nonlinear mixed effect model fitting with nlme

David Lamparter

March 29, 2010

Purpose of nonlinear mixed effects modeling

nonlinearity

- fitting to mechanistic or semimechanistic model with fixed number of parameters
- parsimonious model-specification, few parameters.

mixed effects modeling

- data has grouping structure and parameter estimates are allowed to vary among groups.
- for parsimonious modeling: Parameter variation is modeled by an underlying distribution
- gives information about variation of parameter values between groups.

single level nlme model

$$y_{ij} = f(\phi_i, \boldsymbol{\nu}_{ij}) + \varepsilon_{ij} \quad i = 1, \ldots, M \quad j = 1, \ldots, n_i \tag{1}$$

where ϕ_i is a group-specific parameter vector. $\boldsymbol{\nu}_{ij}$ is a covariate vector and $\varepsilon_{ij} \sim \mathcal{N}_1(0, \sigma^2)$. M is the number of groups, and n_i the number of observations within a group

ϕ_i is modeled via

$$\phi_i = \boldsymbol{A}\beta + \boldsymbol{B}\boldsymbol{b}_i \quad \boldsymbol{b}_i \sim \mathcal{N}(0, \Psi) \tag{2}$$

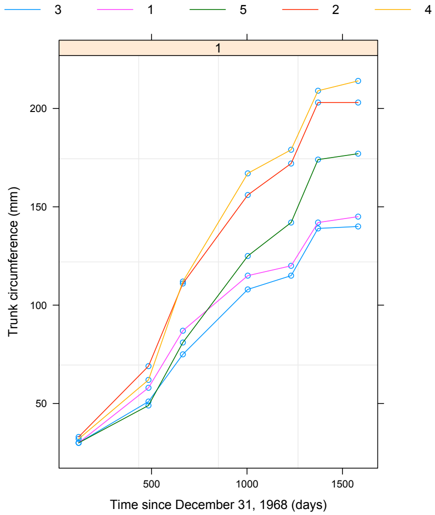
and $\varepsilon_{ij} \perp \boldsymbol{b}_i \quad \forall i, j$

(note: slight generalization later on.)

Orange tree example(1)

Growth of Orange trees

Data: trunk circumference of 5 trees measured over time.



Growth of Orange trees
Data: trunk circumference of 5 trees measured over time.

model:

$$y_{ij} = \frac{\phi_1}{1 + \exp[-(t_{ij} - \phi_2)/\phi_3]} + \varepsilon_{ij} \qquad (3)$$

- ϕ_1: asymptotic height: $t \to \infty \Rightarrow y_{ij} = \phi_1 + \varepsilon_{ij}$ (Asym)
- ϕ_2: time at half-asymptotic height: $t = \phi_2 \Rightarrow y_{ij} = \phi_1/2 + \varepsilon_{ij}$ (xmid)
- ϕ_3: time between $1/2$ and $3/4$ of asymptotic height. (scal)

allows for simple heuristic to find starting estimates.
more elaborate heuristic implemented in function SSlogis:

SSlogis algorithm

Conditional linearizability

1. scale reponse variable y to (0,1)-interval: new response y'

$$y' \approx \frac{1}{1 + \exp[(\phi_2 - x)/\phi_3]}$$

2. take logistic transformation: $z := \log[y'/(1 - y')]$

$$z \approx -(\phi_2 - x)/\phi_3$$

3. fit linear regression for $x = a + bz$. choose $\phi_2(0) = a$, $\phi_3(0) = b$
4. use algorithm for partially linear models (see Golub and Pereyra 73) to fit:

$$y = \frac{\phi_1}{1 + \exp[(\phi_2 - x)/\phi_3]}$$

Overview of procedure

1. plot and structure of data
2. Ignore grouping structure at first: nls-function
3. fit model separately for each group: nlsList-function
4. fit non-linear mixed effect model: nlme-function
5. analyse non-linear mixed effect model, *go back to step 4*

$\rightarrow$ R

the nlme-model

$$y_{ij} = f(\phi_i, \boldsymbol{\nu}_{ij}) + \varepsilon_{ij} \quad i = 1, \ldots, M \quad j = 1, \ldots, n_i \tag{4}$$

becomes

$$y_{ij} = \frac{\phi_1}{1 + \exp[-(t_{ij} - \phi_2)/\phi_3]} + \varepsilon_{ij} \tag{5}$$

and

$$\phi_i = \boldsymbol{A}\beta + \boldsymbol{B}\boldsymbol{b}_i \quad \boldsymbol{b}_i \sim \mathcal{N}(0, \Psi) \tag{6}$$

becomes

$$\begin{bmatrix} \phi_{i1} \\ \phi_{i2} \\ \phi_{i3} \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{\mathbf{A}} \times \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} + \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{\mathbf{B}} \times \begin{bmatrix} b_{1i} \\ b_{2i} \\ b_{3i} \end{bmatrix} \tag{7}$$

The nlme function call

```
> nlme(model, data, fixed, random, groups, start)
```

- 'model' can be a two-sided formula an SSlogis function or an nlsList-Object
- 'data', 'start': clear; 'groups' not needed if groups are specified somewhere else
- 'fixed' gives models for the fixed effects: most natural: list of right-hand side formulas, each one corresponding to a row in the fixed effect matrix

  ```
  fixed=list(Asym~1, xmid~1, scal~1)
  ```

  bug-note: doesn't work. However, abreviation works.

  ```
  fixed=Asym+xmid+scal~1
  ```

  Else specify via nlsList

The nlme function call cont'

```
> nlme(model, data, fixed, random, groups, start)
```

'random' works analog to 'fixed' (including bug). Additionally, use 'pdMat'-objects to specify additionally correlation structure:

```
random=pdDiag(list(Asym~1, xmid~1, scal~1) )
```

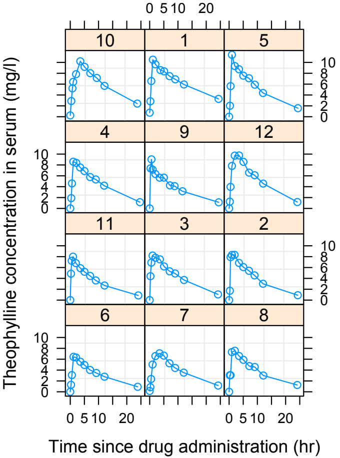
pdMat'-constructor functions available:

- 'pdBlocked': block-diagonal
- 'pdCompSymm': compound-symmetry structure
- 'pdDiag': diagonal
- 'pdIdent': multiple of identity
- 'pdSymm': general positive-definite matrix

→R

the Theophyline example

Setup:
Serum concentration of Theophyline measured in 12 subjects at eleven times after receiving an oral dose.



the Theophyline example model

Model:

$$c_t = \frac{Dk_e k_a}{Cl(k_a - k_e)}[\exp(-k_e t) - \exp(-k_a t)] \tag{8}$$

log-transformed version (ensure positive estimates):

$$c_t = \frac{DlK_e + lK_a - Cl}{\exp lK_a - \exp(lK_e)}(\exp[\exp(lK_e)t] - \exp[-\exp(lKa)t]) \tag{9}$$

where $lK_e = \log(k_e)$, $lK_a = \log(k_a)$ and $lCl = \log(Cl)$

$\rightarrow$R

Using covariates

Random effects model deviations of individual parameter from the fixed effect. But deviation might be explainable by covariate values among groups

example In the Theophyline example also weight of subject is known. Assume, that the subject specific absorbtion rate lKa_i depends linearly on weight Wt_i:

$$\phi_i = \begin{bmatrix} lKe_i \\ lKa_i \\ lCl_i \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & Wt_i \\ 0 & 0 & 1 & 0 \end{bmatrix}}_{\boldsymbol{A}_i} \times \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{bmatrix} + \boldsymbol{B}\boldsymbol{b}_i \qquad (10)$$

Then some variation in lKa_i is explained by weight Wt_i

```
fixed=list(lKe~1, lKa~Wt, lCl~1)
```

Using covariates cont'

Need for more general model formulation:

$$y_{ij} = f(\phi_{ij}, \boldsymbol{\nu}_{ij}) + \varepsilon_{ij} \quad i = 1, \ldots, M \quad j = 1, \ldots, n_i \tag{11}$$

ϕ_{ij} is modeled via

$$\phi_{ij} = \boldsymbol{A}_{ij}\beta + \boldsymbol{B}_{ij}\boldsymbol{b}_i \quad \boldsymbol{b}_i \sim \mathcal{N}(0, \Psi) \tag{12}$$

i.e.: Matrices are allowed to be functions of covariates.
Note: covariate does not need to be constant within one group.
(That is why Matrices are allowed to vary within each observation)

Using covariates cont'

Variation for the particular parameter is explained away. $\Rightarrow$ often random effects drop out. e.g.:

$$\begin{bmatrix} lKe_i \\ lKa_i \\ lCl_i \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & Wt_i \\ 0 & 0 & 1 & 0 \end{bmatrix} \times \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{bmatrix} + \begin{bmatrix} b_{1i} \\ 0 \\ b_{3i} \end{bmatrix} \tag{13}$$

But sometimes additional random effects lead to better fit:

$$\begin{bmatrix} lKe_i \\ lKa_i \\ lCl_i \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & Wt_i \\ 0 & 0 & 1 & 0 \end{bmatrix} \times \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{bmatrix} + \begin{bmatrix} b_{1i} \\ b_{2i} \\ b_{3i} \\ b_{4i} \end{bmatrix} \tag{14}$$

Recommended heuristic procedure

- use forward stepwise approach, testing covariates one at the time.
- fit model without covariate and plot estimated random effects against covariate
- Compare models as usual (AIC, BIC, likelihood ratio)

In Theophyline example with above setup: Fit model without weight covariate,
plot $\hat{b}_{2i}$ vs. Wt_i

updated Procedure

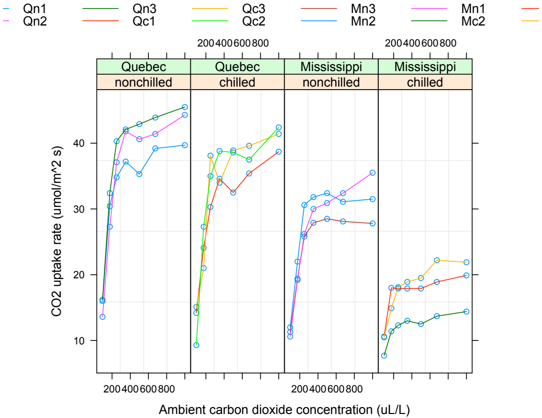
Overview of procedure

1. plot and structure of data
2. Ignore grouping structure at first: nls-function
3. fit model separately for each group: nlsList-function
4. fit non-linear mixed effect model: nlme-function
5. analyse non-linear mixed effect model, *go back to step 4*
6. incorporate Covariates if possible or necessary

Study of cold tolerance in C_4-grass species.

setup:

2 species of grass (Quebec/Missisipi) 6 plants each. Each group divided into 2 groups: control and chilled. (plants were chilled for 14h at $7^\circ C$; after 10h of recovery CO_2 uptake was measured for various ambient CO_2 concentrations.



CO_2 uptake example

Study of cold tolerance in C_4-grass species.

setup:

2 species of grass (Quebec/Missisipi) 6 plants each. Each group divided into 2 groups: control and chilled. (plants were chilled for 14h at $7^\circ C$; after 10h of recovery CO_2 uptake was measured for various ambient CO_2 concentrations.

Model: Offset asymptotic regression model (log-transformed):

$$U(c) = \phi_1(1 - \exp[-\exp(\phi_2)(c - \phi_3)]) \tag{15}$$

- ϕ_1: Asymptote (Asym)
- ϕ_2: log-rate constant (lrc)
- ϕ_3: offset, max CO_2-conc. without uptake (c0)

$\rightarrow$R(flash resulting model)

CO_2 uptake model with covariate

$$\begin{bmatrix} \phi_{i1} \\ \phi_{i2} \\ \phi_{i3} \end{bmatrix} = \begin{bmatrix} 1 & Treat_i & Type_i & Treat_i * Type_i & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} \beta_1 \\ \gamma_1 \\ \gamma_2 \\ \gamma_3 \\ \beta_2 \\ \beta_3 \end{bmatrix} + \begin{bmatrix} b_{1i} \\ b_{2i} \\ 0 \end{bmatrix} \tag{16}$$

where

$$Treat_i = \begin{cases} -1 & \text{Treatment of Plant } i = \text{nochilled} \\ 1 & \text{Treatment of Plant } i = \text{chilled} \end{cases} \tag{17}$$

$$Type_i = \begin{cases} -1 & \text{Type of Plant } i = \text{Quebec} \\ 1 & \text{Type of Plant } i = \text{Mississippi} \end{cases} \tag{18}$$

extended single level nlme model

$$\boldsymbol{y}_i = \boldsymbol{f}(\theta_i, \boldsymbol{\nu}_i) + \varepsilon_i \quad i = 1, \ldots, M \tag{19}$$

$$\phi_i = \boldsymbol{A}_i\beta + \boldsymbol{B}_i\boldsymbol{b}_i \quad \boldsymbol{b}_i \sim \mathcal{N}(0, \Psi) \quad \varepsilon_i \sim \mathcal{N}(0, \sigma\boldsymbol{\Lambda}_i) \tag{20}$$

i.e.: within-group errors are allowed to be correlated and have non-constant variance.

extended single level nlme model: variance functions

Motivation: When plotting residuals against (a) a covariate or (b) the fitted values, we sometimes see trends. How to incorporate this most naturally?

$$\mathrm{Var}(\varepsilon_{ij}|\boldsymbol{b}_i) = \sigma \times g^2(\mu_{ij}, \boldsymbol{\nu_{ij}}, \boldsymbol{\delta}), \quad i = 1, \ldots, M, \quad j = 1, \ldots, n_i \tag{21}$$

- $\mu_{ij} = \mathrm{E}[y_{ij}|\boldsymbol{b}_i]$,
- ν_{ij}: a variance- covariate vector
- $\boldsymbol{\delta}$ is a vector of variance parameters.
- $g(\cdot)$ is the variance function.

Example: if we see increase of variance when plotted against the fitted values, a possible choice for the variance function would be: $\mathrm{Var}(\varepsilon_{ij}|\boldsymbol{b}_i) = \sigma \times |\mu_{ij}|^{2\delta}$

variance functions cont'

$$\mathrm{Var}(\varepsilon_{ij}|\boldsymbol{b}_i) = \sigma \times g^2(\mu_{ij}, \boldsymbol{\nu_{ij}}, \boldsymbol{\delta}), \quad i = 1, \ldots, M, \quad j = 1, \ldots, n_i \tag{22}$$

Problem: Within group error and random effects no longer independent.
Approximation scheme:

- fit without modeling heteroscedasticity. Calculate $\hat{\mu}_{ij} = \boldsymbol{x}_{ij}^T \beta + \boldsymbol{z}_{ij}^T \hat{\boldsymbol{b}}_i$ (i.e: estimate of model fit *given* the random effect)
- use the following approximation:

$$\mathrm{Var}(\varepsilon) \approx \sigma \times g^2(\hat{\mu}_{ij}, \boldsymbol{\nu_{ij}}, \boldsymbol{\delta}), \quad i = 1, \ldots, M, \quad j = 1, \ldots, n_i \tag{23}$$

  (i.e: assuming independence between within-group errors and random effects.)
- repeat until convergence

variance functions cont'

independence Assumption is core of approximation scheme:

$$\mathrm{Var}(\varepsilon) \approx \sigma \times g^2(\hat{\mu}_{ij}, \boldsymbol{\nu_{ij}}, \boldsymbol{\delta}), \quad i = 1, \ldots, M, \quad j = 1, \ldots, n_i \tag{24}$$

assuming independence between within-group errors and random effects.

intuition for independence approximation: if our estimate of $\mu_{ij} = \mathrm{E}[y_{ij}|\boldsymbol{b}i]$ is "good", plugging in $\hat{\mu}_{ij}$ instead of μ_{ij} won't change a lot. thus actually knowing b_i will not give you more information.

varFunc Classes

set of classes of variance functions, for specifying within-group variance models. varFunc-constructor functions available:

- 'varFixed': fixed variance (if variance is linear one covariate.)
- 'varIdent': different variances per stratum
- 'varPower: power of covariate or expected value
- 'varExp': exponential of covariate or expected value
- 'varConstPower': constant plus power of covariate or expected value
- 'varComb':combination of variance functions.

a closer look at varPower

Implemented variance model:

$$\mathrm{Var}(\varepsilon_{ij}) = \sigma^2 |\nu_{ij}|^{2\delta} \quad (25)$$

ν_{ij} may be a covariate or $\mathrm{E}[y_{ij}|\boldsymbol{b}_i]$.
call:

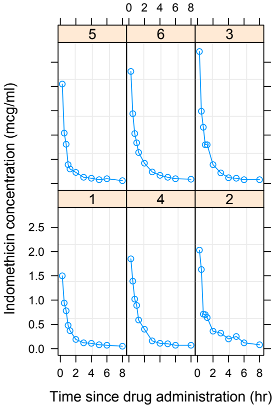
```
weight = varPower(value, form)
```

- value specifies *initial* value for δ.
- form specifies covariate or conditional expected value as right-hand-side formula.

Note: if covariate or conditional expected value takes on zero values, variance is undefined. $\Rightarrow$ use varConstPower
$\rightarrow$R

Optional example: Indomethicin Kinetics

Setup: 6 volunteers received intravenous injections of the same dose of indomethicin and had their plasma concentration measured 11 times.



The Model
already in log-transformed version:

$$y_{ij} = \phi_1 \exp[-\exp(\phi_2)t_j] + \phi_3 \exp[-exp(\phi_4)t_j] + \varepsilon_{ij} \qquad (26)$$

uses the SelfStarting function SSbiexp