

Estimating Parameters in Linear Mixed-Effects Models

A linear mixed-effects model is of the form

$$y = \underbrace{X\beta}_{\text{fixed}} + \underbrace{Zb}_{\text{random}} + \underbrace{\varepsilon}_{\text{error}},$$

where

- y is the n -by-1 response vector, and n is the number of observations.
- X is an n -by- p fixed-effects design matrix.
- β is a p -by-1 fixed-effects vector.
- Z is an n -by- q random-effects design matrix.
- b is a q -by-1 random-effects vector.
- ε is the n -by-1 observation error vector.

The random-effects vector, b , and the error vector, ε , are assumed to have the following independent prior distributions:

$$\begin{aligned} b &\sim N(0, \sigma^2 D(\theta)), \\ \varepsilon &\sim N(0, \sigma^2 I), \end{aligned}$$

where D is a symmetric and positive semidefinite matrix, parameterized by a variance component vector θ , I is an n -by- n identity matrix, and σ^2 is the error variance.

In this model, the parameters to estimate are the fixed-effects coefficients β , and the variance components θ and σ^2 . The two most commonly used approaches to parameter estimation in linear mixed-effects models are maximum likelihood and restricted maximum likelihood methods.

Maximum Likelihood (ML)

The maximum likelihood estimation includes both regression coefficients and the variance components, that is, both fixed-effects and random-effects terms in the likelihood function.

For a linear mixed-effects model defined above, the conditional response of the response variable y given β , b , θ , and σ^2 is

$$y|b, \beta, \theta, \sigma^2 \sim N(X\beta + Zb, \sigma^2 I_n).$$

The likelihood of y given β , θ , and σ^2 is

$$P(y|\beta, \theta, \sigma^2) = \int P(y|b, \beta, \theta, \sigma^2) P(b|\theta, \sigma^2) db,$$

where

$$\begin{aligned} P(b|\theta, \sigma^2) &= \frac{1}{(2\pi\sigma^2)^{q/2}} \frac{1}{|D(\theta)|^{1/2}} \exp\left\{-\frac{1}{2\sigma^2} b^T D^{-1} b\right\} \quad \text{and} \\ P(y|b, \beta, \theta, \sigma^2) &= \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left\{-\frac{1}{2\sigma^2} (y - X\beta - Zb)^T (y - X\beta - Zb)\right\}. \end{aligned}$$

Suppose $\Lambda(\theta)$ is the lower triangular Cholesky factor of $D(\theta)$ and $\Delta(\theta)$ is the inverse of $\Lambda(\theta)$. Then,

$$D(\theta)^{-1} = \Delta(\theta)^T \Delta(\theta).$$

Define

$$r^2(\beta, b, \theta) = b^T \Delta(\theta)^T \Delta(\theta) b + (y - X\beta - Zb)^T (y - X\beta - Zb),$$

and suppose b^* is the value of b that satisfies

$$\left. \frac{\partial r^2(\beta, b, \theta)}{\partial b} \right|_{b^*} = 0$$

for given β and θ . Then, the likelihood function is

$$P(y|\beta, \theta, \sigma^2) = (2\pi\sigma^2)^{-n/2} |D(\theta)|^{-1/2} \exp\left\{-\frac{1}{2\sigma^2} r^2(\beta, b^*(\beta), \theta)\right\} \frac{1}{|\Delta^T \Delta + Z^T Z|^{1/2}}.$$

$P(y|\beta, \theta, \sigma^2)$ is first maximized with respect to β and σ^2 for a given θ . Thus the optimized solutions $\hat{\beta}(\theta)$ and $\hat{\sigma}^2(\theta)$ are obtained as functions of θ . Substituting these solutions into the likelihood function produces $P(y|\hat{\beta}(\theta), \theta, \hat{\sigma}^2(\theta))$. This expression is called a profiled likelihood where β and σ^2 have been profiled out. $P(y|\hat{\beta}(\theta), \theta, \hat{\sigma}^2(\theta))$ is a function of θ , and the algorithm then optimizes it with respect to θ . Once it finds the optimal estimate of θ , the estimates of β and σ^2 are given by $\hat{\beta}(\theta)$ and $\hat{\sigma}^2(\theta)$.

The ML method treats β as fixed but unknown quantities when the variance components are estimated, but does not take into account the degrees of freedom lost by estimating the fixed effects. This causes ML estimates to be biased with smaller variances. However, one advantage of ML over REML is that it is possible to compare two models in terms of their fixed- and random-effects terms. On the other hand, if you use REML to estimate the parameters, you can only compare two models, that are nested in their random-effects terms, with the same fixed-effects design.

Restricted Maximum Likelihood (REML)

Restricted maximum likelihood estimation includes only the variance components, that is, the parameters that parameterize the random-effects terms in the linear mixed-effects model. β is estimated in a second step. Assuming a uniform improper prior distribution for β and integrating the likelihood $P(y|\beta, \theta, \sigma^2)$ with respect to β results in the restricted likelihood $P(y|\theta, \sigma^2)$. That is,

$$P(y|\theta, \sigma^2) = \int P(y|\beta, \theta, \sigma^2) P(\beta) d\beta = \int P(y|\beta, \theta, \sigma^2) d\beta.$$

The algorithm first profiles out $\hat{\sigma}_R^2$ and maximizes remaining objective function with respect to θ to find $\hat{\theta}_R$. The restricted likelihood is then maximized with respect to σ^2 to find $\hat{\sigma}_R^2$. Then, it estimates β by finding its expected value with respect to the posterior distribution

$$P(\beta|y, \hat{\theta}_R, \hat{\sigma}_R^2).$$

REML accounts for the degrees of freedom lost by estimating the fixed effects, and makes a less biased estimation of random effects variances. The estimates of θ and σ^2 are invariant to the value of β and less sensitive to outliers in the data compared to ML estimates. However, if you use REML to estimate the parameters, you can only compare two models that have the identical fixed-effects design matrices and are nested in their random-effects terms.

References

- [1] Pinheiro, J. C., and D. M. Bates. *Mixed-Effects Models in S and S-PLUS*. Statistics and Computing Series, Springer, 2004.
- [2] Hariharan, S. and J. H. Rogers. "Estimation Procedures for Hierarchical Linear Models." *Multilevel Modeling of Educational Data* (A. A. Connell and D. B. McCoach, eds.). Charlotte, NC: Information Age Publishing, Inc., 2008.
- [3] Raudenbush, S. W. and A. S. Bryk. *Hierarchical Linear Models: Applications and Data Analysis Methods*, 2nd ed. Thousand Oaks, CA: Sage Publications, 2002.
- [4] Hox, J. *Multilevel Analysis, Techniques and Applications*. Lawrence Erlbaum Associates, Inc, 2002.
- [5] Snijders, T. and R. Bosker. *Multilevel Analysis*. Thousand Oaks, CA: Sage Publications, 1999.
- [6] McCulloch, C.E., R. S. Shayle, and J. M. Neuhaus. *Generalized, Linear, and Mixed Models*. Wiley, 2008.

See Also

[LinearMixedModel](#) | [fitlme](#) | [fitlmematrix](#)

Related Topics

- [Linear Mixed-Effects Models](#)