Linear Mixed Model Cheat Sheet

Linear Mixed Model

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\tau} + \boldsymbol{Z}\boldsymbol{u} + \boldsymbol{e} \tag{1}$$

where

- $\boldsymbol{y}$ is a $n \times 1$ vector of observations,
- $\boldsymbol{X}$ is a $n \times p$ design matrix of fixed effects $\boldsymbol{\tau}$ with rank $p_0 \leq p$,
- $\boldsymbol{Z}$ is a $n \times q$ design matrix of random effects $\boldsymbol{u}$, and
- $\boldsymbol{e}$ is a $n \times 1$ vector of random error.

Assume

$$\begin{bmatrix}\boldsymbol{u}\\ \boldsymbol{e}\end{bmatrix} \sim N\left(\begin{bmatrix}\boldsymbol{0}\\ \boldsymbol{0}\end{bmatrix}, \begin{bmatrix}\boldsymbol{G}(\boldsymbol{\kappa}_G) & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{R}(\boldsymbol{\kappa}_R)\end{bmatrix}\right).$$

Mixed Model Equation

$$\underbrace{\begin{bmatrix}\boldsymbol{X}^\top\boldsymbol{R}^{-1}\boldsymbol{X} & \boldsymbol{X}^\top\boldsymbol{R}^{-1}\boldsymbol{Z}\\ \boldsymbol{Z}^\top\boldsymbol{R}^{-1}\boldsymbol{X} & \boldsymbol{Z}^\top\boldsymbol{R}^{-1}\boldsymbol{Z}+\boldsymbol{G}^{-1}\end{bmatrix}}_{(\boldsymbol{W}^\top\boldsymbol{R}^{-1}\boldsymbol{W}+\boldsymbol{G}^*)}\underbrace{\begin{bmatrix}\hat{\boldsymbol{\tau}}\\ \tilde{\boldsymbol{u}}\end{bmatrix}}_{\tilde{\boldsymbol{\beta}}} = \underbrace{\begin{bmatrix}\boldsymbol{X}^\top\boldsymbol{R}^{-1}\boldsymbol{y}\\ \boldsymbol{Z}^\top\boldsymbol{R}^{-1}\boldsymbol{y}\end{bmatrix}}_{\boldsymbol{W}^\top\boldsymbol{R}^{-1}\boldsymbol{y}}$$

$$\left.\begin{aligned}\hat{\boldsymbol{\tau}} &= (\boldsymbol{X}^\top\boldsymbol{V}^{-1}\boldsymbol{X})^-\boldsymbol{X}^\top\boldsymbol{V}^{-1}\boldsymbol{y}\\ \tilde{\boldsymbol{u}} &= \boldsymbol{G}\boldsymbol{Z}^\top\boldsymbol{P}_X\boldsymbol{y} = \boldsymbol{G}\boldsymbol{Z}^\top\boldsymbol{V}^{-1}(\boldsymbol{y}-\boldsymbol{X}\hat{\boldsymbol{\tau}})\end{aligned}\right\}\tilde{\boldsymbol{\beta}} = \boldsymbol{C}^-\boldsymbol{W}^\top\boldsymbol{R}^{-1}\boldsymbol{y}$$

Prediction Error Variance

$$\mathrm{var}(\tilde{\boldsymbol{\beta}}-\boldsymbol{\beta}) = \boldsymbol{C}^-$$

$$\begin{bmatrix}(\boldsymbol{X}^\top\boldsymbol{V}^{-1}\boldsymbol{X})^- & -(\boldsymbol{X}^\top\boldsymbol{V}^{-1}\boldsymbol{X})^-\boldsymbol{X}^\top\boldsymbol{V}^{-1}\boldsymbol{Z}\boldsymbol{G}\\ -\boldsymbol{G}\boldsymbol{Z}^\top\boldsymbol{V}^{-1}\boldsymbol{X}(\boldsymbol{X}^\top\boldsymbol{V}^{-1}\boldsymbol{X})^- & \boldsymbol{G}-\boldsymbol{G}\boldsymbol{Z}^\top\boldsymbol{P}_X\boldsymbol{Z}\boldsymbol{G}\end{bmatrix}$$

Non-full rank $\boldsymbol{X}$

- If $\mathrm{rank}(\boldsymbol{X})$ is less than full-rank then the inverse of $\boldsymbol{X}^\top\boldsymbol{V}^{-1}\boldsymbol{X}$ does not exist.
- Let $\boldsymbol{T}$ be a permutation matrix such that $\boldsymbol{X}\boldsymbol{T} = \begin{bmatrix}\boldsymbol{X}_1 & \boldsymbol{X}_2\end{bmatrix}$ where $\boldsymbol{X}_2$ is full rank.
- Let $\boldsymbol{X}^* = \boldsymbol{X}\boldsymbol{T}$ and $\boldsymbol{\tau}^* = \boldsymbol{T}^\top\boldsymbol{\tau}$. Note that $\boldsymbol{X}\boldsymbol{\tau} = \boldsymbol{X}^*\boldsymbol{\tau}^*$.
- Note $\boldsymbol{X}_1 = \boldsymbol{X}_2\boldsymbol{F}$ for some matrix $\boldsymbol{F}$.
- Note $\boldsymbol{T}$ is orthogonal so $\boldsymbol{T}^{-1} = \boldsymbol{T}^\top$.
- Selected generalise inverse of $\boldsymbol{X}^\top\boldsymbol{V}^{-1}\boldsymbol{X}$ is

$$\boldsymbol{T}\begin{bmatrix}\boldsymbol{0} & \boldsymbol{0}\\ \boldsymbol{0} & (\boldsymbol{X}_2^\top\boldsymbol{V}^{-1}\boldsymbol{X}_2)^{-1}\end{bmatrix}\boldsymbol{T}^\top$$

and we select $\boldsymbol{\tau}_1 = \boldsymbol{0}$.

Miscellaneous

- $\boldsymbol{V} = \boldsymbol{Z}\boldsymbol{G}\boldsymbol{Z}^\top + \boldsymbol{R}$
- $\boldsymbol{P}_X = \boldsymbol{V}^{-1} - \boldsymbol{V}^{-1}\boldsymbol{X}(\boldsymbol{X}^\top\boldsymbol{V}^{-1}\boldsymbol{X})^-\boldsymbol{X}^\top\boldsymbol{V}^{-1}$
- $\boldsymbol{P}_X = \boldsymbol{V}^{-1} - \boldsymbol{V}^{-1}\boldsymbol{X}_2(\boldsymbol{X}_2^\top\boldsymbol{V}^{-1}\boldsymbol{X}_2)^{-1}\boldsymbol{X}_2^\top\boldsymbol{V}^{-1}$
- $\boldsymbol{P}_X = \boldsymbol{R}^{-1} - \boldsymbol{R}^{-1}\boldsymbol{W}\boldsymbol{C}^-\boldsymbol{W}^\top\boldsymbol{R}^{-1}$
- $\boldsymbol{V}^{-1} = \boldsymbol{R}^{-1} - \boldsymbol{R}^{-1}\boldsymbol{Z}(\boldsymbol{Z}^\top\boldsymbol{R}^{-1}\boldsymbol{Z}+\boldsymbol{G}^{-1})^{-1}\boldsymbol{Z}^\top\boldsymbol{R}^{-1}$
- $\boldsymbol{G}\boldsymbol{Z}^\top\boldsymbol{V}^{-1} = (\boldsymbol{Z}^\top\boldsymbol{R}^{-1}\boldsymbol{Z}+\boldsymbol{G}^{-1})^{-1}\boldsymbol{Z}^\top\boldsymbol{R}^{-1}$

Residual

- Marginal residual $\hat{e} = \boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\tau}} = \boldsymbol{Z}\tilde{\boldsymbol{u}} + \boldsymbol{e}$.
- Model-based residual $\tilde{e} = \boldsymbol{R}\boldsymbol{P}_X\boldsymbol{y} = \boldsymbol{R}\boldsymbol{V}^{-1}\hat{e}$
- Conditional residual is $r_i = (\boldsymbol{\delta}_i^\top\boldsymbol{P}_X\boldsymbol{\delta}_i)^{-1}\boldsymbol{\delta}_i^\top\boldsymbol{P}_X\boldsymbol{y}$.
- Studentised conditional residual $t_i = r_i/\sqrt{\mathrm{var}(r_i)}$ where $\mathrm{var}(r_i) = (\boldsymbol{\delta}_i^\top\boldsymbol{P}_X\boldsymbol{\delta}_i)^{-1}$.
- Residual sum of squares $\boldsymbol{y}^\top\boldsymbol{P}\boldsymbol{y}$.

Variance Parameters

Suppose $\boldsymbol{L} = \begin{bmatrix}\boldsymbol{L}_1 & \boldsymbol{L}_2\end{bmatrix}$ is a non-singular matrix and $\boldsymbol{L}_1$ and $\boldsymbol{L}_2$ are $n \times p_0$ and $n \times (n-p_0)$ matrices such that

$$\boldsymbol{L}_1^\top\boldsymbol{X}_2 = \boldsymbol{I}_{p_0} \quad \text{and} \quad \boldsymbol{L}_2^\top\boldsymbol{X}_2 = \boldsymbol{0}.$$

Residual likelihood

$$\begin{aligned}\ell_R &= -\frac{n-p_0}{2}\log 2\pi - \frac{1}{2}\log|\boldsymbol{L}_2^\top\boldsymbol{V}\boldsymbol{L}_2| - \frac{1}{2}\boldsymbol{y}_2^\top(\boldsymbol{L}_2^\top\boldsymbol{V}\boldsymbol{L}_2)^{-1}\boldsymbol{y}_2\\ &= -\frac{n-p_0}{2}\log 2\pi - \frac{1}{2}\log|\boldsymbol{L}^\top\boldsymbol{L}| - \frac{1}{2}|\boldsymbol{V}| - \frac{1}{2}\log|\boldsymbol{X}_2^\top\boldsymbol{V}^{-1}\boldsymbol{X}_2| - \frac{1}{2}\boldsymbol{y}^\top\boldsymbol{P}_X\boldsymbol{y}\\ &= \text{constant w.r.t. } \boldsymbol{\kappa} - \frac{1}{2}|\boldsymbol{V}| - \frac{1}{2}\log|\boldsymbol{X}_2^\top\boldsymbol{V}^{-1}\boldsymbol{X}_2| - \frac{1}{2}\boldsymbol{y}^\top\boldsymbol{P}_X\boldsymbol{y}\end{aligned}$$

$$\ell_R = -\frac{1}{2}\left\{\log|\boldsymbol{G}| + \log|\boldsymbol{R}| + \log|\boldsymbol{C}| + \boldsymbol{y}^\top\boldsymbol{P}_X\boldsymbol{y}\right\}$$

Residual score equations

$$\boldsymbol{u}_R(\kappa_i) = -\frac{1}{2}\left\{\mathrm{tr}\left(\boldsymbol{P}_X\dot{\boldsymbol{V}}_i\right) - \boldsymbol{y}^\top\boldsymbol{P}_X\dot{\boldsymbol{V}}_i\boldsymbol{P}_X\boldsymbol{y}\right\}$$

where $\dot{\boldsymbol{V}}_i = \partial\boldsymbol{V}/\partial\kappa_i$.

---

 emi.tanaka@sydney.edu.au

$$u_R(\kappa_{R_i}) = -\frac{1}{2}\left(\mathrm{tr}\left(\boldsymbol{R}^{-1}\dot{\boldsymbol{R}}_i\right) - \mathrm{tr}\left(\boldsymbol{C}^{-}\boldsymbol{W}^\top\boldsymbol{R}^{-1}\dot{\boldsymbol{R}}_i\boldsymbol{R}^{-1}\boldsymbol{W}\right) - \tilde{\boldsymbol{e}}\boldsymbol{R}^{-1}\dot{\boldsymbol{R}}_i\boldsymbol{R}^{-1}\tilde{\boldsymbol{e}}\right)$$
$$u_R(\kappa_{G_i}) = -\frac{1}{2}\left(\mathrm{tr}\left(\boldsymbol{G}^{-1}\dot{\boldsymbol{G}}_i\right) - \mathrm{tr}\left(\boldsymbol{G}^{-1}\dot{\boldsymbol{G}}_i\boldsymbol{G}^{-1}\boldsymbol{C}^{ZZ}\right) - \tilde{\boldsymbol{u}}^\top\boldsymbol{G}^{-1}\dot{\boldsymbol{G}}_i\boldsymbol{G}^{-1}\tilde{\boldsymbol{u}}\right)$$

where $\dot{\boldsymbol{R}}_i = \partial\boldsymbol{R}/\partial\kappa_{R_i}$ and $\dot{\boldsymbol{G}}_i = \partial\boldsymbol{G}/\partial\kappa_{G_i}$.

Quasi Newton-Raphson algorithm

$$\begin{aligned}\boldsymbol{\kappa}^{(m+1)} &= \boldsymbol{\kappa}^{(m)} - \left[\frac{\delta\boldsymbol{u}_R(\boldsymbol{\kappa})}{\delta\boldsymbol{\kappa}^\top}\right]^{-1}_{\boldsymbol{\kappa}=\boldsymbol{\kappa}^{(m)}}\boldsymbol{u}_R(\boldsymbol{\kappa}^{(m)})\\ &= \boldsymbol{\kappa}^{(m)} + \left[\boldsymbol{\mathcal{I}}^{(m)}\right]^{-1}\boldsymbol{u}_R(\boldsymbol{\kappa}^{(m)})\end{aligned}$$

Information Matrix

$$\begin{aligned}\boldsymbol{\mathcal{I}}_o(\kappa_i,\kappa_j) &= \frac{1}{2}\left\{\mathrm{tr}\left(\boldsymbol{P}_X\dot{\boldsymbol{V}}_{ij}\right) - \mathrm{tr}\left(\boldsymbol{P}_X\dot{\boldsymbol{V}}_i\boldsymbol{P}_X\dot{\boldsymbol{V}}_j\right) + \boldsymbol{y}^\top\boldsymbol{P}_X\dot{\boldsymbol{V}}_i\boldsymbol{P}_X\dot{\boldsymbol{V}}_j\boldsymbol{P}_X\boldsymbol{y} - \boldsymbol{y}^\top\boldsymbol{P}_X\dot{\boldsymbol{V}}_{ij}\boldsymbol{P}_X\boldsymbol{y}\right\}\\ \boldsymbol{\mathcal{I}}_e(\kappa_i,\kappa_j) &= \frac{1}{2}\mathrm{tr}\left(\boldsymbol{P}_X\dot{\boldsymbol{V}}_i\boldsymbol{P}_X\dot{\boldsymbol{V}}_j\boldsymbol{P}_X\right)\\ \boldsymbol{\mathcal{I}}_A(\kappa_i,\kappa_j) &= \frac{1}{2}\boldsymbol{y}^\top\boldsymbol{P}_X\dot{\boldsymbol{V}}_i\boldsymbol{P}_X\dot{\boldsymbol{V}}_j\boldsymbol{P}_X\boldsymbol{y} = \frac{1}{2}\boldsymbol{Q}^\top\boldsymbol{P}_X\boldsymbol{Q}\end{aligned}$$

where $\boldsymbol{Q} = \begin{bmatrix}\boldsymbol{q}_1 & \cdots & \boldsymbol{q}_{n_\kappa}\end{bmatrix}$ and $\boldsymbol{q}_i = \dot{\boldsymbol{V}}_i\boldsymbol{P}_X\boldsymbol{y}$ (the working vector for κ_i).

Augmented MMEs

$$\boldsymbol{M} = \begin{bmatrix}\boldsymbol{C}_{QQ} & \boldsymbol{C}_{yQ}^\top & \boldsymbol{C}_{XQ}^\top & \boldsymbol{C}_{ZQ}^\top\\ \boldsymbol{C}_{yQ} & \boldsymbol{C}_{yy} & \boldsymbol{C}_{Xy}^\top & \boldsymbol{C}_{Zy}^\top\\ \boldsymbol{C}_{XQ} & \boldsymbol{C}_{Xy} & \boldsymbol{C}_{XX} & \boldsymbol{C}_{XZ}\\ \boldsymbol{C}_{ZQ} & \boldsymbol{C}_{Zy} & \boldsymbol{C}_{ZX} & \boldsymbol{C}_{ZZ}\end{bmatrix}$$

where $\boldsymbol{C}_{QQ} = \boldsymbol{Q}^\top\boldsymbol{R}^{-1}\boldsymbol{Q}$, $\boldsymbol{C}_{XQ} = \boldsymbol{X}^\top\boldsymbol{R}^{-1}\boldsymbol{Q}$, and so on. Following the absorption of $\boldsymbol{C}$ provides $\tilde{\boldsymbol{\beta}}$, $\boldsymbol{C}_+^{-1}$ (and $\boldsymbol{C}^-$) as well as $\log|\boldsymbol{C}|_+$.

Mean Shift Outlier Model

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\tau} + \boldsymbol{Z}\boldsymbol{u} + \boldsymbol{\delta}_I\phi_I + \boldsymbol{e} \tag{2}$$

where $\boldsymbol{\delta}_I$ is a $n\times 1$ binary vector where the $i\in I$ positions is 1 and 0 else where and ϕ_I is the fixed mean shift effect for the set I .

Solution

$$\begin{aligned}\hat{\boldsymbol{\tau}} &= (\boldsymbol{X}^\top\boldsymbol{P}_{\delta_I}\boldsymbol{X})^{-}\boldsymbol{X}^\top\boldsymbol{P}_{\delta_I}\boldsymbol{y}\\ \hat{\phi}_I &= (\boldsymbol{\delta}_I^\top\boldsymbol{P}_X\boldsymbol{\delta}_I)^{-}\boldsymbol{\delta}_I^\top\boldsymbol{P}_X\boldsymbol{y}\\ \tilde{\boldsymbol{u}} &= \boldsymbol{G}\boldsymbol{Z}^\top\boldsymbol{V}^{-1}(\boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\tau}} - \boldsymbol{\delta}_I\hat{\phi}_I)\end{aligned}$$

where $\boldsymbol{P}_{\boldsymbol{\delta}_I} = \boldsymbol{V}^{-1} - \boldsymbol{V}^{-1}\boldsymbol{\delta}_I(\boldsymbol{\delta}_I^\top\boldsymbol{V}^{-1}\boldsymbol{\delta}_I)^{-}\boldsymbol{\delta}_I^\top\boldsymbol{V}^{-1}$.

Variance Shift / Alternative Outlier Model

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\tau} + \boldsymbol{Z}\boldsymbol{u} + \boldsymbol{\delta}_I o_I + \boldsymbol{e} \tag{3}$$

where $o_I \sim N(0, \sigma_o^2\boldsymbol{I}_{|I|})$.

Solution

$$\begin{aligned}\hat{\boldsymbol{\tau}} &= (\boldsymbol{X}^\top(\boldsymbol{V} + \sigma_o^2\boldsymbol{I}_{|I|})^{-1}\boldsymbol{X})^{-}\boldsymbol{X}^\top(\boldsymbol{V} + \sigma_o^2\boldsymbol{I}_{|I|})^{-1}\boldsymbol{y}\\ \tilde{o}_I &= (\boldsymbol{\delta}_I^\top\boldsymbol{P}_X\boldsymbol{\delta}_I)^{-}\boldsymbol{\delta}_I^\top\boldsymbol{P}_X\boldsymbol{y}\\ \tilde{\boldsymbol{u}} &= \boldsymbol{G}\boldsymbol{Z}^\top(\boldsymbol{V} + \sigma_o^2\boldsymbol{I}_{|I|})^{-1}(\boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\tau}})\end{aligned}$$

where $\boldsymbol{P}_{\boldsymbol{\delta}_I} = \boldsymbol{V}^{-1} - \boldsymbol{V}^{-1}\boldsymbol{\delta}_I(\boldsymbol{\delta}_I^\top\boldsymbol{V}^{-1}\boldsymbol{\delta}_I)^{-}\boldsymbol{\delta}_I^\top\boldsymbol{V}^{-1}$.

References

[1] Charles Roy Henderson. Estimation of changes in herd environment. *Journal of Dairy Science*, 32(Abstract):706, 1949.

[2] Alison B Smith, Brian R Cullis, and Robin Thompson. The analysis of crop cultivar breeding and evaluation trials: an overview of current mixed model approaches. *The Journal of Agricultural Science*, 143(06):449—-462, 2005.