intercept

Treatment Group

LDS Score

Treament x LDS

Conclusions from the Analysis

- Little evidence of a treatment effect
- Significant effect of LDS score – higher LDS predicts worse outcomes
- LDS effect is more pronounced in the middle period of HRSD-17, and towards the end of the 16 weeks for HRSD-RV.
- Interaction is significant in the last 8 weeks of HRSD-17, and marginally significant in the same time period for HRSD-RV.
- Interaction effect – LDS is less predictive of poor response in psychotherapy group than in pharmacotherapy group.
- Univariate fitting – interaction is not significant; wider pointwise credible intervals.