State Space Models for Longitudinal Neuropsychological Outcomes

Zach Baucom

Section 1

Recap of Project 1

The model

We wish to model the data according to a specific SSM, the Local Linear Trend Model (LLT),

$$y_{ij} = \alpha_{ij} + x_{ij}^T \beta + \varepsilon_{ij}$$
$$\alpha_{ij} = \alpha_{i(j-1)} + \eta_{ij}$$

Where $\alpha_0 \sim N(a_0, P_0)$, $\varepsilon_{ij} \sim N(0, \sigma_\varepsilon^2)$, and $\eta_{ij} \sim N(0, \delta_{ij}\sigma_\eta^2)$.

- y_j is an $n \times 1$ observation vector where n indicates the number of subjects.
- α_j is an $n \times 1$ latent state vector.
  - Variation in α_j over time creates a dynamic moving average auto-correlation between observations y_j.
- X_j is an $n \times p$ matrix of time varying covariates.
- β is the linear effect of the columns in X_j.

Results

- The LLT models did a much better job at measuring linear effects than the commonly used linear mixed effect models.
- Among the LLTs, the Bayesian Estimation Process yielded the best results in terms of accuracy and computation time.
- Paper 1 is currently under review at the Annals of Applied Statistics.

Section 2

Project 2 Multivatiate Models

Aims of Project 2

- Cognitive tests, offered by the NACC, are administered to measure different underlying constructs of cognition.
  - e.g. memory, attention, executive, and language.
- Understanding how the tests are related to different constructs is vital for interpreting effects of interest.
- In order to create a model capable of accounting for inter-relatedness between tests, we propose a Multivariate Bayesian Local Linear Trend Model.

Aims

- With the Multivariate Bayesian Local Linear Trend Model we wish to:
  - Gain power by taking advantage of correlations between tests.
  - Get insight into relatedness of underlying cognition levels from each test.
  - Be able to compare linear effects across tests (i.e. does APOE have the same effect for test A and test B?)

The Model

$$\begin{bmatrix} y_{ij1} \\ y_{ij2} \\ \vdots \\ y_{ijK} \end{bmatrix} = \begin{bmatrix} \alpha_{ij1} \\ \alpha_{ij2} \\ \vdots \\ \alpha_{ijK} \end{bmatrix} + \begin{bmatrix} \boldsymbol{x_{ij}}\boldsymbol{\beta_1} \\ \boldsymbol{x_{ij}}\boldsymbol{\beta_2} \\ \vdots \\ \boldsymbol{x_{ij}}\boldsymbol{\beta_K} \end{bmatrix} + \begin{bmatrix} \varepsilon_{ij1} \\ \varepsilon_{ij2} \\ \vdots \\ \varepsilon_{ijK} \end{bmatrix}, \begin{bmatrix} \varepsilon_{ij1} \\ \varepsilon_{ij2} \\ \vdots \\ \varepsilon_{ijK} \end{bmatrix} \sim N(0, \Sigma_\varepsilon)$$

$$\begin{bmatrix} \alpha_{ij1} \\ \alpha_{ij2} \\ \vdots \\ \alpha_{ijK} \end{bmatrix} = \begin{bmatrix} \alpha_{i(j-1)1} \\ \alpha_{i(j-1)2} \\ \vdots \\ \alpha_{i(j-1)K} \end{bmatrix} + \begin{bmatrix} \eta_{ij1} \\ \eta_{ij2} \\ \vdots \\ \eta_{ijK} \end{bmatrix}, \begin{bmatrix} \eta_{ij1} \\ \eta_{ij2} \\ \vdots \\ \eta_{ijK} \end{bmatrix} \sim N(0, \delta_{ij}\Sigma_\eta)$$

Model 1: Correlation is in Measurement eq.

$$\Sigma_\varepsilon = \begin{bmatrix} \sigma_\varepsilon^{2(1)} & \sigma_\varepsilon^{2(1,2)} & \cdots & \sigma_\varepsilon^{2(1,K)} \\ \sigma_\varepsilon^{2(1,2)} & \sigma_\varepsilon^{2(2)} & & \\ \vdots & & \ddots & \\ \sigma_\varepsilon^{2(1,K)} & & & \sigma_\varepsilon^{2(K)} \end{bmatrix}, \Sigma_\eta = \begin{bmatrix} \sigma_\eta^{2(1)} & 0 & \cdots & 0 \\ 0 & \sigma_\eta^{2(2)} & & \\ \vdots & & \ddots & \\ 0 & & & \sigma_\eta^{2(K)} \end{bmatrix}$$

Model 2: Correlation is in State eq.

$$\Sigma_{\varepsilon} = \begin{bmatrix} \sigma_{\varepsilon}^{2(1)} & 0 & \cdots & 0 \\ 0 & \sigma_{\varepsilon}^{2(2)} & & \\ \vdots & & \ddots & \\ 0 & & & \sigma_{\varepsilon}^{2(K)} \end{bmatrix}, \Sigma_{\eta} = \begin{bmatrix} \sigma_{\eta}^{2(1)} & \sigma^{2(1,2)} & \cdots & \sigma_{\eta}^{2(1,K)} \\ \sigma_{\eta}^{2(1,2)} & \sigma_{\eta}^{2(2)} & & \\ \vdots & & \ddots & \\ \sigma_{\eta}^{2(1,K)} & & & \sigma_{\eta}^{2(K)} \end{bmatrix}$$

Model 3: Correlation is in both Measurement & State eq.

$$\Sigma_\varepsilon = \begin{bmatrix} \sigma_\varepsilon^{2(1)} & \sigma_\varepsilon^{2(1,2)} & \cdots & \sigma_\varepsilon^{2(1,K)} \\ \sigma_\varepsilon^{2(1,2)} & \sigma_\varepsilon^{2(2)} & & \\ \vdots & & \ddots & \\ \sigma_\varepsilon^{2(1,K)} & & & \sigma_\varepsilon^{2(K)} \end{bmatrix}, \Sigma_\eta = \begin{bmatrix} \sigma_\eta^{2(1)} & \sigma^{2(1,2)} & \cdots & \sigma_\eta^{2(1,K)} \\ \sigma_\eta^{2(1,2)} & \sigma_\eta^{2(2)} & & \\ \vdots & & \ddots & \\ \sigma_\eta^{2(1,K)} & & & \sigma_\eta^{2(K)} \end{bmatrix}$$

Kalman Filter

For j in $1, 2, ..., J$:

1. Predicted state: $\boldsymbol{\alpha}_{ij|i(j-1)} = \boldsymbol{\alpha}_{i(j-1)|i(j-1)}$
2. Predicted state variance: $\boldsymbol{P_{ij|i(j-1)} = P_{i(j-1)|i(j-1)}} + \sigma^2_\eta$
3. Kalman Gain: $\boldsymbol{K_{ij}} = \boldsymbol{P_{ij|i(j-1)}(P_{ij|i(j-1)} + \sigma^2_\varepsilon})^{-1}$
4. Updated state estimate: $\boldsymbol{\alpha}_{ij|ij} = \boldsymbol{\alpha}_{ij|i(j-1)} + \boldsymbol{K_{ij}}(\boldsymbol{\tilde{y}_{ij} - \alpha_{ij|i(j-1)}})$
5. Updated state covariance: $\boldsymbol{P_{ij|ij}} = (1 - \boldsymbol{K_{ij}})\boldsymbol{P_{ij|i(j-1)}}$

Kalman Smoother

For j^* in J, J-1, . . . , 1:

1. Smoothed predicted state:
   $\boldsymbol{\alpha}_{iJ|i(j^*-1)} = \boldsymbol{\alpha}_{i(j^*-1)|i(j^*-1)} + \boldsymbol{L}_{i(j^*-1)}(\boldsymbol{\alpha}_{ij^*|iJ} - \boldsymbol{\alpha}_{ij^*|i(j^*-1)})$
2. Smoothed predicted state variance:
   $\boldsymbol{P}_{\boldsymbol{i(j^*-1)|iJ}} = \boldsymbol{P}_{\boldsymbol{i(j^*-1)|i(j^*-1)}} - \boldsymbol{L}^2_{\boldsymbol{i(j^*-1)}}(\boldsymbol{P}_{\boldsymbol{ij^*|iJ}} - \boldsymbol{P}_{\boldsymbol{ij^*|i(j^*-1)}})$

Where $\boldsymbol{L}_{\boldsymbol{i(j^*-1)}} = \boldsymbol{P}_{\boldsymbol{i(j^*-1)|i(j^*-1)}} \boldsymbol{P}^{-1}_{\boldsymbol{ij^*|i(j^*-1)}}$.

Backward Sampler

Let $\psi = \{\Sigma_\eta, \Sigma_\varepsilon, \boldsymbol{\beta}\}$,

$$P_\psi(\boldsymbol{\alpha}_{i(0:J)}|\boldsymbol{y}_{i(1:J)}) = P_\psi(\boldsymbol{\alpha}_{iJ}|\boldsymbol{y}_{i(1:J)})P_\psi(\boldsymbol{\alpha}_{i(J-1)}|\boldsymbol{\alpha}_{i(J)}, \boldsymbol{y}_{1:(J-1)})...P_\psi(\boldsymbol{\alpha}_{i0}|\boldsymbol{\alpha}_{i1})$$

Therefore we need the following densities for j in $1, 2, ..., J-1$ and i in $1, 2, ..., N$:

$$P_\psi(\boldsymbol{\alpha}_{ij}|\boldsymbol{\alpha}_{i(j+1)}, \boldsymbol{y}_{i(1:j)}) \propto P_\psi(\boldsymbol{\alpha}_{ij}|\boldsymbol{y}_{i(1:j)})P_\psi(\boldsymbol{\alpha}_{i(j+1)}|\boldsymbol{\alpha}_{ij})$$

Backward Sampler

- From the Kalman Filter we calculate $\boldsymbol{\alpha}_{ij}|\boldsymbol{y}_{i(1:j)} \sim N_\psi(\boldsymbol{\alpha}_{ij|ij}, \boldsymbol{P}_{ij|ij})$ and $\boldsymbol{\alpha}_{i(j+1)}|\boldsymbol{\alpha}_{ij} \sim N_\psi(\boldsymbol{\alpha}_{ij}, \Sigma_\eta)$.
- After combining the two densities $\boldsymbol{m}_{ij} = E_\psi(\boldsymbol{\alpha}_{ij}|\boldsymbol{\alpha}_{i(j+1)}, \boldsymbol{y}_{i(1:j)}) = \boldsymbol{\alpha}_{ij|ij} + \boldsymbol{L}_{ij}(\boldsymbol{\alpha}_{i(j+1)} - \boldsymbol{\alpha}_{i(j+1)|ij})$ and $\boldsymbol{R}_{ij} = \mathrm{Var}_\psi(\boldsymbol{\alpha}_{ij}|\boldsymbol{\alpha}_{i(j+1)}, \boldsymbol{y}_{i(1:j)}) = \boldsymbol{P}_{ij|ij} - \boldsymbol{L}_{ij}^2\boldsymbol{P}_{i(j+1)|ij}$.
- Because of normality, the posterior distribution for $\boldsymbol{\alpha}_{ij}$ is $N(\boldsymbol{m}_{ij}, \boldsymbol{R}_{ij})$.

β Posterior

$$
\begin{aligned}
P(Y, \boldsymbol{\alpha}, \Sigma_\eta, \Sigma_\varepsilon | \beta) &= P(\boldsymbol{y}_1, ..., \boldsymbol{y}_J, \boldsymbol{\alpha}_1, ..., \boldsymbol{\alpha}_J, \Sigma_\eta, \Sigma_\varepsilon | \beta) \\
=& P(\boldsymbol{y}_J | \boldsymbol{y}_1, ..., \boldsymbol{y}_{T-1}, \boldsymbol{\alpha}_1, ..., \boldsymbol{\alpha}_J, \Sigma_\eta, \Sigma_\varepsilon \beta) \\
& \times P(\boldsymbol{y}_1, ..., \boldsymbol{y}_{T-1}, \boldsymbol{\alpha}_1, ..., \boldsymbol{\alpha}_J, \Sigma_\eta, \Sigma_\varepsilon | \beta) \\
=& P(\boldsymbol{y}_J | \boldsymbol{\alpha}_J, \Sigma_\varepsilon \beta) P(\boldsymbol{\alpha}_{T-1} | \boldsymbol{y}_1, ..., \boldsymbol{y}_{T-1}, \boldsymbol{\alpha}_1, ..., \boldsymbol{\alpha}_{T-1}, \Sigma_\eta, \Sigma_\varepsilon, \beta) \\
& \times P(\boldsymbol{y}_1, ..., \boldsymbol{y}_{T-1}, \boldsymbol{\alpha}_1, ..., \boldsymbol{\alpha}_{T-1}, \Sigma_\eta, \Sigma_\varepsilon | \beta) \\
=& P(\boldsymbol{y}_J | \boldsymbol{\alpha}_J, \Sigma_\varepsilon \beta) P(\boldsymbol{\alpha}_{T-1} | \boldsymbol{\alpha}_{T-1}, \Sigma_\eta) \\
& \times P(\boldsymbol{y}_1, ..., \boldsymbol{y}_{T-1}, \boldsymbol{\alpha}_1, ..., \boldsymbol{\alpha}_{T-1}, \Sigma_\eta, \Sigma_\varepsilon | \beta) \\
=& P(\Sigma_\varepsilon) P(\Sigma_\eta) \left(\prod_{k=0}^{J} P(\boldsymbol{\alpha}_k | \boldsymbol{\alpha}_{k-1}, \Sigma_\eta) \right) \prod_{j=1}^{J} P(\boldsymbol{y}_j | \boldsymbol{\alpha}_j, \Sigma_\varepsilon, \beta) \\
\propto & \prod_{j=1}^{J} P(\boldsymbol{y}_j | \boldsymbol{\alpha}_j, \Sigma_\varepsilon, \beta)
\end{aligned}
$$

β Posterior

Due to independence from subject to subject we can write the -2 log likelihood as,

$$-2logP(Y, \boldsymbol{\alpha}, \Sigma_\eta, \Sigma_\varepsilon|\boldsymbol{\beta}) \propto \sum_{i=1}^{N}\sum_{j=1}^{T}(\boldsymbol{y}_{ij} - \boldsymbol{\alpha}_{ij} - \mathbb{X}_{ij}\boldsymbol{\beta})'\Sigma_\varepsilon^{-1}(\boldsymbol{y}_{ij} - \boldsymbol{\alpha}_{ij} - \boldsymbol{\beta})$$

where,

$$y_j = \begin{bmatrix} y_{ij1} \\ y_{ij2} \\ \vdots \\ y_{ijK} \end{bmatrix}, \alpha_j = \begin{bmatrix} \alpha_{ij1} \\ \alpha_{ij2} \\ \vdots \\ \alpha_{ijK} \end{bmatrix}, \mathbb{X}_{ij} = \begin{bmatrix} \boldsymbol{x}_{ij} & 0 & \dots & 0 \\ 0 & \boldsymbol{x}_{ij} & & \\ \vdots & & \ddots & \\ 0 & & & \boldsymbol{x}_{ij} \end{bmatrix} \beta = \begin{bmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \\ \vdots \\ \boldsymbol{\beta}_K \end{bmatrix}$$

β Posterior

Next, we write the -2 log likelihood in terms of β,

$$-2logP(Y, \boldsymbol{\alpha}, \sigma_\eta^2, \sigma_\varepsilon^2 | \boldsymbol{\beta}) \propto \boldsymbol{\beta}'(\sum_{i=1}^{N}\sum_{j=1}^{T} \mathbb{X}_{ij}' \Sigma_\varepsilon^{-1} \mathbb{X}_{ij})\boldsymbol{\beta} - 2\sum_{i=1}^{N}\sum_{j=1}^{T}(y_{ij} - \alpha_{ij})' \Sigma_\varepsilon^{-1} \mathbb{X}_{ij}\boldsymbol{\beta}$$

Similarly, the prior can be written in a similar format,

$$\begin{aligned}-2logP(\boldsymbol{\beta}) &\propto (\boldsymbol{\beta} - \theta)'\Sigma_\beta^{-1}(\boldsymbol{\beta} - \theta) \\ &\propto \boldsymbol{\beta}'\Sigma_\beta^{-1}\boldsymbol{\beta} - 2\theta\Sigma_\beta^{-1}\boldsymbol{\beta}\end{aligned}$$

β Posterior

Combining the two we are able to get the proportionality of the posterior,

$$\beta'(\sum_{i=1}^{N}\sum_{j=1}^{T}\mathbb{X}_{ij}'\Sigma_{\varepsilon}^{-1}\mathbb{X}_{ij})\beta - 2\sum_{i=1}^{N}\sum_{j=1}^{T}(y_{ij}-\alpha_{ij})'\Sigma_{\varepsilon}^{-1}\mathbb{X}_{ij}\beta + \beta'\Sigma_{\beta}^{-1}\beta - 2\theta\Sigma_{\beta}^{-1}\beta$$

$$= \beta'(\sum_{i=1}^{N}\sum_{j=1}^{T}\mathbb{X}_{ij}'\Sigma_{\varepsilon}^{-1}\mathbb{X}_{ij} + \Sigma_{\beta}^{-1})\beta - 2(\sum_{i=1}^{N}\sum_{j=1}^{T}(y_{ij}-\alpha_{ij})'\Sigma_{\varepsilon}^{-1}\mathbb{X}_{ij} + \theta\Sigma_{\beta}^{-1})\beta$$

β Posterior

Which, after completing the square, has the form

$$(\boldsymbol{\beta} - \boldsymbol{\Sigma}^{-1}B)'\boldsymbol{\Sigma}(\boldsymbol{\beta} - \boldsymbol{\Sigma}^{-1}B)$$

Where, $B = \sum_{i=1}^{N}\sum_{j=1}^{T}(y_{ij} - \alpha_{ij})'\Sigma_{\varepsilon}^{-1}\mathbb{X}_{ij} + \theta\Sigma_{\beta}^{-1}$ and $\boldsymbol{\Sigma} = \sum_{i=1}^{N}\sum_{j=1}^{T}\mathbb{X}_{ij}'\Sigma_{\varepsilon}^{-1}\mathbb{X}_{ij} + \Sigma_{\beta}^{-1}p$.

Thus, the posterior for β has the following distribution,

$$\boldsymbol{\beta}|... \sim N(\Sigma^{-1}B, \Sigma^{-1})$$

Σ_ε Posterior

For the posterior of Σ_ε, if we assume the matrix is non-diagnol as in Model 1, then we give Σ_ε an inverse-Wishart distribution.

$$\Sigma_\varepsilon \sim \text{Inv-Wishart}_{\nu_\varepsilon}(\Lambda_\varepsilon)$$
$$P(\Sigma_\varepsilon) \propto |\Sigma_\varepsilon|^{-(\nu_\varepsilon+p+1)/2} \exp(-tr(\Lambda_\varepsilon \Sigma_\varepsilon^{-1}))$$

$$P(Y, \boldsymbol{\alpha}, \Sigma_\eta, \beta | \Sigma_\varepsilon) \propto$$
$$|\Sigma_\varepsilon|^{-(NJ)/2} \exp(-\sum_{i=1}^{N}\sum_{j=1}^{J}(\boldsymbol{y}_{ij} - \boldsymbol{\alpha}_{ij} - \mathbb{X}_{ij}\boldsymbol{\beta})'\Sigma_\varepsilon^{-1}(\boldsymbol{y}_{ij} - \boldsymbol{\alpha}_{ij} - \mathbb{X}_{ij}\boldsymbol{\beta})/2$$

Σ_ε Posterior

$$P(\Sigma_\varepsilon | Y, \boldsymbol{\alpha}, \Sigma_\eta, \boldsymbol{\beta}) = P(Y, \boldsymbol{\alpha}, \Sigma_\eta, \boldsymbol{\beta} | \Sigma_\varepsilon) P(\Sigma_\varepsilon)$$

$$\propto |\Sigma_\varepsilon|^{-(NJ+\nu_\varepsilon+p+1)/2} \exp(-tr(\Lambda_\varepsilon + \sum_{i=1}^{N}\sum_{j=1}^{J}(\boldsymbol{y}_{ij} - \boldsymbol{\alpha}_{ij} - \mathbb{X}_{ij}\boldsymbol{\beta})^2)\Sigma_\varepsilon^{-1})$$

$$\Sigma_\varepsilon | \ldots \sim \text{Inv-Wishart}_{\nu_\varepsilon + NJ}(\Lambda_\varepsilon + \sum_{i=1}^{N}\sum_{j=1}^{J}(\boldsymbol{y}_{ij} - \boldsymbol{\alpha}_{ij} - \mathbb{X}_{ij}\boldsymbol{\beta})^2)$$

The posterior for the diagnol matrix Σ_ε can be calculated independently for each test as was previously shown [Paper1] using the inverse-gamma for each test.

Σ_η Posterior

Similarly, if Σ_η is non-diagonal:

$$\Sigma_\eta \sim \text{Inv-Wishart}_{\nu_\eta}(\Lambda_\eta)$$
$$P(\Sigma_\eta) \propto |\Sigma_\eta|^{-(\nu_\eta+p+1)/2} \exp(-tr(\Lambda_\eta \Sigma_\eta^{-1}))$$

$$P(Y, \boldsymbol{\alpha}, \Sigma_\eta, \boldsymbol{\beta} | \Sigma_\eta) \propto$$
$$|\Sigma_\eta|^{-(N(J-1))/2} \exp(-\sum_{i=1}^{N}\sum_{j=2}^{J}(\boldsymbol{\alpha}_{ij} - \boldsymbol{\alpha}_{i(j-1)})' \Sigma_\eta^{-1} (\boldsymbol{\alpha}_{ij} - \boldsymbol{\alpha}_{i(j-1)})/2$$

Σ_η Posterior

$$P(\Sigma_\eta|Y,\boldsymbol{\alpha},\Sigma_\eta,\beta) = P(Y,\boldsymbol{\alpha},\Sigma_\eta,\beta|\Sigma_\eta)P(\Sigma_\eta)$$

$$\propto |\Sigma_\eta|^{-(N(J-1)+\nu_\eta+p+1)/2}\exp(-tr(\Lambda_\eta + \sum_{i=1}^{N}\sum_{j=2}^{J}(\boldsymbol{\alpha}_{ij} - \boldsymbol{\alpha}_{i(j-1)})^2)\Sigma_\eta^{-1})$$

$$\Sigma_\eta|\ldots \sim \text{Inv-Wishart}_{\nu_\eta+N(J-1)}(\Lambda_\eta + \sum_{i=1}^{N}\sum_{j=2}^{J}(\boldsymbol{\alpha}_{ij} - \boldsymbol{\alpha}_{i(j-1)})^2)$$

β Standardization

$$
\begin{aligned}
\boldsymbol{y}_{ij} &= \boldsymbol{\alpha}_{ij}^{(m)} + \mathbb{X}_{ij}\boldsymbol{\beta}^{(m)} + \boldsymbol{\varepsilon}_{ij} \\
\boldsymbol{y}_{ij} - \boldsymbol{\alpha}_{ij}^{(m)} &= \mathbb{X}_{ij}\boldsymbol{\beta}^{(m)} + \boldsymbol{\varepsilon}_{ij} \\
(\boldsymbol{y}_{ij} - \boldsymbol{\alpha}_{ij}^{(m)})(\Sigma_{\varepsilon}^{(m)})^{-1/2} &= \mathbb{X}_{ij}\boldsymbol{\beta}^{(m)}(\Sigma_{\varepsilon}^{(m)})^{-1/2} + \boldsymbol{\varepsilon}_{ij}(\Sigma_{\varepsilon}^{(m)})^{-1/2} \\
\boldsymbol{y}_{ij}^{(m)*} &= \mathbb{X}_{ij}\boldsymbol{\beta}^{(m)}(\Sigma_{\varepsilon}^{(m)})^{-1/2} + \boldsymbol{\varepsilon}_{ij}^{*}
\end{aligned}
$$

Now, $\text{var}(\varepsilon_{ij}^{*}) = I$. Each test is now independent and effect comparison can be made using $\beta^{*(m)} = \boldsymbol{\beta}^{(m)}(\Sigma_{\varepsilon}^{(m)})^{-1/2}$.

Fully Simulated Data Analysis

We simulated 100 subjects with between 2-13 observations from the following model:

$$\begin{bmatrix} y_{ij1} \\ y_{ij2} \\ y_{ij3} \end{bmatrix} = \begin{bmatrix} \alpha_{ij1} \\ \alpha_{ij2} \\ \alpha_{ij3} \end{bmatrix} + \begin{bmatrix} \mathbf{x_{ij}}\boldsymbol{\beta_1} \\ \mathbf{x_{ij}}\boldsymbol{\beta_2} \\ \mathbf{x_{ij}}\boldsymbol{\beta_3} \end{bmatrix} + \begin{bmatrix} \varepsilon_{ij1} \\ \varepsilon_{ij2} \\ \varepsilon_{ij3} \end{bmatrix}, \begin{bmatrix} \varepsilon_{ij1} \\ \varepsilon_{ij2} \\ \varepsilon_{ij3} \end{bmatrix} \sim N(0, \begin{bmatrix} 15 & 2.4 & 1 \\ 2.4 & 15 & 1 \\ 1 & 1 & 10 \end{bmatrix})$$

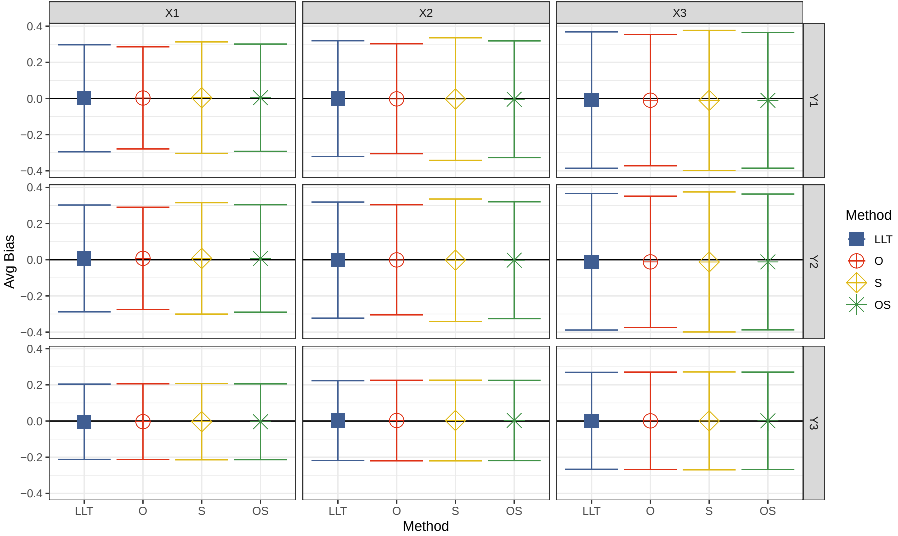
$$\begin{bmatrix} \alpha_{ij1} \\ \alpha_{ij2} \\ \alpha_{ij3} \end{bmatrix} = \begin{bmatrix} \alpha_{i(j-1)1} \\ \alpha_{i(j-1)2} \\ \alpha_{i(j-1)3} \end{bmatrix} + \begin{bmatrix} \eta_{ij1} \\ \eta_{ij2} \\ \eta_{ij3} \end{bmatrix}, \begin{bmatrix} \eta_{ij1} \\ \eta_{ij2} \\ \eta_{ij3} \end{bmatrix} \sim N(0, \delta_{ij} \begin{bmatrix} 5 & 3.7 & 0 \\ 3.7 & 5 & 0 \\ 0 & 0 & 2 \end{bmatrix})$$

Parameter coverage

Table 1: Linear effect coverage percentage.

Test	Variable	Beta	LLT	O	S	OS
Y1	X1	4	96.2%	94.8%	96.8%	96.1%
Y1	X2	2	94.6%	93.4%	95.5%	94.5%
Y1	X3	1	95.2%	94.1%	95.8%	95%
Y2	X1	-3	95%	94.4%	95.8%	95.7%
Y2	X2	0	94.5%	93.1%	95.9%	95.2%
Y2	X3	1	94.9%	94%	96.8%	96.4%
Y3	X1	0	95.7%	96.2%	96%	95.7%
Y3	X2	0	95%	94.7%	94.9%	95.2%
Y3	X3	0	94.5%	94.8%	94.9%	94.9%

Bias and Variability



Variance Coverage

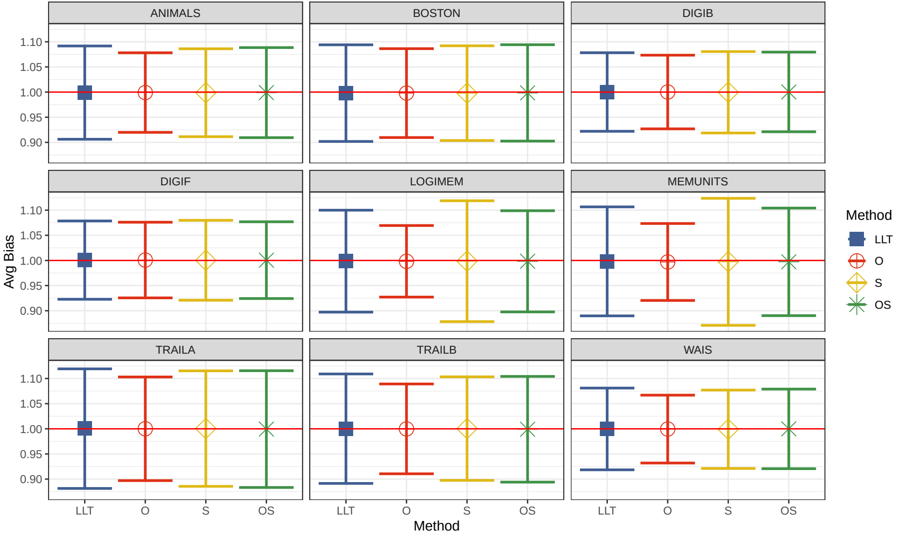
Parameter	True value	LLT	O	S	OS
Observation Error					
1,1	15	14.964 (94.4%)	15.752 (88.2%)	14.273 (86.7%)	15.075 (94.3%)
1,2	2.4	-	4.976 (2.4%)	-	2.14 (91.7%)
2,2	15	14.948 (94.1%)	15.733 (87.9%)	14.248 (85.1%)	15.066 (93.8%)
1,3	1	-	0.997 (94.6%)	-	0.979 (94.2%)
2,3	1	-	1.008 (94.8%)	-	0.997 (92.6%)
3,3	10	10.021 (93.7%)	10.021 (93.5%)	9.988 (93.7%)	10.056 (93.1%)
State Process					
1,1	5	4.835 (94%)	4.168 (78%)	5.776 (82.8%)	4.941 (94.2%)
1,2	3.714	-	-	5.007 (44.7%)	3.891 (93.7%)
2,2	5	4.849 (92.4%)	4.183 (78.9%)	5.8 (82%)	4.952 (92.7%)
1,3	0	-	-	0.47 (71.9%)	-0.009 (93.8%)
2,3	0	-	-	0.47 (73.8%)	-0.015 (93.9%)
3,3	2	1.894 (92.6%)	1.943 (93.8%)	1.977 (93%)	1.939 (93%)

Real Data Simulation

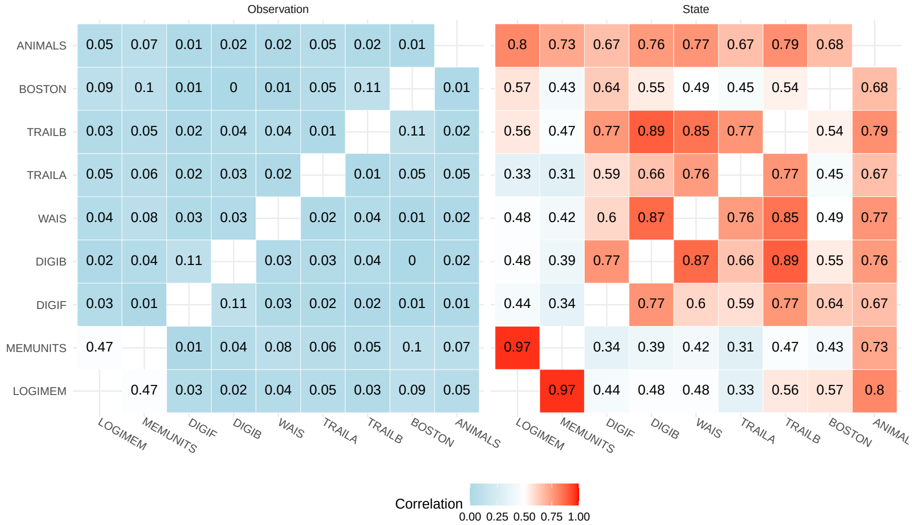
Simulated Parameter Coverage

	LLT	O	S	OS
ANIMALS	94.1%	90.9%	93.9%	94.9%
BOSTON	93.5%	91.8%	92.1%	92.9%
DIGIB	94.1%	92.9%	93.9%	94.1%
DIGIF	94.3%	93.8%	95.5%	94.2%
LOGIMEM	96.1%	86.2%	97.1%	97%
MEMUNITS	96.5%	87.7%	97.9%	96.6%
TRAILA	93.9%	89.8%	93.8%	92.9%
TRAILB	95.2%	88.6%	93.2%	94.4%
WAIS	94.9%	90.3%	94.2%	95.1%

Simulated Parameter Bias and Variability

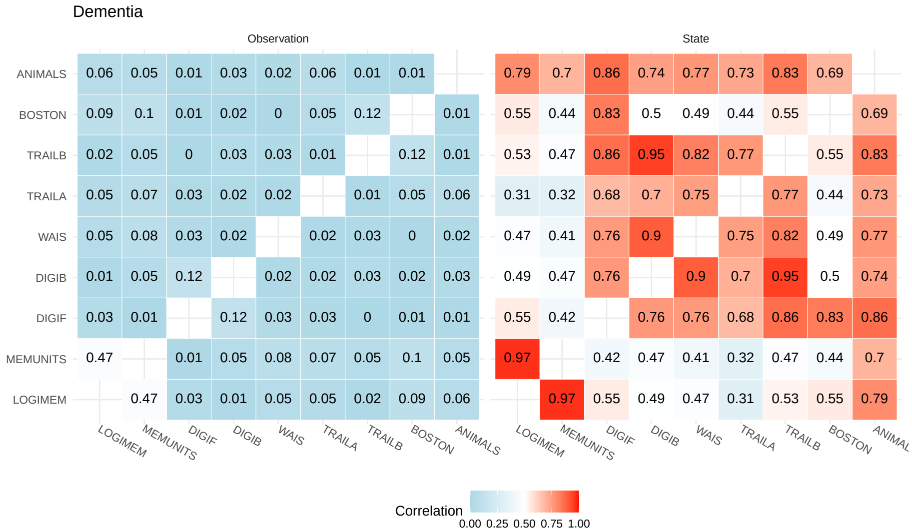


Covariance Matrices

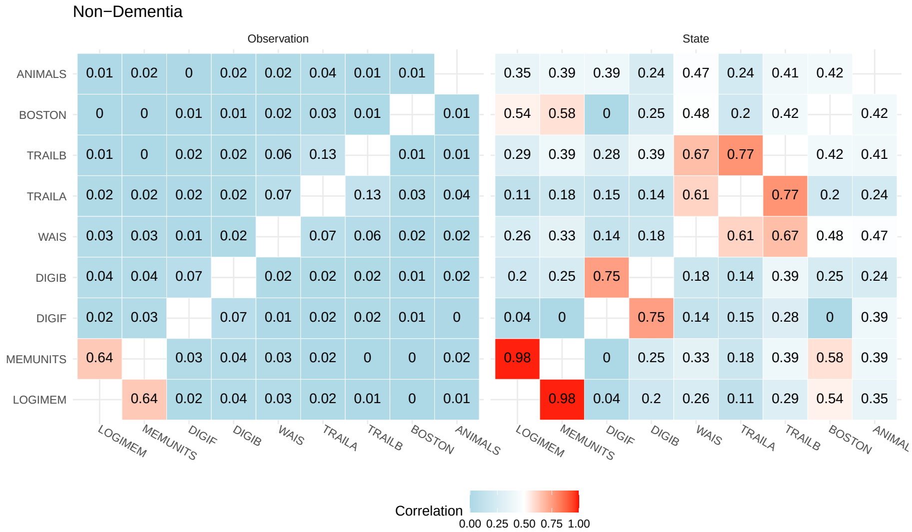


Data Analysis

Dementia Transitioners



Non-Dementia Transitioners



Section 3

Project 3 Overview

Project 3 Aims

- We wish to better measure the underlying relatedness between cognition tests and how they load to certain domains.
- We propose the use of an LLT factor analysis.

Project 3

$$
\begin{bmatrix} y_{ij1} \\ y_{ij2} \\ \vdots \\ y_{ijK} \end{bmatrix} = G \begin{bmatrix} \alpha_{ij1} \\ \alpha_{ij2} \\ \vdots \\ \alpha_{ijq} \end{bmatrix} + \begin{bmatrix} \mathbf{x_{ij}} \beta^{\mathbf{(1)}} \\ \mathbf{x_{ij}} \beta^{\mathbf{(2)}} \\ \vdots \\ \mathbf{x_{ij}} \beta^{\mathbf{(K)}} \end{bmatrix} + \begin{bmatrix} \varepsilon_{ij1} \\ \varepsilon_{ij2} \\ \vdots \\ \varepsilon_{ijK} \end{bmatrix}, \begin{bmatrix} \varepsilon_{ij1} \\ \varepsilon_{ij2} \\ \vdots \\ \varepsilon_{ijK} \end{bmatrix} \sim N(0, \sigma_\varepsilon^2 I)
$$

$$
\begin{bmatrix} \alpha_{ij1} \\ \alpha_{ij2} \\ \vdots \\ \alpha_{ijq} \end{bmatrix} = \begin{bmatrix} \alpha_{i(j-1)1} \\ \alpha_{i(j-1)2} \\ \vdots \\ \alpha_{i(j-1)q} \end{bmatrix} + \begin{bmatrix} \eta_{ij1} \\ \eta_{ij2} \\ \vdots \\ \eta_{ijq} \end{bmatrix}, \begin{bmatrix} \eta_{ij1} \\ \eta_{ij2} \\ \vdots \\ \eta_{ijq} \end{bmatrix} \sim N(0, \delta_{ij} \sigma_\eta^2 I)
$$

Where $G \in R^{K \times q}$ is a factor loading matrix. We will again model this in the Bayesian context and give each entry of G a normal prior ($G_{s,t} \sim N(0, \sigma_G^2)$).

Project 3

$$\begin{bmatrix} y_{ij1} \\ y_{ij2} \\ \vdots \\ y_{ijK} \end{bmatrix} = G \begin{bmatrix} \alpha_{ij1} \\ \alpha_{ij2} \\ \vdots \\ \alpha_{ijK} \end{bmatrix} + \begin{bmatrix} \boldsymbol{x_{ij}} \beta^{(1)} \\ \boldsymbol{x_{ij}} \beta^{(2)} \\ \vdots \\ \boldsymbol{x_{ij}} \beta^{(K)} \end{bmatrix} + \begin{bmatrix} \varepsilon_{ij1} \\ \varepsilon_{ij2} \\ \vdots \\ \varepsilon_{ijK} \end{bmatrix}, \begin{bmatrix} \varepsilon_{ij1} \\ \varepsilon_{ij2} \\ \vdots \\ \varepsilon_{ijK} \end{bmatrix} \sim N(0, \sigma_\varepsilon^2 I)$$

$$\begin{bmatrix} \alpha_{ij1} \\ \alpha_{ij2} \\ \vdots \\ \alpha_{ijK} \end{bmatrix} = \begin{bmatrix} \alpha_{i(j-1)1} \\ \alpha_{i(j-1)2} \\ \vdots \\ \alpha_{i(j-1)K} \end{bmatrix} + \begin{bmatrix} \eta_{ij1} \\ \eta_{ij2} \\ \vdots \\ \eta_{ijK} \end{bmatrix}, \begin{bmatrix} \eta_{ij1} \\ \eta_{ij2} \\ \vdots \\ \eta_{ijK} \end{bmatrix} \sim \tau N(0, \delta_{ij}\sigma_\eta^2 I) + (1-\tau) N(0, w)$$

\end{equation*}

Where G is now a $K \times K$ matrix and each chain of α has a spike and slab distribution. The vector τ is of length K and controls which chains of α will be set to 0. Using this spike and slab we are able to see how the underlying cognition levels naturally cluster together.

Timeline

- Project 1 is finished.
- Start coding project 3 immediately while finishing a write up for project 2 by the end of March.
- Finish project 3 in June/July.
- Defend September.

Thank you!

- Recommendations?
- Questions?