

Nonlinear mixed models — Documentation and Methods

Nonlinear mixed-effects models—What are they and why use them?

“Nonlinear mixed-effects (NLME) models are mixed-effects models in which some, or all, of the fixed and random effects occur nonlinearly in the model function (Pinheiro and Bates 2000, p. 306).”

- Extension of linear mixed-effects (LME) models in which the conditional expectation of the response given the random effects is allowed to be a nonlinear function
- Extension of nonlinear regression models for independent data in which random effects are incorporated in the coefficients to allow them to vary by group

LME vs. NLME

Why use NLME models instead of LME models?

Interpretability, parsimony, and validity beyond the observed range of the data

Costs of using NLME instead of LME models:

- Marginal likelihood function has no closed-form expression
 - approximate likelihood function used for parameter estimation
 - estimation algorithms more computationally extensive
 - less reliable inference results
- Starting estimates for fixed-effects coefficients required

Similarities of NLME and LME models:

- Both models used with grouped data
- Describe a response variable as a function of covariates, taking into account the correlation among observations in the same group
- Random effects used to represent within-group dependence
- Assumptions about random effects and within-group errors identical
(Pinheiro and Bates 2000, p. 274-277)

NLS vs. NLME

Nonlinear regression models with fixed effects only are fit by nonlinear least squares (NLS). Using NLS instead of NLME for grouped data prevents us from understanding the true structure of the data and from considering different sources of variability that may be of interest in themselves (Pinheiro and Bates 2000, p. 279-280).

Notation and some theory

The notation used is based on Pinheiro and Bates (2000).

Notation

For a single-level of grouping, the j th observation on the i th group is modeled as

$$y_{ij} = f(\phi_{ij}, \nu_{ij}) + \varepsilon_{ij} \quad i = 1, \dots, M, \quad j = 1, \dots, n_i \quad [1]$$

where M is the number of groups, n_i is the number of observations on the i th group, $\varepsilon_{ij} \sim N(0, \sigma^2)$ is the within-group error, and f is a general, real-valued, differentiable function of a group-specific parameter vector ϕ_{ij} and a covariate vector ν_{ij} , which is nonlinear in at least one component of ϕ_{ij} , which is modeled as:

$$\phi_{ij} = A_{ij}\beta + B_{ij}b_i \quad b_i \sim N(0, \Psi) \quad [2]$$

β is a p -dimensional vector of fixed effects and b_i is a q -dimensional vector of random effects associated with the i th group with variance-covariance matrix Ψ . Matrices A_{ij} and B_{ij} are of appropriate dimensions and depend on the group and on the values of covariates at the j th observation.

This single-level model can be extended to data grouped according to multiple, nested factors by modifying the model for the random effects. For example, a model with two levels of nesting the k th observation on the j th second-level group of the i th first-level group is modeled as:

$$y_{ijk} = f(\phi_{ijk}, \nu_{ijk}) + \varepsilon_{ijk} \quad i = 1, \dots, M, \quad j = 1, \dots, M_i, \quad k = 1, \dots, n_{ij} \quad [3]$$

M is the number of first-level groups, M_i is the number of second-level groups within the i th first-level group, n_{ij} is the number of observations in the j th second-level group of the i th first-level group, and ε_{ijk} is the normally distributed within-group error term. The group-specific parameter vector ϕ_{ijk} is expressed as:

$$\phi_{ijk} = A_{ijk}\beta + B_{i,jk}b_i + B_{ijk}b_{ij} \quad b_i \sim N(0, \Psi_1), \quad b_{ij} \sim N(0, \Psi_2) \quad [4]$$

Likelihood estimation

The maximum likelihood (ML) estimation in mixed-effects models is based on the marginal density of the response \mathbf{y} . For a model with Q levels of nesting it is calculated as:

$$p(\mathbf{y} | \beta, \sigma^2, \Psi_1, \dots, \Psi_Q) = \int p(\mathbf{y} | \mathbf{b}, \beta, \sigma^2) p(\mathbf{b} | \Psi_1, \dots, \Psi_Q) d\mathbf{b} \quad [5]$$

where $p(\mathbf{y} | \beta, \sigma^2, \Psi_1, \dots, \Psi_Q)$ is the marginal density of \mathbf{y} , $p(\mathbf{y} | \mathbf{b}, \beta, \sigma^2)$ is the conditional density of \mathbf{y} given the random effects \mathbf{b} , and $p(\mathbf{b} | \Psi_1, \dots, \Psi_Q)$ is the marginal distribution of \mathbf{b} .

A closed-form expression for the integral in equation 5 is generally not available since non-linearity prevents integration over the random effects.

→ approximation needed, e.g.,

- Lindstrom and Bates (1990): LME approximation to the log-likelihood function → currently implemented in `nlme()`

For more details see Pinheiro and Bates (2000, Section 7.2).

Inference and predictions

Inference results based on the LME approximation to the log-likelihood function by Lindstrom and Bates (1990)

Compare nested NLME models through likelihood ratio tests

Note: For REML estimation only models with identical fixed and random effects structures can be compared! REML = restricted maximum likelihood, default estimation method for `lme()`.

→ fit different nested models in which the random-effects structure changes and apply likelihood-ratio tests

Information criterion statistics (e.g., AIC and BIC; ‘the smaller the value the better the model’) can be used to compare NLME models (nested and non-nested)

Hypotheses about fixed effects should be tested using t and F tests

Inference results for NLME models based on the LME approximation are ‘approximately asymptotic’ → less reliable than the asymptotic inference results for LME models.

Predictions for `nlme` models can be obtained at different levels of nesting or at the population level

Assumptions

1. Observations corresponding to different groups are independent

2. Within-group errors ε_{ij} are independent and identically distributed as $N(0, \sigma^2)$ and independent of the random effects b_i

- a. Test assumption of equal variances with residual plots and normality assumption with normal probability plots for residuals
- b. Remedies: variance-covariance structure of the within-group errors can be decomposed into two independent components: a variance structure and a correlation structure (for details see Pinheiro and Bates 2000, Chapter 5; Zuur et al. 2009, Chapters 4, 6, and 7)
 - use variance functions available in `nlme()` to model within-group heteroscedasticity, e.g.,
 - fixed variance
 - different variances per stratum
 - power of covariate

- exponential of covariate
- use correlation structures available `nlme()` in to represent the correlation structure of within-group errors, e.g.,
 - compound symmetry
 - general
 - autoregressive of order 1
 - continuous-time AR(1)
 - autoregressive-moving average
- 3. Random effects are normally distributed with mean zero and variance-covariance matrix Ψ**
 - a. Test assumption with normal probability plots for deviations of coefficients from average
- 4. Random effects corresponding to different groups are independent**

Using covariates with nlme

Random effects in mixed-effects model represent deviations of the individual parameters from the fixed effects

These deviations can arise from unexplained intergroup variation, but often can be partially explained by differences in covariate values among groups

Including covariates in the model to explain intergroup variation generally reduces number of random effects in the model and leads to a better understanding of the mechanisms producing the response

Systematic pattern in a given random effect with respect to a covariate would indicate that the covariate should be included in the model

Model building strategies

Starting values

Nonlinear mixed-effects models require starting estimates for the fixed parameters

Some general recommendations for determining reasonable starting values are available:

- Choose parameters that have meaningful graphical interpretations → determine initial estimates from a plot if parameters have graphical interpretation
- Take advantage of partially linear models so that initial estimates are needed only for those parameters that enter the model nonlinearly
- Refine estimates of some parameters by iterating on them while holding all other parameters fixed
- Fit nonlinear fixed-effects model and use parameter estimates as starting values

Finding starting values can be tedious; if possible use `selfStart` functions = functions that contain an auxiliary function to calculate initial parameter estimates

For more details on starting values see, for example, Pinheiro and Bates (2000) and Ritz and Streibig (2008, Chapter 3).

Which effects should be fixed and which should be random?

Crucial step in model-building: decide which of the coefficients in the model need random effects to account for their between-subject variation and which can be treated as purely fixed effects

- plot of individual confidence intervals from `nlsList()` fit
- Boxplots of residuals by group → if residuals mostly negative for some groups and mostly positive for other groups, then include a random group effect

First possible modeling strategy: start out with model that has random effects for all parameters

- convergence issues when the number of random effects is large relative to the number of groups
- assuming diagonal ψ in model helps to avoid convergence problems
- Near-zero estimate for standard deviation of a random parameter → term could be dropped from model
- Use augmented prediction plots to determine whether NLME model can accommodate group effects

Second possible modeling strategy: fit baseline model (minimum fixed and random effects)

- Add random effects
 - Include heteroscedastic variance structure
 - Correlation structure
 - Extra random effects (e.g., random slope)
- Once a reasonable model is found, add more fixed effects

Note: Fixed effects explain variation, random effects organize unexplained variation. Adding random effects adds information and improves diagnostic compatibility, but explains no more variation!

nlme function in R

Nonlinear mixed-effects models are fit in R using the `nlme()` function in the `nlme` package.

typical call: `nlme(model, data, fixed, random, groups, start)`

`fixed` and `random` are formulas defining the structures of the fixed and random effects in the model, they are used to specify the matrices A_{ij} and B_{ij} in Equation 2.

Each parameter in the model usually has an associated fixed effect, but may or may not have an associated random effect

By default, if `random` is omitted, all fixed effects in model are assumed to have associated random effect

`start` → starting values for the fixed effects must be given; starting estimates for remaining parameters are generated automatically; finding good starting values is important!!!

The default method is `method="ML"` maximizing the log-likelihood; `method="REML"` maximizes the restricted log-likelihood

Note: `nlme()` default is `method="ML"` while `lme()` default is `method="REML"`

Objects returned by `nlme` are of class `nlme` and inherit from class `lme` → all methods available for `lme` objects can be applied to `nlme` objects!

Variance functions and correlation structures for `nlme()` are specified in the same way as for `lme()` → Same diagnostic plots available

`nlme` fits can be very slow! → setting `verbose=TRUE` helps to monitor progress

Some convergence error messages can be resolved by altering the `control` argument (e.g., increase number of iterations); see `?nlmeControl` for more details

References

- Lindstrom, M.J. and Bates, D.M. 1990. Nonlinear mixed-effects models for repeated measures data. *Biometrics* 46:673-687.
- Pinheiro, J.C., and Bates, D.M. 2000. Mixed-effects models in S and S-Plus. Springer, New York, 528p.
- Ritz, C. and Streibig, J.C. 2008. Nonlinear regression with R. Springer, New York, 148p.
- Venables, W.N., and Ripley, B.D. 2002. Modern applied statistics with S. Springer, New York, 495p.
- Zuur, A.F., Ieno, E.N., Walker, N.J., Saveliev, A.A., and Smith, G.M. 2009. Mixed effects models and extensions in Ecology with R. Springer, New York, 574p.