MIXTURE OF MULTIVARIATE t LINEAR MIXED MODELS FOR MULTI-OUTCOME LONGITUDINAL DATA WITH HETEROGENEITY

Wan-Lun Wang

Feng Chia University

Supplementary Material

This document presents some detailed derivations and results for the preliminary analysis of PBCseq data and simulation studies.

S.1 Details for the AECM Algorithm

Define $\mathbf{y} = \{\mathbf{y}_i\}_{i=1}^n$, $\mathbf{u} = \{\mathbf{u}_i\}_{i=1}^n$, $\boldsymbol{\tau} = \{\tau_i\}_{i=1}^n$, and $\mathbf{b} = \{\mathbf{b}_{i1}, \mathbf{b}_{i2}, \ldots, \mathbf{b}_{iG}\}_{i=1}^n$. Let $\phi_d(\cdot|\boldsymbol{\mu}, \boldsymbol{\Omega})$ be the pdf of d-variate normal distribution with mean vector $\boldsymbol{\mu}$ and variance-covariance matrix $\boldsymbol{\Omega}$, and $\mathcal{G}(\cdot|a, b)$ be the pdf of gamma distribution with shape a and inverse scale b. The AECM algorithm for the FM-MtLMM can be summarized as follows.

The 1st cycle

E-step: Under hierarchy (3.4), the complete-data log-likelihood function of $\boldsymbol{\Theta}$ is

$$\ell_c^{[1]}(\boldsymbol{\Theta}|\mathbf{y}, \mathbf{u}) = \sum_{i=1}^n \sum_{g=1}^G u_{ig} \Big\{ \log(w_g) + \log t_{n_i}(\mathbf{y}_i | \mathbf{X}_i \boldsymbol{\beta}_g, \boldsymbol{\Lambda}_{ig}, \nu_g) \Big\}. \tag{S.1}$$

Evaluate the conditional expectation of complete-data log-likelihood function (S.1) given the observed data $\mathbf{y}$ and current values $\hat{\boldsymbol{\Theta}}^{(h)} = (\hat{\boldsymbol{\Theta}}_1^{(h)}, \hat{\boldsymbol{\Theta}}_2^{(h)}, \hat{\boldsymbol{\Theta}}_3^{(h)})$, which gives

$$\begin{aligned} Q^{[1]}(\boldsymbol{\Theta}|\hat{\boldsymbol{\Theta}}^{(h)}) \quad = \quad & \sum_{i=1}^n \sum_{g=1}^G \hat{u}_{ig}^{(h)} \Big\{ \log(w_g) + \log\Gamma(\frac{\nu_g + n_i}{2}) - \log\Gamma(\frac{\nu_g}{2}) \\ & + \frac{1}{2} \Big[\log|\boldsymbol{\Lambda}_{ig}^{-1}| - n_i \log(\nu_g) - (\nu_g + n_i) \log\big(1 + \Delta_{ig}/\nu_g\big) \Big] \Big\}, \end{aligned} \tag{S.2}$$

where $\hat{u}_{ig}^{(h)}$ is calculated as (4.1) evaluated at $\boldsymbol{\Theta} = \hat{\boldsymbol{\Theta}}^{(h)}$.

CM-step: Update $\hat{w}_g^{(h)}$ by maximizing (S.2), yielding $\hat{w}_g^{(h+1)} = \sum_{i=1}^n \hat{u}_{ig}^{(h)}/n$.

CML-step: Update $\hat{\rho}_g^{(h)}$ and $\hat{\nu}_g^{(h)}$ by maximizing the constraint actual log-likelihood functions:

$$\hat{\rho}_g^{(h+1)} = \arg\max_{\rho_g\in(0,1)} \Big\{ \sum_{i=1}^n \log\Big[\hat{w}_g^{(h)} t_{n_i}\big(\mathbf{y}_i|\mathbf{X}_i\hat{\boldsymbol\beta}_g^{(h)}, \mathbf{Z}_i\hat{\mathbf{D}}_g^{(h)}\mathbf{Z}_i^{\mathrm{T}} + \hat{\boldsymbol\Sigma}_g^{(h)}\otimes\mathbf{C}_i(\rho_g), \hat{\nu}_g^{(h)}\big) + \sum_{g'\neq g}^G \hat{w}_{g'}^{(h)} t_{n_i}\big(\mathbf{y}_i|\mathbf{X}_i\hat{\boldsymbol\beta}_{g'}^{(h)}, \mathbf{Z}_i\hat{\mathbf{D}}_{g'}^{(h)}\mathbf{Z}_i^{\mathrm{T}} + \hat{\boldsymbol\Sigma}_{g'}^{(h)}\otimes\mathbf{C}_i(\hat{\rho}_{g'}^{(h)}), \hat{\nu}_{g'}^{(h)}\big)\Big]\Big\},$$

and

$$\hat{\nu}_g^{(h+1)} = \arg\max_{\nu_g\in[2,\infty)} \Big\{\sum_{i=1}^n \log\Big[\hat{w}_g^{(h)} t_{n_i}\big(\mathbf{y}_i|\mathbf{X}_i\hat{\boldsymbol\beta}_g^{(h)}, \hat{\boldsymbol\Lambda}_{ig}^{(h)}, \nu_g\big) + \sum_{g'\neq g}^G \hat{w}_{g'}^{(h)} t_{n_i}\big(\mathbf{y}_i|\mathbf{X}_i\hat{\boldsymbol\beta}_{g'}^{(h)}, \hat{\boldsymbol\Lambda}_{ig'}^{(h)}, \hat{\nu}_{g'}^{(h)}\big)\Big]\Big\},$$

sequentially. In the case of $\nu_1 = \cdots = \nu_G = \nu$, we update the estimate of common DOF $\hat{\nu}^{(h)}$ by

$$\hat{\nu}^{(h+1)} = \arg\max_{\nu\in[2,\infty)} \Big\{\sum_{i=1}^n \log\Big[\sum_{g=1}^G \hat{w}_g^{(h)} t_{n_i}(\mathbf{y}_i|\mathbf{X}_i\hat{\boldsymbol\beta}_g^{(h)}, \hat{\boldsymbol\Lambda}_{ig}^{(h)}, \nu)\Big]\Big\}.$$

The 2nd cycle

E-step: Under hierarchy (3.5), the complete-data log-likelihood function of $\boldsymbol\Theta$ is

$$\ell_c^{[2]}(\boldsymbol\Theta|\mathbf{y},\mathbf{u},\boldsymbol\tau) = \sum_{i=1}^n\sum_{g=1}^G u_{ig}\Big\{\log(w_g) + \log\phi_{n_i}(\mathbf{y}_i|\mathbf{X}_i\boldsymbol\beta_g, \tau_i^{-1}\boldsymbol\Lambda_{ig}) + \log\mathcal{G}(\tau_i|\nu_g/2,\nu_g/2)\Big\}. \tag{S.3}$$

Evaluate the conditional expectation of complete-data log-likelihood function (S.3) given $\mathbf{y}$ and $\hat{\boldsymbol\Theta}^{(h+1/3)} = (\hat{\boldsymbol\Theta}_1^{(h+1)}, \hat{\boldsymbol\Theta}_2^{(h)}, \hat{\boldsymbol\Theta}_3^{(h)})$. The resulting $Q^{[2]}$ function is

$$Q^{[2]}(\boldsymbol\Theta|\hat{\boldsymbol\Theta}^{(h+1/3)}) = \sum_{i=1}^n\sum_{g=1}^G \hat{u}_{ig}^{(h)}\Big\{\log(w_g) - \log\Gamma(\frac{\nu_g}{2}) + \frac{1}{2}\Big[\log|\boldsymbol\Lambda_{ig}^{-1}| - \hat{\tau}_{ig}^{(h)}\big(\Delta_{ig}+\nu_g\big) + \nu_g\Big(\hat{\kappa}_{ig}^{(h)} + \log\big(\frac{\nu_g}{2}\big)\Big)\Big]\Big\}, \tag{S.4}$$

where $\hat{u}_{ig}^{(h)}$ is calculated by (4.1) evaluated at $\boldsymbol\Theta = \hat{\boldsymbol\Theta}^{(h+1/3)}$,

$$\hat{\tau}_{ig}^{(h)} = E[\tau_i|\mathbf{y}_i, u_{ig}=1, \hat{\boldsymbol\Theta}^{(h+1/3)}] = (\hat{\nu}_g^{(h)} + n_i)\Big/(\hat{\nu}_g^{(h)} + \hat{\Delta}_{ig}^{(h)}), \tag{S.5}$$

and

$$\hat{\kappa}_{ig}^{(h)} = E[\log\tau_i|\mathbf{y}_i, u_{ig}=1, \hat{\boldsymbol\Theta}^{(h+1/3)}] = \mathcal{D}G\Big(\frac{\hat{\nu}_g^{(h)} + n_i}{2}\Big) - \log\Big(\frac{\hat{\nu}_g^{(h)} + \hat{\Delta}_{ig}^{(h)}}{2}\Big) \tag{S.6}$$

with $\hat{\Delta}_{ig}^{(h)}$ being Δ_{ig} evaluated at $\mathbf{\Theta} = \hat{\mathbf{\Theta}}^{(h+1/3)}$.

CM-step: Letting the first partial derivative of (S.4) with respect to $\boldsymbol{\beta}_g$ equal zero yields the updated estimate $\hat{\boldsymbol{\beta}}_g^{(h+1)}$, given in (4.2).

The 3rd cycle

E-step: The complete-data log-likelihood function of $\mathbf{\Theta}$ under hierarchy (3.6) is

$$\begin{aligned}\ell_c^{[3]}(\mathbf{\Theta}|\mathbf{y},\mathbf{u},\boldsymbol{\tau},\mathbf{b}) &= \sum_{i=1}^{n}\sum_{g=1}^{G} u_{ig}\Big\{\log(w_g) + \log\phi_{n_i}(\mathbf{y}_i|\mathbf{X}_i\boldsymbol{\beta}_g + \mathbf{Z}_i\mathbf{b}_{ig}, \tau_i^{-1}\mathbf{R}_{ig}) \\ &\quad + \log\phi_q(\mathbf{b}_{ig}|\mathbf{0}, \tau_i^{-1}\mathbf{D}_g) + \log\mathcal{G}(\tau_i|\nu_g/2, \nu_g/2)\Big\}. \qquad \text{(S.7)}\end{aligned}$$

Evaluating the conditional expectation of complete-data log-likelihood function (S.7) given $\mathbf{y}$ and $\hat{\mathbf{\Theta}}^{(h+2/3)} = (\hat{\mathbf{\Theta}}_1^{(h+1)}, \hat{\mathbf{\Theta}}_2^{(h+1)}, \hat{\mathbf{\Theta}}_3^{(h)})$ gives

$$\begin{aligned}Q^{[3]}(\mathbf{\Theta}|\hat{\mathbf{\Theta}}^{(h+2/3)}) &= \sum_{i=1}^{n}\sum_{g=1}^{G} \hat{u}_{ig}^{(h)}\Big\{\log(w_g) - \log\Gamma(\frac{\nu_g}{2}) + \frac{1}{2}\Big[\log|\mathbf{R}_{ig}^{-1}| + \log|\mathbf{D}_g^{-1}| \\ &\quad + \nu_g\big(\log(\frac{\nu_g}{2}) + \hat{\kappa}_{ig}^{(h)} - \hat{\tau}_{ig}^{(h)}\big) - \mathrm{tr}(\mathbf{R}_{ig}^{-1}\widehat{\tau\mathbf{E}}_{ig}^{(h)}(\boldsymbol{\beta}_g)) \\ &\quad - \mathrm{tr}(\mathbf{D}_g^{-1}\widehat{\tau\mathbf{B}}_{ig}^{(h)})\Big]\Big\}, \qquad \text{(S.8)}\end{aligned}$$

where $\hat{u}_{ig}^{(h)}$, $\hat{\tau}_{ig}^{(h)}$, and $\hat{\kappa}_{ig}^{(h)}$ are given in (4.1), (S.5), and (S.6), respectively, evaluated at $\mathbf{\Theta} = \hat{\mathbf{\Theta}}^{(h+2/3)}$;

$$\begin{aligned}\widehat{\tau\mathbf{E}}_{ig}^{(h)}(\boldsymbol{\beta}_g) &= E\big[\tau_i\boldsymbol{\varepsilon}_{ig}\boldsymbol{\varepsilon}_{ig}^{\mathrm{T}}|\mathbf{y}_i, u_{ig}=1, \hat{\mathbf{\Theta}}^{(h+2/3)}\big] \\ &= \hat{\tau}_{ig}^{(h)}\hat{\boldsymbol{\varepsilon}}_{ig}^{(h)}\hat{\boldsymbol{\varepsilon}}_{ig}^{(h)^{\mathrm{T}}} + (\mathbf{I}_{n_i} - \hat{\mathbf{R}}_{ig}^{(h)}\hat{\mathbf{\Lambda}}_{ig}^{(h)^{-1}})\hat{\mathbf{R}}_{ig}^{(h)} \qquad \text{(S.9)}\end{aligned}$$

and

$$\widehat{\tau\mathbf{B}}_{ig}^{(h)} = E\big[\tau_i\mathbf{b}_{ig}\mathbf{b}_{ig}^{\mathrm{T}}|\mathbf{y}_i, u_{ig}=1, \hat{\mathbf{\Theta}}^{(h+2/3)}\big] = \hat{\tau}_{ig}^{(h)}\hat{\mathbf{b}}_{ig}^{(h)}\hat{\mathbf{b}}_{ig}^{(h)^{\mathrm{T}}} + \hat{\mathbf{V}}_{\boldsymbol{b}_{ig}}^{(h)}$$

with

$$\begin{aligned}\hat{\mathbf{b}}_{ig}^{(h)} &= E\big[\mathbf{b}_{ig}|\mathbf{y}_i, u_{ig}=1, \hat{\mathbf{\Theta}}^{(h+2/3)}\big] = \hat{\mathbf{D}}_g^{(h)}\mathbf{Z}_i^{\mathrm{T}}\hat{\mathbf{\Lambda}}_{ig}^{(h)^{-1}}(\mathbf{y}_i - \mathbf{X}_i\hat{\boldsymbol{\beta}}_g^{(h)}), \qquad \text{(S.10)} \\ \hat{\boldsymbol{\varepsilon}}_{ig}^{(h)} &= E\big[\boldsymbol{\varepsilon}_{ig}|\mathbf{y}_i, u_{ig}=1, \hat{\mathbf{\Theta}}^{(h+2/3)}\big] = \mathbf{y}_i - \mathbf{X}_i\boldsymbol{\beta}_g - \mathbf{Z}_i\hat{\mathbf{b}}_{ig}^{(h)},\end{aligned}$$

and $\hat{\mathbf{V}}_{\boldsymbol{b}_{ig}}^{(h)}$ being $\mathbf{V}_{\boldsymbol{b}_{ig}}$ defined in Proposition 1 and evaluated at $\mathbf{\Theta} = \hat{\mathbf{\Theta}}^{(h+2/3)}$.

CM-step: Update $\hat{\mathbf{D}}_g^{(h)}$ and $\hat{\mathbf{\Sigma}}_g^{(h)} = [\hat{\sigma}_{g,ls}^{(h)}]$ by equaling the first partial derivatives of (S.8) with respect to $\mathbf{D}_g$ and each entry of $\mathbf{\Sigma}_g^{-1}$ zero, which gives $\hat{\mathbf{D}}_g^{(h+1)}$ and $\hat{\sigma}_{g,ls}^{(h+1)}$ in (4.3) and (4.4), respectively.

S.2 Individual Score Vector and Hessian Matrix for the g-th Component

The individual score vector $\mathbf{s}^{(i)}_{\boldsymbol{\theta}_g}$ for the g-th component contains the following entries:

$$
\begin{aligned}
s^{(i)}_{w_g} &= 1/w_g, \\
\mathbf{s}^{(i)}_{\boldsymbol{\beta}_g} &= \Big(\frac{\nu_g+n_i}{\nu_g+\Delta_{ig}}\Big)\mathbf{X}_i^{\mathrm{T}}\boldsymbol{\Lambda}_{ig}^{-1}(\mathbf{y}_i-\mathbf{X}_i\boldsymbol{\beta}_g), \\
\big[\mathbf{s}^{(i)}_{\boldsymbol{\alpha}_g}\big]_l &= \frac{1}{2}\Big\{\Big(\frac{\nu_g+n_i}{\nu_g+\Delta_{\boldsymbol{y}_{ig}}}\Big)(\mathbf{y}_i-\mathbf{X}_i\boldsymbol{\beta}_g)^{\mathrm{T}}\boldsymbol{\Lambda}_{ig}^{-1}\dot{\boldsymbol{\Lambda}}_{igl}\boldsymbol{\Lambda}_{ig}^{-1}(\mathbf{y}_i-\mathbf{X}_i\boldsymbol{\beta}_g)-\mathrm{tr}(\boldsymbol{\Lambda}_{ig}^{-1}\dot{\boldsymbol{\Lambda}}_{igl})\Big\}, \\
\mathbf{s}^{(i)}_{\nu_g} &= \frac{1}{2}\Big\{\mathcal{D}_{\mathcal{G}}\Big(\frac{\nu_g+n_i}{2}\Big)-\mathcal{D}_{\mathcal{G}}\Big(\frac{\nu_g}{2}\Big)-\frac{n_i}{\nu_g}-\log\Big(1+\frac{\Delta_{ig}}{\nu_g}\Big)+\frac{(\nu_g+n_i)\Delta_{ig}}{(\nu_g+\Delta_{ig})\nu_g}\Big\},
\end{aligned}
$$

for $l=1,\ldots,d^*$, $d^*=q(q+1)/2+r(r+1)/2+1$, where

$$
\dot{\boldsymbol{\Lambda}}_{ig}=\frac{\partial\boldsymbol{\Lambda}_{ig}}{\partial(\boldsymbol{\alpha}_g)_l}=\begin{cases}\mathbf{Z}_i\frac{\partial\mathbf{D}_g}{\partial(\boldsymbol{\alpha}_g)_l}\mathbf{Z}_i^{\mathrm{T}} & \text{if } (\boldsymbol{\alpha}_g)_l=\mathrm{vech}(\mathbf{D}_g),\\[2ex] \frac{\partial\boldsymbol{\Sigma}_g}{\partial(\boldsymbol{\alpha}_g)_l}\otimes\mathbf{C}_i(\rho_g) & \text{if } (\boldsymbol{\alpha}_g)_l=\mathrm{vech}(\boldsymbol{\Sigma}_g),\\[2ex] \boldsymbol{\Sigma}_g\otimes\frac{\partial\mathbf{C}_g(\rho_g)}{\partial(\boldsymbol{\alpha}_g)_l} & \text{if } (\boldsymbol{\alpha}_g)_l=\rho_g.\end{cases}
$$

Besides, the component submatrix $\mathbf{H}^{(i)}_{\boldsymbol{\theta}_g\boldsymbol{\theta}_g}$, which is the negative of individual Hessian matrix and symmetric, can be written as

$$
\mathbf{H}^{(i)}_{\boldsymbol{\theta}_g\boldsymbol{\theta}_g}=\begin{bmatrix} H^{(i)}_{w_gw_g} & \mathbf{H}^{(i)}_{w_g\boldsymbol{\beta}_g} & \mathbf{H}^{(i)}_{w_g\boldsymbol{\alpha}_g} & \mathbf{H}^{(i)}_{w_g\nu_g}\\ & \mathbf{H}^{(i)}_{\boldsymbol{\beta}_g\boldsymbol{\beta}_g} & \mathbf{H}^{(i)}_{\boldsymbol{\beta}_g\boldsymbol{\alpha}_g} & \mathbf{H}^{(i)}_{\boldsymbol{\beta}_g\nu_g}\\ & & \mathbf{H}^{(i)}_{\boldsymbol{\alpha}_g\boldsymbol{\alpha}_g} & \mathbf{H}^{(i)}_{\boldsymbol{\alpha}_g\nu_g}\\ & & & H^{(i)}_{\nu_g\nu_g}\end{bmatrix},
$$

where $H^{(i)}_{w_gw_g}=1/w_g^2$, $\mathbf{H}^{(i)}_{w_g\boldsymbol{\beta}_g}=\mathbf{0}$, $\mathbf{H}^{(i)}_{w_g\boldsymbol{\alpha}_g}=\mathbf{0}$,

$$
\begin{aligned}
\mathbf{H}^{(i)}_{\boldsymbol{\beta}_g\boldsymbol{\beta}_g} &= \Big(\frac{\nu_g+n_i}{\nu_g+\Delta_{ig}}\Big)\Big[\mathbf{X}_i^{\mathrm{T}}\boldsymbol{\Lambda}_{ig}^{-1}\mathbf{X}_i+\Big(\frac{2}{\nu_g+\Delta_{ig}}\Big)\mathbf{X}_i^{\mathrm{T}}\boldsymbol{\Lambda}_{ig}^{-1}(\mathbf{y}_i-\mathbf{X}_i\boldsymbol{\beta}_g)(\mathbf{y}_i-\mathbf{X}_i\boldsymbol{\beta}_g)^{\mathrm{T}}\\
&\quad\times\boldsymbol{\Lambda}_{ig}^{-1}\mathbf{X}_i\Big],\\
\big[\mathbf{H}^{(i)}_{\boldsymbol{\beta}_g\boldsymbol{\alpha}_g}\big]_l &= \Big(\frac{\nu_g+n_i}{\nu_g+\Delta_{ig}}\Big)\Big[\mathbf{X}_i^{\mathrm{T}}\boldsymbol{\Lambda}_{ig}^{-1}\dot{\boldsymbol{\Lambda}}_{igl}\boldsymbol{\Lambda}_{ig}^{-1}(\mathbf{y}_i-\mathbf{X}_i\boldsymbol{\beta}_g)+(\nu_g+\Delta_{ig})^{-1}\mathbf{X}_i^{\mathrm{T}}\boldsymbol{\Lambda}_{ig}^{-1}\\
&\quad\times(\mathbf{y}_i-\mathbf{X}_i\boldsymbol{\beta}_g)(\mathbf{y}_i-\mathbf{X}_i\boldsymbol{\beta}_g)^{\mathrm{T}}\boldsymbol{\Lambda}_{ig}^{-1}\dot{\boldsymbol{\Lambda}}_{igl}\boldsymbol{\Lambda}_{ig}^{-1}(\mathbf{y}_i-\mathbf{X}_i\boldsymbol{\beta}_g)\Big],\\
\mathbf{H}^{(i)}_{\boldsymbol{\beta}_g\nu_g} &= \frac{(n_i-\Delta_{ig})}{(\nu_g+\Delta_{ig})^2}\mathbf{X}_i^{\mathrm{T}}\boldsymbol{\Lambda}_{ig}^{-1}(\mathbf{y}_i-\mathbf{X}_i\boldsymbol{\beta}_g),
\end{aligned}
$$

$$\begin{aligned}
\big[\mathbf{H}^{(i)}_{\boldsymbol{\alpha}_g\boldsymbol{\alpha}_g}\big]_{lv} &= -\frac{1}{2}\Big\{\mathrm{tr}(\boldsymbol{\Lambda}^{-1}_{ig}\dot{\boldsymbol{\Lambda}}_{igv}\boldsymbol{\Lambda}^{-1}_{ig}\dot{\boldsymbol{\Lambda}}_{igl}) - \mathrm{tr}(\boldsymbol{\Lambda}^{-1}_{ig}\ddot{\boldsymbol{\Lambda}}_{ijlv}) + \frac{(\nu_g+n_i)}{(\nu_g+\Delta_{ig})^2}(\mathbf{y}_i-\mathbf{X}_i\boldsymbol{\beta}_g)^{\mathrm{T}} \\
&\quad \times\boldsymbol{\Lambda}^{-1}_{ig}\dot{\boldsymbol{\Lambda}}_{igv}\boldsymbol{\Lambda}^{-1}_{ig}(\mathbf{y}_i-\mathbf{X}_i\boldsymbol{\beta}_g)(\mathbf{y}_i-\mathbf{X}_i\boldsymbol{\beta}_g)^{\mathrm{T}}\boldsymbol{\Lambda}^{-1}_{ig}\dot{\boldsymbol{\Lambda}}_{igl}\boldsymbol{\Lambda}^{-1}_{ig}(\mathbf{y}_i-\mathbf{X}_i\boldsymbol{\beta}_g) \\
&\quad -\Big(\frac{\nu_g+n_i}{\nu_g+\Delta_{ig}}\Big)(\mathbf{y}_i-\mathbf{X}_i\boldsymbol{\beta}_g)^{\mathrm{T}}\big[2\boldsymbol{\Lambda}^{-1}_{ig}\dot{\boldsymbol{\Lambda}}_{igv}\boldsymbol{\Lambda}^{-1}_{ig}\dot{\boldsymbol{\Lambda}}_{igl}\boldsymbol{\Lambda}^{-1}_{ig} - \boldsymbol{\Lambda}^{-1}_{ig}\ddot{\boldsymbol{\Lambda}}_{iglv}\boldsymbol{\Lambda}^{-1}_{ig}\big] \\
&\quad \times(\mathbf{y}_i-\mathbf{X}_i\boldsymbol{\beta}_g)\Big\}, \\
\big[\mathbf{H}^{(i)}_{\boldsymbol{\alpha}_g\nu_g}\big]_{l} &= -\frac{1}{2}\Big[\frac{1}{(\nu_g+\Delta_{ig})} - \frac{(\nu_g+n_i)}{(\nu_g+\Delta_{ig})^2}\Big](\mathbf{y}_i-\mathbf{X}_i\boldsymbol{\beta}_g)^{\mathrm{T}}\boldsymbol{\Lambda}^{-1}_{ig}\dot{\boldsymbol{\Lambda}}_{igl}\boldsymbol{\Lambda}^{-1}_{ig}(\mathbf{y}_i-\mathbf{X}_i\boldsymbol{\beta}_g),
\end{aligned}$$

and

$$\mathbf{H}^{(i)}_{\nu_g\nu_g} = -\frac{1}{4}\Big\{\mathcal{T}_{\mathcal{G}}\big(\frac{\nu_g+n_i}{2}\big) - \mathcal{T}_{\mathcal{G}}\big(\frac{\nu_g}{2}\big) + \frac{2n_i}{\nu_g^2} + \frac{4\Delta_{ig}}{(\nu_g+\Delta_{ig})\nu_g} - \frac{2(\nu_g+n_i)\Delta_{ig}(2\nu_g+\Delta_{ig})}{(\nu_g+\Delta_{ig})^2\nu_g^2}\Big\},$$

for $l, v = 1, \ldots, d^*$.

S.3 Supplementary Tables for Preliminary Analysis of PBCseq Data and Simulation Studies

Table S1. Observed correlations between two responses at exactly monitored years (diagonal entries) and observed autocorrelations for single response over time (upper-triangular entries for lbili marker; lower-triangular entries for lalbumin marker).

		logarithm of serum bilirubin (lbili)															
		Baseline	Month6	Year1	Year2	Year3	Year4	Year5	Year6	Year7	Year8	Year9	Year10	Year11	Year12	Year13	Year14
	Baseline	**−0.39**	0.89	0.85	0.80	0.72	0.65	0.59	0.52	0.52	0.49	0.62	0.40	0.39	0.45	0.45	0.43
	Month 6	0.49	**−0.39**	0.89	0.81	0.78	0.70	0.64	0.56	0.58	0.54	0.60	0.56	0.60	0.68	0.58	0.38
	Year 1	0.37	0.53	**−0.37**	0.90	0.85	0.80	0.74	0.68	0.67	0.64	0.67	0.51	0.60	0.62	0.62	0.68
	Year 2	0.33	0.49	0.46	**−0.49**	0.93	0.89	0.84	0.78	0.78	0.72	0.70	0.72	0.69	0.71	0.87	0.74
	Year 3	0.42	0.55	0.52	0.65	**−0.46**	0.95	0.92	0.86	0.87	0.84	0.72	0.73	0.68	0.75	0.89	0.80
	Year 4	0.29	0.49	0.49	0.54	0.67	**−0.45**	0.96	0.92	0.91	0.89	0.79	0.80	0.74	0.82	0.89	0.86
logarithm of	Year 5	0.37	0.45	0.52	0.54	0.55	0.71	**−0.40**	0.95	0.93	0.92	0.80	0.84	0.78	0.85	0.94	0.94
serum albumin	Year 6	0.24	0.33	0.43	0.32	0.51	0.59	0.64	**−0.50**	0.95	0.94	0.81	0.84	0.81	0.89	0.94	0.96
(lalbumin)	Year 7	0.08	0.31	0.34	0.38	0.52	0.62	0.49	0.64	**−0.38**	0.96	0.86	0.88	0.86	0.90	0.96	0.98
	Year 8	0.17	0.42	0.44	0.29	0.52	0.51	0.46	0.74	0.72	**−0.43**	0.89	0.94	0.92	0.93	0.95	1.00
	Year 9	0.18	0.52	0.62	0.42	0.60	0.63	0.70	0.75	0.48	0.82	**−0.61**	0.93	0.93	0.94	0.94	0.99
	Year 10	−0.14	0.09	0.36	−0.14	0.17	0.31	0.40	0.53	0.58	0.59	0.73	**−0.47**	0.97	0.99	0.99	1.00
	Year 11	−0.23	−0.14	0.47	0.07	−0.29	0.55	0.34	0.22	0.42	0.30	0.34	0.26	**−0.65**	0.99	0.95	1.00
	Year 12	0.11	−0.04	0.22	−0.07	0.27	0.46	0.51	0.13	0.12	0.28	0.46	0.81	0.59	**−0.67**	0.97	1.00
	Year 13	0.23	0.38	0.47	−0.41	0.08	0.66	0.48	0.04	0.76	0.29	0.31	0.84	0.46	0.65	**−0.46**	0.99
	Year 14	0.44	−0.86	0.99	−0.33	0.43	1.00	1.00	0.96	0.94	0.56	1.00	1.00	0.88	0.84	0.91	**−0.91**

Table S2. Simulation results of estimation accuracy for fixed effects under case (a). The scores for MSE and MARE have been multiplied by 10 for clear presentation.

Parameter	Criteria	$n = 20$				$n = 50$				$n = 100$			
		$\nu = 5$		$\nu = 50$		$\nu = 5$		$\nu = 50$		$\nu = 5$		$\nu = 50$	
		MT	UT	MT	UT	MT	UT	MT	UT	MT	UT	MT	UT
β_{111}	MSE	4.397	5.836	5.275	5.898	1.614	1.755	1.141	1.280	0.740	0.773	0.539	0.550
	MARE	5.114	5.985	5.764	6.014	3.188	3.326	2.670	2.831	2.060	2.186	1.872	1.930
β_{112}	MSE	0.087	0.138	0.087	0.114	0.038	0.045	0.021	0.030	0.016	0.021	0.019	0.020
	MARE	0.355	0.451	0.373	0.423	0.245	0.265	0.180	0.221	0.157	0.179	0.175	0.180
β_{121}	MSE	4.535	6.403	6.136	7.305	1.664	2.125	2.240	2.801	0.967	1.184	1.055	1.035
	MARE	2.646	3.181	3.069	3.305	1.632	1.805	1.899	2.091	1.309	1.421	1.302	1.339
β_{122}	MSE	0.140	0.197	0.176	0.256	0.059	0.075	0.063	0.094	0.036	0.050	0.032	0.039
	MARE	0.231	0.273	0.256	0.305	0.155	0.176	0.156	0.178	0.114	0.140	0.113	0.122
β_{131}	MSE	6.695	11.609	5.273	8.232	1.702	3.704	1.425	2.598	0.753	1.411	0.760	1.217
	MARE	6.314	8.436	5.741	6.814	3.232	4.689	2.955	3.756	2.147	2.858	2.114	2.662
β_{132}	MSE	0.140	0.232	0.116	0.146	0.046	0.121	0.044	0.111	0.020	0.043	0.023	0.047
	MARE	0.463	0.601	0.421	0.445	0.275	0.423	0.242	0.369	0.173	0.241	0.195	0.242
β_{211}	MSE	3.909	5.426	4.370	4.277	1.326	1.654	1.621	1.541	0.595	0.699	0.730	0.748
	MARE	4.819	5.522	5.321	5.259	2.989	3.335	3.209	3.171	1.988	2.138	2.108	2.171
β_{212}	MSE	0.087	0.138	0.076	0.083	0.029	0.038	0.035	0.042	0.013	0.016	0.014	0.015
	MARE	0.247	0.312	0.229	0.242	0.146	0.166	0.161	0.177	0.100	0.107	0.100	0.101
β_{221}	MSE	6.672	8.743	5.695	6.002	2.214	2.496	1.859	2.054	1.029	1.252	1.018	1.041
	MARE	6.273	7.163	5.886	6.255	3.731	4.045	3.327	3.620	2.505	2.871	2.604	2.605
β_{222}	MSE	0.175	0.207	0.139	0.221	0.055	0.076	0.064	0.099	0.029	0.045	0.024	0.035
	MARE	0.344	0.379	0.320	0.394	0.201	0.219	0.218	0.241	0.135	0.168	0.138	0.165
β_{231}	MSE	5.352	26.720	3.839	5.792	1.754	5.142	1.550	3.280	1.003	2.037	0.977	2.159
	MARE	1.459	2.405	1.234	1.530	0.832	1.239	0.784	1.031	0.649	0.859	0.614	0.860
β_{232}	MSE	0.153	0.732	0.103	0.153	0.052	0.099	0.043	0.073	0.020	0.040	0.029	0.043
	MARE	0.599	1.005	0.554	0.652	0.392	0.492	0.333	0.428	0.237	0.313	0.292	0.364

MT: finite mixture of multivariate t linear mixed model (FM-MtLMM); UT: finite mixture of (univariate) t linear mixed models (FM-tLMMs)

Table S3. Simulation results of estimation accuracy for fixed effects under case (b). The scores for MSE and MARE have been multiplied by 10 for clear presentation.

Parameter	Criteria	$n = 20$				$n = 50$				$n = 100$			
		$\nu = 5$		$\nu = 50$		$\nu = 5$		$\nu = 50$		$\nu = 5$		$\nu = 50$	
		MT	UT	MT	UT	MT	UT	MT	UT	MT	UT	MT	UT
β_{111}	MSE	3.949	5.104	3.478	4.217	1.886	1.643	1.166	1.263	1.090	0.913	0.902	0.871
	MARE	4.907	5.518	4.607	4.956	3.573	3.316	2.731	2.867	2.633	2.531	2.335	2.340
β_{112}	MSE	0.081	0.097	0.089	0.127	0.051	0.040	0.024	0.028	0.016	0.017	0.016	0.017
	MARE	0.367	0.413	0.375	0.438	0.287	0.248	0.202	0.225	0.161	0.171	0.156	0.163
β_{121}	MSE	7.406	10.916	5.517	5.265	3.029	2.201	2.056	2.168	1.746	1.603	1.065	1.104
	MARE	3.495	4.020	2.956	2.798	2.137	1.873	1.850	1.861	1.635	1.591	1.357	1.370
β_{122}	MSE	0.174	0.278	0.154	0.195	0.084	0.089	0.051	0.062	0.044	0.049	0.027	0.034
	MARE	0.269	0.322	0.249	0.282	0.177	0.178	0.147	0.161	0.137	0.146	0.104	0.118
β_{131}	MSE	8.449	16.121	3.626	5.811	2.108	2.350	2.245	2.556	0.854	0.833	0.893	1.414
	MARE	7.096	11.319	4.684	5.908	3.746	3.806	3.709	3.973	2.384	2.358	2.449	2.927
β_{132}	MSE	0.156	1.891	0.096	0.172	0.063	0.076	0.053	0.067	0.021	0.027	0.024	0.032
	MARE	0.493	0.746	0.387	0.492	0.323	0.340	0.280	0.315	0.182	0.212	0.190	0.223
β_{211}	MSE	5.108	4.913	3.075	3.101	2.232	1.924	1.604	1.608	0.864	0.793	0.761	0.778
	MARE	5.775	5.605	4.335	4.304	3.820	3.420	3.028	3.082	2.309	2.198	2.229	2.251
β_{212}	MSE	0.089	0.125	0.065	0.075	0.052	0.044	0.040	0.044	0.018	0.019	0.013	0.016
	MARE	0.250	0.285	0.213	0.230	0.192	0.178	0.167	0.169	0.112	0.114	0.100	0.108
β_{221}	MSE	7.966	8.821	6.368	7.229	3.238	7.690	2.499	2.703	1.531	1.390	0.961	1.054
	MARE	6.918	7.443	6.395	6.859	4.403	6.149	3.918	4.077	3.144	2.992	2.566	2.639
β_{222}	MSE	0.236	0.322	0.157	0.159	0.085	1.075	0.072	0.074	0.053	0.048	0.029	0.034
	MARE	0.390	0.430	0.339	0.337	0.240	0.325	0.221	0.224	0.194	0.183	0.139	0.154
β_{231}	MSE	5.063	6.875	3.881	5.610	2.084	3.526	1.827	2.338	1.250	1.118	0.808	1.227
	MARE	1.403	1.533	1.181	1.418	0.934	1.185	0.860	0.981	0.724	0.671	0.551	0.691
β_{232}	MSE	0.137	0.215	0.127	0.193	0.048	0.101	0.043	0.077	0.025	0.028	0.024	0.030
	MARE	0.625	0.764	0.596	0.680	0.345	0.511	0.345	0.418	0.270	0.281	0.253	0.267

MT: finite mixture of multivariate t linear mixed model (FM-MtLMM); UT: finite mixture of (univariate) t linear mixed models (FM-tLMMs)

Department of Statistics, Graduate Institute of Statistics and Actuarial Science, Feng Chia University, Taichung 40724, Taiwan.

E-mail: wlunwang@fcu.edu.tw