Mixed Effects Model Estimation

Ryan Brill

August 2025

1 Model

We consider the mixed effects model

$$y = X\beta + Zb + \varepsilon, \tag{1.1}$$

where

$y \in \mathbb{R}^n$	response vector,
$X \in \mathbb{R}^{n\times p}$	fixed effects design matrix,
$Z \in \mathbb{R}^{n\times q}$	random effects design matrix,
$\beta \in \mathbb{R}^p$	fixed effects coefficients,
$b \in \mathbb{R}^q$	random effects coefficients,
$\varepsilon \in \mathbb{R}^n$	residual noise.

Distributional assumptions.

$$\begin{aligned} b &\sim N(0, \Sigma_\theta), \\ \varepsilon &\sim N(0, \sigma^2 I_n), \\ b &\perp \varepsilon, \\ X, Z &\text{ known and fixed}, \\ \beta &\text{ unknown and fixed}. \end{aligned}$$

Covariance structure. $\Sigma_\theta \in \mathbb{R}^{q\times q}$ is a symmetric positive semidefinite covariance matrix. There exists a unique lower triangular matrix Λ_θ such that

$$\Sigma_\theta = \Lambda_\theta \Lambda_\theta^T.$$

The parameters θ are the $q(q+1)/2$ nonzero lower triangular entries of Λ_θ to be estimated.

2 Two Log-Likelihoods

2.1 Maximum Likelihood (ML)

We first find the marginal distribution of y:

$$\begin{aligned} y \mid b &\sim N(X\beta + Zb, \sigma^2 I_n), \\ b &\sim N(0, \Sigma_\theta) \\ \Rightarrow Zb &\sim N(0, Z\Sigma_\theta Z^T), \\ \Rightarrow y &\sim N\left(X\beta, V_\theta\right), \end{aligned}$$

where

$$V_\theta = Z\Sigma_\theta Z^T + \sigma^2 I_n. \tag{2.1}$$

The log-likelihood is

$$\ell_{\mathrm{ML}}(\beta, \theta, \sigma^2) \propto -\frac{1}{2}\left[\log|V_\theta| + (y - X\beta)^T V_\theta^{-1}(y - X\beta)\right]. \tag{2.2}$$

2.2 Restricted Maximum Likelihood (REML)

Now, we will integrate out β. Find an error contrast matrix $A \in \mathbb{R}^{n\times(n-p)}$ satisfying

$$A^T X = 0, \quad A^T A = I_{n-p}.$$

Such an A exists since X is full rank, by the Rank–Nullity Theorem: $\mathrm{Nullity}(X^T) = \mathrm{Nullity}(X) = n - \mathrm{Rank}(X) = n - p$.

Then

$$A^T y \sim N_{n-p}\left(0, A^T V_\theta A\right),$$

which is free of β. The REML log-likelihood becomes

$$\ell_{\mathrm{REML}}(\theta, \sigma^2) \propto -\frac{1}{2}\left[\log|A^T V_\theta A| + y^T P_\theta y\right], \tag{2.3}$$

where

$$P_\theta = A(A^T V_\theta A)^{-1} A^T.$$

A ton of linear algebra (not shown) will allow you to re-write the REML and P_θ just in terms of V_θ and X, without A:

$$\ell_{\mathrm{REML}}(\theta, \sigma^2) \propto -\frac{1}{2}\left[\log|V_\theta| + \log|X^T V_\theta^{-1} X| + y^T P_\theta y\right], \tag{2.4}$$

where

$$P_\theta = V_\theta^{-1} - V_\theta^{-1} X(X^T V_\theta^{-1} X)^{-1} X^T V_\theta^{-1}.$$

3 Estimation Procedure

1. Initialize θ (e.g., entries of Λ_θ) and σ^2 (e.g., $\widehat{\mathrm{Var}}(y)$).
2. Given (θ, σ^2), estimate β by maximum likelihood via GLS:

$$\hat{\beta} = (X^T V_\theta^{-1} X)^{-1} X^T V_\theta^{-1} y, \qquad V_\theta = Z \Sigma_\theta Z^T + \sigma^2 I_n.$$

3. Update the variance parameters: maximize $\ell_{\mathrm{REML}}(\theta, \sigma^2)$ numerically.
4. Iterate 2-3 until convergence (e.g., relative parameter & objective change $< 10^{-6}$).
5. After convergence (reporting/inference):
   - Fixed effects: $\hat{\beta}$ with $\mathrm{Var}(\hat{\beta}) = (X^T V_{\hat{\theta}}^{-1} X)^{-1}$.
   - Random effects (BLUPs): $\hat{b} = \Sigma_{\hat{\theta}} Z^T V_{\hat{\theta}}^{-1}(y - X\hat{\beta})$ with conditional $\mathrm{Var}(b \mid y) = (\Sigma_{\hat{\theta}}^{-1} + Z^T(\hat{\sigma}^{-2} I) Z)^{-1}$.
6. Numerical tips. Use Cholesky factorizations for solves and log-determinants; avoid explicit inverses.