Multivariate Skew-normal Linear Mixed Models for Multi-outcome Longitudinal Data

Dr. Tsung I. Lin

(Joint work with Dr. Wan-Lun Wang)

Institute of Statistics

National Chung Hsing University

August 25, 2014

Outline

1. Introduction

Linear mixed models (LMM; Laird and Ware, 1982)

$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{Z}\boldsymbol{b} + \boldsymbol{\varepsilon}$$

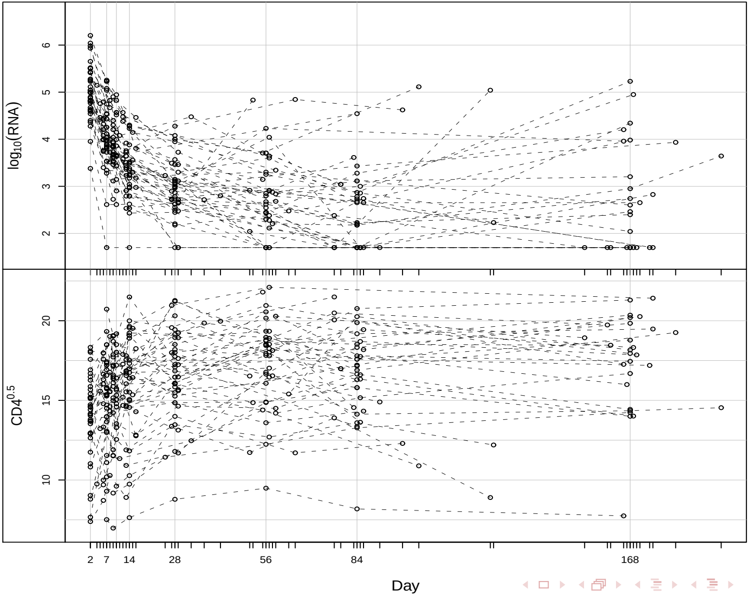
$\boldsymbol{b} \sim$ **Normal** $\quad \boldsymbol{\varepsilon} \sim$ **Normal**

- Shah et al. (1997) proposed the multivariate linear mixed model (MLMM) for multi-outcome longitudinal data.
- Lin and Lee (2008) proposed the skew-normal linear mixed model (SNLMM) to handle asymmetric single-outcome longitudinal data.
- Goal: Extend the SNLMM to the multivariate skew-normal linear mixed model (MSNLMM).

2. ACTG 315 Clinical Trials (Lederman et al. 1998)

- 53 HIV-1 infected patients were collected from 3 clinical centers.
  - 5 weeks prior to study, patients discontinued any antiretroviral therapy and then were treated with a combination of potent antiviral drugs (ritonavir, 3TC, and AZT).
  - monitored at days 0, 2, 7, 10 and weeks 2, 4, 8, 12, 24, and 48.
- Response variables:
  - virologic marker (plasma HIV-1 RNA copies)
  - immunologic marker (CD4+ T cell counts)
- We concentrate on 48 patients with the two markers completely observed for the first 24 weeks (168 days) of treatment.

$\log_{10}$RNA and $CD4^{0.5}$ for 48 HIV-1 Infected Patients



- Diagonal entries
  - Between-responses correlations at exactly monitored days.
- Upper-triangular entries
  - Between-occasion correlations for $\log_{10}$ RNA copies.
- Lower-triangular entries
  - Between-occasion correlations for CD4$^{0.5}$ cell counts.

		$\log_{10}$ RNA							
		day 2	day 7	day 10	week 2	week 4	week 8	week 12	week 24
CD4$^{0.5}$	day 2	-0.3231	0.8375	0.7631	0.6850	0.5770	0.4728	0.3136	0.4197
	day 7	0.7405	-0.1542	0.8785	0.8167	0.6520	0.4285	0.3039	0.4333
	day 10	0.6942	0.8447	-0.1195	0.8381	0.6860	0.5398	0.2502	0.4436
	day 14	0.5739	0.7335	0.7786	-0.0494	0.7662	0.3912	0.3810	0.3233
	day 28	0.5216	0.7356	0.7433	0.7951	0.0592	0.5970	0.5708	0.3903
	day 56	0.6869	0.6689	0.7663	0.7301	0.7967	0.0768	0.6733	0.3015
	day 84	0.4265	0.5625	0.5926	0.6858	0.6828	0.7273	-0.4579	0.4180
	day 168	0.5952	0.6908	0.7243	0.7540	0.8120	0.8093	0.8565	0.0318

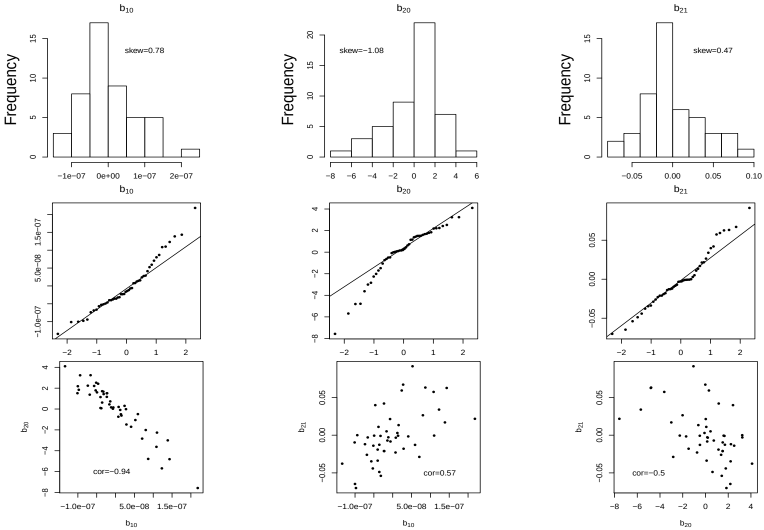
Preliminary Analysis for ACTG 315 data

- Let $\boldsymbol{y}_{i1} = \log_{10} \text{RNA}_i$, $\boldsymbol{y}_{i2} = \text{CD4}_i^{0.5}$, $i = 1, \ldots, 48$.
- Fit a bivariate LMM (Shah et al., 1997) to the data:

$$\begin{bmatrix} \boldsymbol{y}_{i1} \\ \boldsymbol{y}_{i2} \end{bmatrix} = \boldsymbol{I}_2 \otimes [\mathbf{1}_i : \boldsymbol{t}_i : \boldsymbol{t}_i^2 : \text{rna}_i \mathbf{1}_i]\boldsymbol{\beta} + \begin{bmatrix} \mathbf{1}_i & \mathbf{0}_i & \mathbf{0}_i \\ \mathbf{0}_i & \mathbf{1}_i & \boldsymbol{t}_i \end{bmatrix} \begin{bmatrix} b_{i10} \\ b_{i20} \\ b_{i21} \end{bmatrix} + \begin{bmatrix} \boldsymbol{e}_{i1} \\ \boldsymbol{e}_{i2} \end{bmatrix}$$

  - $\boldsymbol{I}_d$ is an identity matrix of order d;
  - $\mathbf{0}_i$ and $\mathbf{1}_i$ are $s_i \times 1$ vectors with each entry being 0 and 1, respectively;
  - $\boldsymbol{t}_i = (t_{i1}, \ldots, t_{is_i})$ with $t_{ik} = \text{day}_{ik}/7$ for $k = 1, \ldots, s_i$;
  - rna_i is the baseline $\log_{10}$RNA (time independent covariate);
  - b_{i10}: random intercepts for $\log_{10}$RNA;
  - (b_{i20}) random intercepts and (b_{i21}) random slopes for $\text{CD4}^{0.5}$.

Empirical Bayes estimates of random effects obtained by fitting bivariate LMM



Multivariate Skew-Normal Distribution (Azzalini and Dalla Valle 1996)

- $\boldsymbol{X} \sim \mathcal{SN}_d(\boldsymbol{\mu}, \boldsymbol{\Omega}, \boldsymbol{\lambda})$: the MSN distribution with location vector $\boldsymbol{\mu} \in \mathbb{R}^d$, scale covariance matrix $\boldsymbol{\Omega}$ and skewness vector $\boldsymbol{\lambda} \in \mathbb{R}^d$.
- Probability density function:

$$f(\boldsymbol{x}; \boldsymbol{\mu}, \boldsymbol{\Omega}, \boldsymbol{\lambda}) = 2\phi_d(\boldsymbol{x}|\boldsymbol{\mu}, \boldsymbol{\Omega})\Phi(\boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{\Omega}^{-1/2}(\boldsymbol{x} - \boldsymbol{\mu}))$$

- Stochastic representation:

$$\boldsymbol{x} = \boldsymbol{\mu} + \boldsymbol{\Omega}^{1/2}\boldsymbol{\delta}\gamma + \boldsymbol{\Omega}^{1/2}(\boldsymbol{I}_d - \boldsymbol{\delta}\boldsymbol{\delta}^{\mathrm{T}})^{1/2}\boldsymbol{z}, \quad \gamma \perp \boldsymbol{z},$$

$\gamma \sim \mathcal{TN}(0, 1; (0, \infty))$, $\boldsymbol{z} \sim \mathcal{N}_d(\mathbf{0}, \boldsymbol{I}_d)$, $\boldsymbol{\delta} = \boldsymbol{\lambda}(1 + \boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{\lambda})^{-1/2}$.

- Hierarchical formulation:

$$\boldsymbol{x}|\gamma \sim \mathcal{N}_d(\boldsymbol{\mu} + \boldsymbol{\Omega}^{1/2}\boldsymbol{\delta}\gamma, \boldsymbol{\Omega}^{1/2}(\boldsymbol{I}_d - \boldsymbol{\delta}\boldsymbol{\delta}^{\mathrm{T}})\boldsymbol{\Omega}^{1/2}),\ \gamma \sim \mathcal{TN}(0, 1; (0, \infty)).$$

Notation for the MSNLMM

- $i = 1, \ldots, N$, $j = 1, \ldots, r$, $t = 1, \ldots,, s_i$
- $\boldsymbol{Y}_i = [\boldsymbol{y}_{i1} : \cdots : \boldsymbol{y}_{ir}]$: $s_i \times r$ outcome matrix of subject i
  - $\boldsymbol{y}_{ij} = (y_{ij1}, \ldots, y_{ijs_i})^{\mathrm{T}}$: a $s_i \times 1$ response vector corresponding to the jth variable.

$$\boldsymbol{y}_i = \text{vec}(\boldsymbol{Y}_i) : \; n_i \times 1 \text{ vector} \;\; n_i = s_i r.$$

- $\boldsymbol{E}_i = [\boldsymbol{e}_{i1} : \cdots : \boldsymbol{e}_{ir}]$: $s_i \times r$ within-subject errors matrix.

$$\boldsymbol{e}_i = \text{vec}(\boldsymbol{E}_i) : \; n_i \times 1 \text{ vector.}$$

- $\boldsymbol{X}_i = \text{diag}\{\boldsymbol{X}_{i1}, \ldots, \boldsymbol{X}_{ir}\}$: $n_i \times p$ design matrix for fixed effects
- $\boldsymbol{Z}_i = \text{diag}\{\boldsymbol{Z}_{i1}, \ldots, \boldsymbol{Z}_{ir}\}$: $n_i \times q$ design matrix for random effects
  - $p = \sum_{j=1}^r p_j$ and $q = \sum_{j=1}^r q_j$; $p_j = \text{rank}(\boldsymbol{X}_{ij})$ and $q_j = \text{rank}(\boldsymbol{Z}_{ij})$

3. Multivariate Skew-Normal Linear Mixed Model

MSNLMM for subject i

$$\boldsymbol{y}_i = \boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{b}_i + \boldsymbol{e}_i \quad \text{with}$$

$$\begin{bmatrix} \boldsymbol{b}_i \\ \boldsymbol{e}_i \end{bmatrix} \sim \mathcal{SN}_{q+n_i}\left(\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \boldsymbol{D} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{R}_i \end{bmatrix}, \begin{bmatrix} \boldsymbol{\lambda} \\ \mathbf{0} \end{bmatrix} \right) \tag{1}$$

- $\boldsymbol{\beta} = (\boldsymbol{\beta}_1^{\mathrm{T}}, \ldots, \boldsymbol{\beta}_r^{\mathrm{T}})^{\mathrm{T}}$: ($p$-dimensional) regression coefficients
- $\boldsymbol{R}_i = \boldsymbol{\Sigma} \otimes \boldsymbol{C}_i$: $n_i \times n_i$ kronecker product matrix.

Damped Exponential Correlation (DEC)

$$\boldsymbol{C}_i = \boldsymbol{C}_i(\phi, \xi; \boldsymbol{t}_i) = \left[\phi^{|t_{ik} - t_{ik'}|^{\xi}}\right], \quad 0 \le \phi < 1, \quad \xi \ge 0,$$

$\xi = 1$: DEC $\rightarrow$ CAR(1) and $\phi = 0$: DEC $\rightarrow$ UNC

1. Marginal distribution:

$$\boldsymbol{y}_i \sim \mathcal{SN}_{n_i}(\boldsymbol{X}_i\boldsymbol{\beta}, \boldsymbol{\Lambda}_i, \boldsymbol{\lambda}_{\boldsymbol{y}_i}) \tag{2}$$

- $\boldsymbol{\Lambda}_i = \boldsymbol{Z}_i\boldsymbol{D}\boldsymbol{Z}_i^{\mathrm{T}} + \boldsymbol{\Sigma}\otimes\boldsymbol{C}_i$, $\boldsymbol{\lambda}_{\boldsymbol{y}_i} = (1 + \boldsymbol{d}_i^{\mathrm{T}}\boldsymbol{\Psi}_i^{-1}\boldsymbol{d}_i)^{-1/2}\boldsymbol{\Lambda}_i^{1/2}\boldsymbol{\Psi}_i^{-1}\boldsymbol{d}_i$.
- $\boldsymbol{\Psi}_i = \boldsymbol{\Lambda}_i - \boldsymbol{d}_i\boldsymbol{d}_i^{\mathrm{T}}$, $\boldsymbol{d}_i = \boldsymbol{Z}_i\boldsymbol{F}\boldsymbol{\delta}$, $\boldsymbol{\delta} = \boldsymbol{\delta}(\boldsymbol{\lambda}) = \boldsymbol{\lambda}/\sqrt{1+\boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{\lambda}} \in (-1,1)^q$.

2. Two-level hierarchy:

$$\boldsymbol{y}_i|\gamma_i \sim \mathcal{N}_{n_i}(\boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{d}_i\gamma_i, \boldsymbol{\Psi}_i), \quad \gamma_i \sim \mathcal{TN}(0,1,(0,\infty)). \tag{3}$$

3. Three-level hierarchy:

$$\begin{aligned} \boldsymbol{y}_i|\boldsymbol{b}_i,\gamma_i &\sim \mathcal{N}_{n_i}\big(\boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{b}_i, \boldsymbol{\Sigma}\otimes\boldsymbol{C}_i\big), \\ \boldsymbol{b}_i|\gamma_i &\sim \mathcal{N}_q(\boldsymbol{\alpha}\gamma_i, \boldsymbol{W}), \quad \gamma_i \sim \mathcal{TN}(0,1,(0,\infty)). \end{aligned} \tag{4}$$

- $\boldsymbol{\alpha} = \boldsymbol{F}\boldsymbol{\delta}$ and $\boldsymbol{W} = \boldsymbol{D} - \boldsymbol{\alpha}\boldsymbol{\alpha}^{\mathrm{T}}$

1. Marginal distribution:

$$\boldsymbol{y}_i \sim \mathcal{SN}_{n_i}(\boldsymbol{X}_i\boldsymbol{\beta}, \boldsymbol{\Lambda}_i, \boldsymbol{\lambda}_{\boldsymbol{y}_i}) \tag{2}$$

- $\boldsymbol{\Lambda}_i = \boldsymbol{Z}_i\boldsymbol{D}\boldsymbol{Z}_i^{\mathrm{T}} + \boldsymbol{\Sigma} \otimes \boldsymbol{C}_i$, $\boldsymbol{\lambda}_{\boldsymbol{y}_i} = (1 + \boldsymbol{d}_i^{\mathrm{T}}\boldsymbol{\Psi}_i^{-1}\boldsymbol{d}_i)^{-1/2}\boldsymbol{\Lambda}_i^{1/2}\boldsymbol{\Psi}_i^{-1}\boldsymbol{d}_i$.
- $\boldsymbol{\Psi}_i = \boldsymbol{\Lambda}_i - \boldsymbol{d}_i\boldsymbol{d}_i^{\mathrm{T}}$, $\boldsymbol{d}_i = \boldsymbol{Z}_i\boldsymbol{F}\boldsymbol{\delta}$, $\boldsymbol{\delta} = \boldsymbol{\delta}(\boldsymbol{\lambda}) = \boldsymbol{\lambda}/\sqrt{1 + \boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{\lambda}} \in (-1, 1)^q$.

2. Two-level hierarchy:

$$\boldsymbol{y}_i|\gamma_i \sim \mathcal{N}_{n_i}(\boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{d}_i\gamma_i, \boldsymbol{\Psi}_i), \quad \gamma_i \sim \mathcal{TN}(0, 1, (0, \infty)). \tag{3}$$

3. Three-level hierarchy:

$$\begin{aligned} \boldsymbol{y}_i|\boldsymbol{b}_i, \gamma_i &\sim \mathcal{N}_{n_i}\big(\boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{b}_i, \boldsymbol{\Sigma} \otimes \boldsymbol{C}_i\big), \\ \boldsymbol{b}_i|\gamma_i &\sim \mathcal{N}_q(\boldsymbol{\alpha}\gamma_i, \boldsymbol{W}), \quad \gamma_i \sim \mathcal{TN}(0, 1, (0, \infty)). \end{aligned} \tag{4}$$

- $\boldsymbol{\alpha} = \boldsymbol{F}\boldsymbol{\delta}$ and $\boldsymbol{W} = \boldsymbol{D} - \boldsymbol{\alpha}\boldsymbol{\alpha}^{\mathrm{T}}$

1. Marginal distribution:

$$\boldsymbol{y}_i \sim \mathcal{SN}_{n_i}(\boldsymbol{X}_i\boldsymbol{\beta}, \boldsymbol{\Lambda}_i, \boldsymbol{\lambda}_{\boldsymbol{y}_i}) \tag{2}$$

- $\boldsymbol{\Lambda}_i = \boldsymbol{Z}_i\boldsymbol{D}\boldsymbol{Z}_i^{\mathrm{T}} + \boldsymbol{\Sigma} \otimes \boldsymbol{C}_i$, $\boldsymbol{\lambda}_{\boldsymbol{y}_i} = (1 + \boldsymbol{d}_i^{\mathrm{T}}\boldsymbol{\Psi}_i^{-1}\boldsymbol{d}_i)^{-1/2}\boldsymbol{\Lambda}_i^{1/2}\boldsymbol{\Psi}_i^{-1}\boldsymbol{d}_i$.
- $\boldsymbol{\Psi}_i = \boldsymbol{\Lambda}_i - \boldsymbol{d}_i\boldsymbol{d}_i^{\mathrm{T}}$, $\boldsymbol{d}_i = \boldsymbol{Z}_i\boldsymbol{F}\boldsymbol{\delta}$, $\boldsymbol{\delta} = \boldsymbol{\delta}(\boldsymbol{\lambda}) = \boldsymbol{\lambda}/\sqrt{1 + \boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{\lambda}} \in (-1, 1)^q$.

2. Two-level hierarchy:

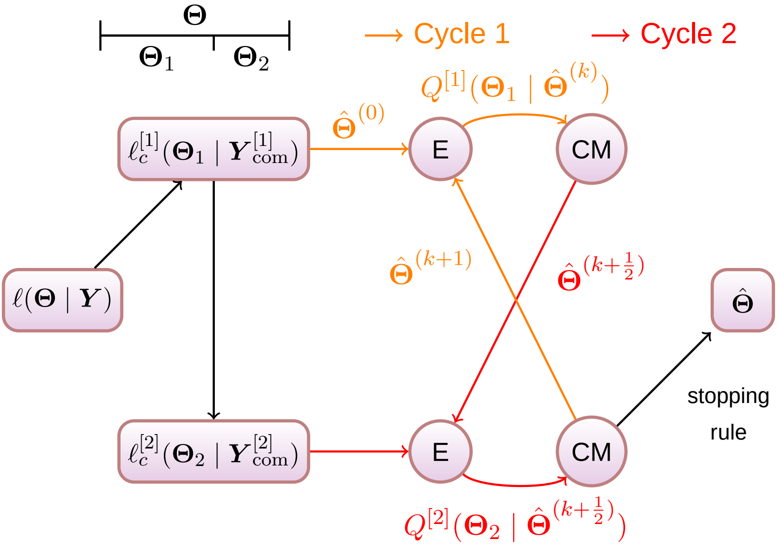
$$\boldsymbol{y}_i|\gamma_i \sim \mathcal{N}_{n_i}(\boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{d}_i\gamma_i, \boldsymbol{\Psi}_i), \quad \gamma_i \sim \mathcal{TN}(0, 1, (0, \infty)). \tag{3}$$

3. Three-level hierarchy:

$$\begin{aligned} \boldsymbol{y}_i|\boldsymbol{b}_i, \gamma_i &\sim \mathcal{N}_{n_i}\big(\boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{b}_i, \boldsymbol{\Sigma} \otimes \boldsymbol{C}_i\big), \\ \boldsymbol{b}_i|\gamma_i &\sim \mathcal{N}_q(\boldsymbol{\alpha}\gamma_i, \boldsymbol{W}), \quad \gamma_i \sim \mathcal{TN}(0, 1, (0, \infty)). \end{aligned} \tag{4}$$

- $\boldsymbol{\alpha} = \boldsymbol{F}\boldsymbol{\delta}$ and $\boldsymbol{W} = \boldsymbol{D} - \boldsymbol{\alpha}\boldsymbol{\alpha}^{\mathrm{T}}$.

Alternating Expectation Conditional Maximization



4.1 The AECM algorithm for MSNLMM

- Partition parameters $\boldsymbol{\theta}$ as $\boldsymbol{\theta}_1 = (\boldsymbol{\beta}, \boldsymbol{\alpha})$, $\boldsymbol{\theta}_2 = (\boldsymbol{F}, \boldsymbol{\Sigma}, \boldsymbol{\delta})$, $\boldsymbol{\theta}_3 = (\phi, \xi)$
- Initial values: $\hat{\boldsymbol{\theta}}^{(0)} = (\hat{\boldsymbol{\beta}}^{(0)}, \hat{\boldsymbol{F}}^{(0)}, \hat{\boldsymbol{\Sigma}}^{(0)}, \hat{\boldsymbol{\delta}}^{(0)}, \hat{\phi}^{(0)}, \hat{\xi}^{(0)})$

Cycle I

The complete data is $\boldsymbol{Y}_{\text{arg}}^{[1]} = (\boldsymbol{y}, \gamma)$, and

$$\ell_c^{[1]}(\boldsymbol{\theta}|\boldsymbol{Y}_{\text{arg}}^{[1]}) = -\frac{1}{2}\sum_{i=1}^{N}\Big\{\log|\boldsymbol{\Psi}_i| + \text{tr}\big(\boldsymbol{\Psi}_i^{-1}\boldsymbol{\Omega}_{1i}\big) + \gamma_i^2\Big\},$$

where $\boldsymbol{\Omega}_{1i} = (\boldsymbol{y}_i - \boldsymbol{X}_i\boldsymbol{\beta} - \boldsymbol{Z}_i\boldsymbol{\alpha}\gamma_i)(\boldsymbol{y}_i - \boldsymbol{X}_i\boldsymbol{\beta} - \boldsymbol{Z}_i\boldsymbol{\alpha}\gamma_i)^{\text{T}}$.

E-step: Evaluate $Q^{[1]}(\boldsymbol{\theta}|\hat{\boldsymbol{\theta}}^{(k)}) = E[\ell_c^{[1]}(\boldsymbol{\theta}|\boldsymbol{Y}_{\text{arg}}^{[1]})|\boldsymbol{y}, \hat{\boldsymbol{\theta}}^{(k)}]$.

CM-step: Update $\hat{\boldsymbol{\theta}}_1^{(k)} = (\hat{\boldsymbol{\beta}}^{(k)}, \hat{\boldsymbol{\alpha}}^{(k)})$ by maximizing $Q^{[1]}(\boldsymbol{\theta}|\hat{\boldsymbol{\theta}}^{(k)})$ over $\boldsymbol{\beta}$ and $\boldsymbol{\alpha}$

Cycle II

The complete data is $\boldsymbol{Y}_{\text{aug}}^{[2]} = (\boldsymbol{Y}_{\text{aug}}^{[1]}, \boldsymbol{b})$, and

$$\ell_c^{[2]}(\boldsymbol{\theta}|\boldsymbol{Y}_{\text{aug}}^{[2]}) = -\frac{1}{2}\sum_{i=1}^{N}\Big\{\log|\boldsymbol{\Sigma}\otimes\boldsymbol{C}_i| + \log|\boldsymbol{W}| + \text{tr}\big((\boldsymbol{\Sigma}\otimes\boldsymbol{C}_i)^{-1}\boldsymbol{\Omega}_{2i}\big) + \text{tr}\big(\boldsymbol{W}^{-1}\boldsymbol{\Omega}_{3i}\big) + \gamma_i^2\Big\},$$

where $\boldsymbol{\Omega}_{2i} = \boldsymbol{e}_i\boldsymbol{e}_i^{\text{T}}$, $\boldsymbol{\Omega}_{3i} = (\boldsymbol{b}_i - \boldsymbol{\alpha}\gamma_i)(\boldsymbol{b}_i - \boldsymbol{\alpha}\gamma_i)^{\text{T}}$, $\boldsymbol{e}_i = \boldsymbol{y}_i - \boldsymbol{X}_i\boldsymbol{\beta} - \boldsymbol{Z}_i\boldsymbol{b}_i$.

E-step: Evaluate $Q^{[2]}(\boldsymbol{\theta}|\hat{\boldsymbol{\theta}}^{(k)}) = E[\ell_c^{[2]}(\boldsymbol{\theta}|\boldsymbol{Y}_{\text{arg}}^{[2]})|\boldsymbol{y}, \hat{\boldsymbol{\theta}}^{(k+1/3)}]$.

CM-step: Update $\hat{\boldsymbol{W}}^{(k)}$ and $\hat{\boldsymbol{\Sigma}}^{(k)}$ by maximizing $Q^{[2]}(\boldsymbol{\theta}|\hat{\boldsymbol{\theta}}^{(k+1/3)})$ over $\boldsymbol{W}$ and σ_{jl}, and then update $\hat{\boldsymbol{D}}^{(k)}$ and $\hat{\boldsymbol{\delta}}^{(k)}$.

CML-step: Calculate $\hat{\boldsymbol{\theta}}_3^{(k+1)} = (\hat{\phi}^{(k+1)}, \hat{\xi}^{(k+1)})$ by maximizing the constrained log-likelihood function evaluated at $\boldsymbol{\theta}_1 = \hat{\boldsymbol{\theta}}_1^{(k+1)}$ and $\boldsymbol{\theta}_2 = \hat{\boldsymbol{\theta}}_2^{(k+1)}$.

Closed-form expressions

$$
\begin{bmatrix} \hat{\boldsymbol{\beta}}^{(k+1)} \\ \hat{\boldsymbol{\alpha}}^{(k+1)} \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^N \boldsymbol{X}_i^{\mathrm{T}} \hat{\boldsymbol{\Psi}}_i^{(k)^{-1}} \boldsymbol{X}_i & \sum_{i=1}^N \hat{\gamma}_i^{(k)} \boldsymbol{X}_i^{\mathrm{T}} \hat{\boldsymbol{\Psi}}_i^{(k)^{-1}} \boldsymbol{Z}_i \\ \sum_{i=1}^N \hat{\gamma}_i^{(k)} \boldsymbol{Z}_i^{\mathrm{T}} \hat{\boldsymbol{\Psi}}_i^{(k)^{-1}} \boldsymbol{X}_i & \sum_{i=1}^N \hat{\gamma}_i^{2(k)} \boldsymbol{Z}_i^{\mathrm{T}} \hat{\boldsymbol{\Psi}}_i^{(k)^{-1}} \boldsymbol{Z}_i \end{bmatrix}^{-1} \times \begin{bmatrix} \sum_{i=1}^N \boldsymbol{X}_i^{\mathrm{T}} \hat{\boldsymbol{\Psi}}_i^{(k)^{-1}} \boldsymbol{y}_i \\ \sum_{i=1}^N \hat{\gamma}_i^{(k)} \boldsymbol{Z}_i^{\mathrm{T}} \hat{\boldsymbol{\Psi}}_i^{(k)^{-1}} \boldsymbol{y}_i \end{bmatrix},
$$

$$
\hat{\boldsymbol{W}}^{(k+1)} = \frac{1}{N} \sum_{i=1}^N \hat{\boldsymbol{\Omega}}_{3i}^{(k+1/3)}(\hat{\boldsymbol{\alpha}}^{(k+1)}),
$$

$$
\hat{\boldsymbol{D}}^{(k+1)} = \hat{\boldsymbol{W}}^{(k+1)} + \hat{\boldsymbol{\alpha}}^{(k+1)} \hat{\boldsymbol{\alpha}}^{(k+1)^{\mathrm{T}}},
$$

$$
\hat{\boldsymbol{\delta}}^{(k+1)} = \hat{\boldsymbol{F}}^{(k+1)^{-1}} \hat{\boldsymbol{\alpha}}^{(k+1)}, \qquad \hat{\boldsymbol{\lambda}}^{(k+1)} = \hat{\boldsymbol{\delta}}^{(k+1)} (1 - \hat{\boldsymbol{\delta}}^{(k+1)^{\mathrm{T}}} \hat{\boldsymbol{\delta}}^{(k+1)})^{-1/2},
$$

$$
\hat{\sigma}_{jl}^{(k+1)} = \begin{cases} \left(\sum_{i=1}^N s_i\right)^{-1} \sum_{i=1}^N \mathrm{tr}\left(\hat{\boldsymbol{C}}_i^{(k)^{-1}} \hat{\Omega}_{ijl}^{(k+1/3)}\right) & \text{for } j = l, \\ \left(2\sum_{i=1}^N s_i\right)^{-1} \sum_{i=1}^N \mathrm{tr}\left(\hat{\boldsymbol{C}}_i^{(k)^{-1}} (\hat{\Omega}_{ijl}^{(k+1/3)} + \hat{\Omega}_{ilj}^{(k+1/3)})\right) & \text{for } j \neq l. \end{cases}
$$

4.2 Random effects estimation

- The posterior density of $\boldsymbol{b}_i$

$$f(\boldsymbol{b}_i|\boldsymbol{y}_i) = \phi_q(\boldsymbol{\mu}_{\boldsymbol{b}_i|\boldsymbol{y}_i}, \boldsymbol{\Sigma}_{\boldsymbol{b}_i|\boldsymbol{y}_i})\frac{\Phi(\boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{D}^{-1/2}\boldsymbol{b}_i)}{\Phi(\eta_i)},$$

$$\begin{aligned}\boldsymbol{\mu}_{\boldsymbol{b}_i|\boldsymbol{y}_i} &= \boldsymbol{D}\boldsymbol{Z}_i^{\mathrm{T}}\boldsymbol{\Lambda}_i^{-1}(\boldsymbol{y}_i - \boldsymbol{X}_i\boldsymbol{\beta}), \quad \boldsymbol{\Sigma}_{\boldsymbol{b}_i|\boldsymbol{y}_i} = \left[\boldsymbol{D}^{-1} + \boldsymbol{Z}_i^{\mathrm{T}}(\boldsymbol{\Sigma}\otimes\boldsymbol{C}_i)^{-1}\boldsymbol{Z}_i\right]^{-1}, \\ \eta_i &= (1 + \boldsymbol{d}_i\boldsymbol{\Psi}_i^{-1}\boldsymbol{d}_i)^{-\frac{1}{2}}\boldsymbol{d}_i^{\mathrm{T}}\boldsymbol{\Psi}_i^{-1}(\boldsymbol{y}_i - \boldsymbol{X}_i\boldsymbol{\beta}).\end{aligned}$$

- The posterior mean

$$\tilde{\boldsymbol{b}}_i(\boldsymbol{\theta}) = E(\boldsymbol{b}_i|\boldsymbol{y}_i) = \boldsymbol{u}_{\boldsymbol{b}_i}(\mu_{\gamma_i} + \kappa_i\sigma_{\gamma_i}) + \boldsymbol{v}_{\boldsymbol{b}_i}, \tag{5}$$

$$\begin{aligned}\boldsymbol{u}_{\boldsymbol{b}_i} &= \left(\boldsymbol{Z}_i^{\mathrm{T}}(\boldsymbol{\Sigma}\otimes\boldsymbol{C}_i)^{-1}\boldsymbol{Z}_i + \boldsymbol{W}^{-1}\right)^{-1}\boldsymbol{W}^{-1}\boldsymbol{\alpha}, \quad \boldsymbol{W} = \boldsymbol{D} - \boldsymbol{\alpha}\boldsymbol{\alpha}^{\mathrm{T}}, \\ \boldsymbol{v}_{\boldsymbol{b}_i} &= \left(\boldsymbol{Z}_i^{\mathrm{T}}(\boldsymbol{\Sigma}\otimes\boldsymbol{C}_i)^{-1}\boldsymbol{Z}_i + \boldsymbol{W}^{-1}\right)^{-1}\boldsymbol{Z}_i^{\mathrm{T}}(\boldsymbol{\Sigma}\otimes\boldsymbol{C}_i)^{-1}(\boldsymbol{y}_i - \boldsymbol{X}_i\boldsymbol{\beta}), \\ \mu_{\gamma_i} &= \sigma_{\gamma_i}\eta_i, \quad \sigma_{\gamma_i}^2 = 1 - \boldsymbol{d}_i^{\mathrm{T}}\boldsymbol{\Lambda}_i^{-1}\boldsymbol{d}_i, \quad \kappa_i = \phi(\eta_i)/\Phi(\eta_i).\end{aligned}$$

- The empirical Bayes estimate $\hat{\boldsymbol{b}}_i = \tilde{\boldsymbol{b}}_i(\hat{\boldsymbol{\theta}})$

5. Application: ACTG 315 data revisited

- Let $y_{i1,k}$ and $y_{i2,k}$ be $\log_{10}$ RNA and CD4$^{0.5}$ responses, respectively, at the kth time point for subject i.
- RI-RI scenario:

$$\begin{aligned} y_{i1,k} &= \beta_{10} + \beta_{11} t_{ik} + \beta_{12} t_{ik}^2 + \beta_{13}\text{rna}_i + b_{i10} + e_{i1,t}, \\ y_{i2,k} &= \beta_{20} + \beta_{21} t_{ik} + \beta_{22} t_{ik}^2 + \beta_{23}\text{rna}_i + b_{i20} + e_{i2,t}, \end{aligned}$$

- RI-RIS scenario:

$$\begin{aligned} y_{i1,k} &= \beta_{10} + \beta_{11} t_{ik} + \beta_{12} t_{ik}^2 + \beta_{13}\text{rna}_i + b_{i10} + e_{i1,t}, \\ y_{i2,k} &= \beta_{20} + \beta_{21} t_{ik} + \beta_{22} t_{ik}^2 + \beta_{23}\text{rna}_i + b_{i20} + b_{i21} t_{i,k} + e_{i2,t}, \end{aligned}$$

- For $\boldsymbol{C}_i$, we adopt the DEC structure and two reduced cases, namely uncorrelated (UNC) and CAR(1).
- For model comparison, the MLMM is also fitted.

Table 1: Summary of model selection criteria

	Model	RI-RI			RI-RIS		
		UNC	CAR1	DEC	UNC	CAR1	DEC
AIC	MLMM	2110.706	2064.147	2054.823	2111.710	2070.165	2060.814
	MSNLMM	2096.949	2050.342	2051.074	2099.910	2058.300	2057.077
BIC	MLMM	2136.903	2092.215	2084.762	2143.520	2103.847	2096.367
	MSNLMM	2126.888	2082.152	2084.756	2137.334	2097.596	2098.244

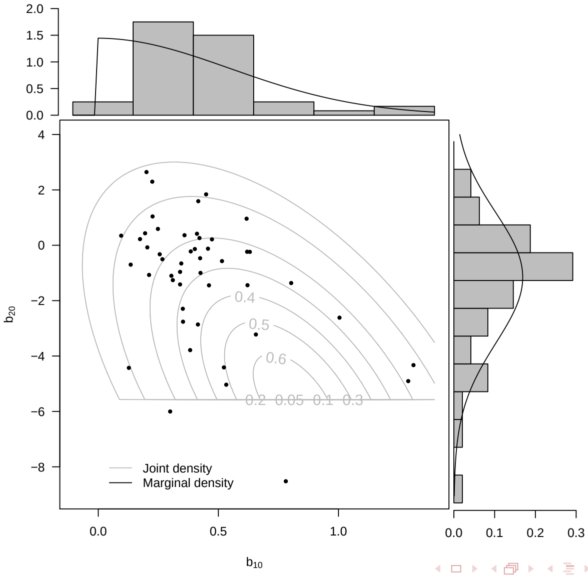
AIC $= 2m - 2\ell_{\max}$ and BIC $= m \log N - 2\ell_{\max}$

Table 2: ML estimates along with the associated standard errors for the MSNLMM with RI-RI and AR1 errors.

	$\hat{\boldsymbol{\beta}}$		$\hat{\boldsymbol{F}}$		$\hat{\boldsymbol{\Sigma}}$
β_{10}	0.7196 (0.990)	f_{11}	0.4839 (0.151)	σ_{11}	0.4457 (0.041)
β_{11}	–0.2971 (0.028)	f_{12}	–0.2659 (0.211)	σ_{12}	–0.3792 (0.078)
β_{12}	0.0098 (0.001)	f_{22}	2.6418 (0.737)	σ_{22}	3.5848 (0.415)
β_{13}	0.6591 (0.188)				$\hat{\rho}_{12} = -0.3$
β_{20}	16.7921 (5.917)				
β_{21}	0.3621 (0.106)				
β_{22}	–0.0108 (0.005)				
β_{23}	–0.1969 (1.072)				

	$\hat{\boldsymbol{C}}_i$		$\hat{\boldsymbol{\delta}}$
ϕ	0.3757 (0.075)	δ_1	0.8764 (0.303)
ξ	1.0000 (——)	δ_2	–0.4817 (0.550)

Scatter plot, contour lines, histograms and density corves for estimated (b_{10}, b_{20})



Summary and Future Research

- A multivariate extension of skew-normal linear mixed model (MSNLMM) is introduced.
- A computationally feasible AECM algorithm is developed for carrying out ML estimation.
- The analysis of the ACTG 315 data illustrates the usefulness of the MSNLMM.
- Future work:
  - Establish missing-data imputation techniques to handle incomplete multiple repeated measures.
  - Develop a joint modelling of the time-to-event data and multivariate longitudinal data.

THANK YOU

FOR

YOUR ATTENTION!