NLME package in R

Jiang Qi

Department of Statistics
Renmin University of China

June 7, 2010
The problem

- Grouped data, or Hierarchical data: correlations between subunits within subjects.
The problem

- Grouped data, or Hierarchical data: correlations between subunits within subjects.
- It arises in many areas as diverse as agriculture, biology, economics, manufacturing, and geophysics.
The problem

- Grouped data, or Hierarchical data: correlations between subunits within subjects.
- It arises in many areas as diverse as agriculture, biology, economics, manufacturing, and geophysics.
- The most popular means to model Grouped data is Mixed Effect Model.
Grouped data, or Hierarchical data: correlations between subunits within subjects.

It arises in many areas as diverse as agriculture, biology, economics, manufacturing, and geophysics.

The most popular means to model Grouped data is Mixed Effect Model.

Mixed Effect Model decomposes the outcome of an observation as fixed effect (population mean) and random effect (group specific), and account for the correlation structure of variations among groups.
General formulation for Linear Mixed Effect Model (LME) described by Laird and Ware (1982):

\[ y = X_i \beta + Z_i b_i + \epsilon_i \]

where the fixed effects \( \beta \), random effects \( b_i \) occur linearly in the model.
The non-linear mixed effect model extends from the linearity assumption to allow for more flexible function forms.
The non-linear mixed effect model extends from the linearity assumption to allow for more flexible function forms.

At the first level, the $j$-th observation of $i$-th group is modeled as,

$$y_{ij} = f(t_{ij}, \phi_i)$$

At the second level, the parameter $\phi_i$ is modeled as,

$$\phi_{ij} = A_{ij}\beta + B_{ij}b_i,$$

where $b_i \sim N(0, \Psi)$, $\beta$ is a $p$-dimensional vector of fixed effects and $b_i$ is a $q$-dimensional random effects vector associated with the $i$-th group.
Data is contained in the NLME library, from a study of the cold tolerance of a C4 grass species. A total of 12 fourweek-old plants, 6 from Quebec and 6 from Mississippi, were divided into two groups: control plants that were kept at 26 degree and chilled plants that were subject to 14 h of chilling at 7 degree. Uptake rates (in μ mol/m2s) were measured for each plant at seven concentrations of ambient CO2 (μ L/L).
Carbon Dioxide Uptake

- Data is contained in the NLME library, from a study of the cold tolerance of a C4 grass species. A total of 12 fourweek-old plants, 6 from Quebec and 6 from Mississippi, were divided into two groups: control plants that were kept at 26 degree and chilled plants that were subject to 14 h of chilling at 7 degree. Uptake rates (in $\mu$ mol/m²s) were measured for each plant at seven concentrations of ambient CO2 ($\mu$ L/L).

- The objective of the experiment was to evaluate the effect of plant type and chilling treatment on the CO2 uptake.
> library(nlme)
> plot(CO2, outer = ~Treatment * Type, layout = c(4, 1))

**Figure:** CO2 uptake versus ambient CO2 by treatment and type f, 6 from Quebec and 6 from Mississippi. Half the plants of each type were chilled overnight before the measurements were taken.
An asymptotic regression model with an offset is used in Potvin et al (1990) to represent the expected CO2 uptake rate $U(c)$ as a function of the ambient CO2 concentration $c$:

$$U(c) = \phi_1 \{1 - \exp[-\exp(\phi_2)(c - \phi_3)]\}$$
> x = seq(2.8, 4, 0.01)
> f = function(a, b, c, x) {
+   return(a * (1 - exp(-exp(b) * (x - c))))
+ }
> y = f(1.5, 2, 3, x)
> plot(x, y, type = "l", ylim = c(-2, 3))
> grid(5, 5, lwd = 2)
When a standard non-linear model is inadequate for the data, and a mixed effect model is necessary to describe the discrepancy and correlation among the group structure?
Necessity for NLme

- When a standard non-linear model is inadequate for the data, and a mixed effect model is necessary to describe the discrepancy and correlation among the group structure?

- Some exploratory data analysis is indispensable, and R has provide us with a function `nlsList()` to help us realize this goal.
Necessity for NLme

When a standard non-linear model is inadequate for the data, and a mixed effect model is necessary to describe the discrepancy and correlation among the group structure?

Some exploratory data analysis is indispensable, and R has provide us with a function `nlsList()` to help us realize this goal.

`nlsList()` produce separate fits of a nonlinear model for each group in a groupedData object. The number of parameters are number of group $\times 3$. 
When a standard non-linear model is inadequate for the data, and a mixed effect model is necessary to describe the discrepancy and correlation among the group structure?

Some exploratory data analysis is indispensable, and R has provide us with a function \textit{nlsList()} to help us realize this goal.

\textit{nlsList()} produce separate fits of a nonlinear model for each group in a groupedData object. The number of parameters are number of group $\times 3$.

These separate fits by group are a powerful tool for model building with nonlinear mixed-effects models, because the individual estimates can suggest the type of random-effects structure to use.
A typical call to `nlsList` is `nlsList(model, data)`.
A typical call to \texttt{nlsList} is \texttt{nlsList(model, data)}.

Note that \texttt{nlsList()} requires initial value for the model.
A typical call to `nlsList` is `nlsList(model, data)`.

Note that `nlsList()` requires initial value for the model.

A `selfStart()` function, which specify the function formula and can automatically generate individual initial estimates for each group, is recommended. (See details in the book *Mixed Effects Models in S and S-Plus* by Pinheiro and Bates).
A typical call to `nlsList` is `nlsList(model, data)`.

Note that `nlsList()` requires initial value for the model.

A `selfStart()` function, which specify the function formula and can automatically generate individual initial estimates for each group, is recommended. (See details in the book «Mixed Effects Models in S and S-Plus» by Pinheiro and Bates).

In `nlme` library, C02 data has been assigned a `SSasympOff`.
> CO2.list <- nlsList(SSasympOff, CO2)

Call:
  Model: uptake ~ SSasympOff(conc, Asym, lrc, c0) | Plant
  Data: CO2
Coefficients:

<table>
<thead>
<tr>
<th></th>
<th>Asym</th>
<th>lrc</th>
<th>c0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Qn1</td>
<td>38.13978</td>
<td>-4.380647</td>
<td>51.22324</td>
</tr>
<tr>
<td>Qn2</td>
<td>42.87169</td>
<td>-4.665728</td>
<td>55.85816</td>
</tr>
<tr>
<td>Qn3</td>
<td>44.22800</td>
<td>-4.486118</td>
<td>54.64958</td>
</tr>
<tr>
<td>Qc1</td>
<td>36.42873</td>
<td>-4.861741</td>
<td>31.07538</td>
</tr>
<tr>
<td>Qc2</td>
<td>40.68370</td>
<td>-4.945218</td>
<td>35.08889</td>
</tr>
<tr>
<td>Qc3</td>
<td>39.81950</td>
<td>-4.463838</td>
<td>72.09422</td>
</tr>
<tr>
<td>Mn1</td>
<td>28.48285</td>
<td>-4.591566</td>
<td>46.97188</td>
</tr>
<tr>
<td>Mn2</td>
<td>32.12827</td>
<td>-4.466157</td>
<td>56.03863</td>
</tr>
<tr>
<td>Mn3</td>
<td>34.08481</td>
<td>-5.064579</td>
<td>36.40805</td>
</tr>
<tr>
<td>Mc1</td>
<td>13.55520</td>
<td>-4.560851</td>
<td>13.05675</td>
</tr>
<tr>
<td>Mc2</td>
<td>18.53506</td>
<td>-3.465158</td>
<td>67.84877</td>
</tr>
<tr>
<td>Mc3</td>
<td>21.78723</td>
<td>-5.142256</td>
<td>-20.39998</td>
</tr>
</tbody>
</table>

Degrees of freedom: 84 total; 48 residual
Residual standard error: 1.79822
The `intervals()` function returns the information about confidence intervals of estimates.

```r
> plot(intervals(CO2.list))
```

**Figure**: Confidence Intervals for parameters Asym, lrc.
The plot of the individual confidence intervals from fm1CO2.lis indicates that there is substantial between-plant variation in the parameter $\phi_1 = \text{Asym}$, and only moderate variation among $\phi_2 = \text{lrc}$, and $\phi_3 = \text{c0}$. 
The plot of the individual confidence intervals from fm1CO2.lis indicates that there is substantial between-plant variation in the parameter $\phi_1 = \text{Asym}$, and only moderate variation among $\phi_2 = \text{lrc}$, and $\phi_3 = \text{c0}$.

Fitting Nonlinear Mixed-Effects Models is appropriate.
The plot of the individual confidence intervals from fm1CO2.lis indicates that there is substantial between-plant variation in the parameter \( \phi_1 = \text{Asym} \), and only moderate variation among \( \phi_2 = \text{lrc} \), and \( \phi_3 = c0 \).

Fitting Nonlinear Mixed-Effects Models is appropriate.

However, we cannot decide the structure of random effects, so we fit a full mixed effect model with all the parameters as mixed effect. The model can be written:

\[
uij = \phi_{1i}1 - \exp[-\exp(\phi_{2i})(c_{ij} - \phi_{3i})] + \varepsilon_{ij}
\]

\[
\phi_{1i} = \beta_1 + b_{1i}, \phi_{2i} = \beta_2 + b_{2i}, \phi_{3i} = \beta_3 + b_{3i}
\]

\[
b_i \sim N(0, \Psi), \varepsilon_{ij} \sim N(0, \sigma^2)
\]
the object produced by \textit{nlsList()} CO2.list can be used to implement a full mixed effect model.

\begin{verbatim}
> CO2.nlme <- nlme(fm1CO2.lis)
> CO2.nlme

Nonlinear mixed-effects model fit by maximum likelihood
  Model: uptake ~ SSasympOff(conc, Asym, lrc, c0)
      . . . . . .

Random effects:
  Formula: list(Asym ~ 1, lrc ~ 1, c0 ~ 1)
  Level: Plant
  Structure: General positive-definite, Log-Cholesky parametrization

     StdDev Corr
     Asym  9.5105877 Asym  lrc
     lrc   0.1285620  -0.162
     c0    10.3742988  1.000  -0.140
     Residual  1.7665298
\end{verbatim}
A strong correlation between \(c_0\) and \(lrc\), suggests that only one of the random effects is needed.

\[
> \text{CO2.nlme2} \leftarrow \text{update(CO2.nlme, random = Asym + lrc} \sim 1) \]

Nonlinear mixed-effects model fit by maximum likelihood

Model: \(\text{uptake} \sim \text{SSasympOff(conc, Asym, lrc, c0)}\)

Data: \(\text{CO2}\)

Log-likelihood: \(-202.7583\)
Fixed: \(\text{list(Asym} \sim 1, \text{lrc} \sim 1, \text{c0} \sim 1)\)

\begin{align*}
\text{Asym} & \quad \text{lrc} & \quad \text{c0} \\
32.411764 & \quad -4.560265 & \quad 49.343573
\end{align*}

Random effects:
Formula: \(\text{list(Asym} \sim 1, \text{lrc} \sim 1)\)
Level: \(\text{Plant}\)
Structure: General positive-definite, Log-Cholesky parametrization

\begin{align*}
\text{StdDev} & \quad \text{Corr} \\
\text{Asym} & \quad 9.6593926 & \quad \text{Asym} \\
\text{lrc} & \quad 0.1995124 & \quad -0.777 \\
\text{Residual} & \quad 1.8079224
\end{align*}

Jiang Qi

NLME package in R
> anova(CO2.nlme, CO2.nlme2)

<table>
<thead>
<tr>
<th>Model</th>
<th>df</th>
<th>AIC</th>
<th>BIC</th>
<th>logLik</th>
<th>Test L.Ratio</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>CO2.nlme</td>
<td>1</td>
<td>422.6212</td>
<td>446.9293</td>
<td>-201.3106</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CO2.nlme2</td>
<td>2</td>
<td>419.5167</td>
<td>436.5324</td>
<td>-202.7583</td>
<td>1 vs 2</td>
<td>2.89549</td>
</tr>
</tbody>
</table>
Add covariates into mixed effect model

- Now the random effects accommodate individual deviations from the fixed effects.
Now the random effects accommodate individual deviations from the fixed effects.

The primary question of interest for the CO2 data is the effect of plant type and chilling treatment on the individual model parameters $\phi_i$. 
Add covariates into mixed effect model

- Now the random effects accommodate individual deviations from the fixed effects.
- The primary question of interest for the CO2 data is the effect of plant type and chilling treatment on the individual model parameters $\phi_i$.
- So, we want to know whether the variation between groups can be attributed as the plant type and chilling treatment.
Add covariates into mixed effect model

- Now the random effects accommodate individual deviations from the fixed effects.
- The primary question of interest for the CO2 data is the effect of plant type and chilling treatment on the individual model parameters $\phi_i$.
- So, we want to know whether the variation between groups can be attributed as the plant type and chilling treatment.
- A plot of predicted random effects of parameters against covariates is appropriate to excavate the correlation.
> CO2.nlmeRE <- ranef(CO2.nlme2, augFrame = T)

<table>
<thead>
<tr>
<th></th>
<th>Asym</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Qn1</td>
<td>6.17</td>
<td>0.048361985</td>
<td>Quebec</td>
<td>nonchilled</td>
<td>435</td>
<td>33.22857</td>
</tr>
<tr>
<td>Qn2</td>
<td>10.5325882</td>
<td>-0.172842971</td>
<td>Quebec</td>
<td>nonchilled</td>
<td>435</td>
<td>35.15714</td>
</tr>
<tr>
<td>Qn3</td>
<td>12.2180908</td>
<td>-0.057987100</td>
<td>Quebec</td>
<td>nonchilled</td>
<td>435</td>
<td>37.61429</td>
</tr>
<tr>
<td>Qc1</td>
<td>3.3521234</td>
<td>-0.075586358</td>
<td>Quebec</td>
<td>chilled</td>
<td>435</td>
<td>29.97143</td>
</tr>
<tr>
<td>Qc3</td>
<td>7.4743083</td>
<td>-0.192416381</td>
<td>Quebec</td>
<td>chilled</td>
<td>435</td>
<td>32.58571</td>
</tr>
<tr>
<td>Qc2</td>
<td>7.9284657</td>
<td>-0.180323624</td>
<td>Quebec</td>
<td>chilled</td>
<td>435</td>
<td>32.70000</td>
</tr>
<tr>
<td>Mn3</td>
<td>-4.0733486</td>
<td>0.033449394</td>
<td>Mississippi</td>
<td>nonchilled</td>
<td>435</td>
<td>24.11429</td>
</tr>
<tr>
<td>Mn2</td>
<td>-0.1419773</td>
<td>0.005645756</td>
<td>Mississippi</td>
<td>nonchilled</td>
<td>435</td>
<td>27.34286</td>
</tr>
<tr>
<td>Mn1</td>
<td>0.2406596</td>
<td>-0.193859245</td>
<td>Mississippi</td>
<td>nonchilled</td>
<td>435</td>
<td>26.40000</td>
</tr>
<tr>
<td>Mc2</td>
<td>-18.7991627</td>
<td>0.319367709</td>
<td>Mississippi</td>
<td>chilled</td>
<td>435</td>
<td>12.14286</td>
</tr>
<tr>
<td>Mc3</td>
<td>-13.1168244</td>
<td>0.299428913</td>
<td>Mississippi</td>
<td>chilled</td>
<td>435</td>
<td>17.30000</td>
</tr>
<tr>
<td>Mc1</td>
<td>-11.7865217</td>
<td>0.166761922</td>
<td>Mississippi</td>
<td>chilled</td>
<td>435</td>
<td>18.00000</td>
</tr>
</tbody>
</table>

Use the general function `plot()`. A onesided formula on the right-hand side, with covariates separated by the * operator, results in a dotplot of the estimated random effects versus all combinations of the unique values of the variables named in the formula.
> plot(CO2.nlmeRE, form = ~ Type * Treatment )

**Figure:** Predicted random effects Virus Covariates
The figure shows a strong relationship between the parameter asym and the covariates: Asym decreases when the plants are chilled and is higher among Quebec plants than Mississippi plants.
Add covariates into mixed effect model

- The figure shows a strong relationship between the parameter asym and the covariates: Asym decreases when the plants are chilled and is higher among Quebec plants than Mississippi plants.
- The increase in Asym from chilled to nonchilled plants is larger among Mississippi plants than Quebec plants, suggesting an interaction between Type and Treatment.
Add covariates into mixed effect model

- The figure shows a strong relationship between the parameter asym and the covariates: Asym decreases when the plants are chilled and is higher among Quebec plants than Mississippi plants.
- The increase in Asym from chilled to nonchilled plants is larger among Mississippi plants than Quebec plants, suggesting an interaction between Type and Treatment.
- We include both covariates in the model to explain the Asym plant-to-plant variation.
> CO2.nlme3 <- update(CO2.nlme2, fixed = list(Asym ~ Type * 
+ Treatment, lrc + c0 ~ 1), start = c(32.412, 0, 0, 0, -4.5603, 
+ 49.344))
> summary(CO2.nlme3)

Nonlinear mixed-effects model fit by maximum likelihood

  Model: uptake ~ SSasympOff(conc, Asym, lrc, c0)
  Data: CO2

              AIC     BIC  logLik
  393.6765 417.9847 -186.8383

Random effects:
  Formula: list(Asym ~ 1, lrc ~ 1)
  Level: Plant

Structure: General positive-definite, Log-Cholesky parametrization

          StdDev   Corr
  Asym.(Intercept) 2.9298913 As.(I)
  lrc            0.1637446  -0.906
  Residual        1.8495592

Jiang Qi

NLME package in R
### Fixed effects: list(Asym ~ Type * Treatment, lrc + c0 ~ 1)

<table>
<thead>
<tr>
<th>Term</th>
<th>Value</th>
<th>Std.Error</th>
<th>DF</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asym.(Intercept)</td>
<td>42.17337</td>
<td>1.345946</td>
<td>67</td>
<td>31.33364</td>
</tr>
<tr>
<td>Asym.TypeMississippi</td>
<td>-11.82431</td>
<td>1.537304</td>
<td>67</td>
<td>-7.69159</td>
</tr>
<tr>
<td>Asym.Treatmentchilled</td>
<td>-5.23756</td>
<td>1.464784</td>
<td>67</td>
<td>-3.57565</td>
</tr>
<tr>
<td>Asym.TypeMississippi:Treatmentchilled</td>
<td>-4.78084</td>
<td>2.353801</td>
<td>67</td>
<td>-2.03111</td>
</tr>
<tr>
<td>lrc</td>
<td>-4.58925</td>
<td>0.084821</td>
<td>67</td>
<td>-54.10511</td>
</tr>
<tr>
<td>c0</td>
<td>49.48163</td>
<td>4.456630</td>
<td>67</td>
<td>11.10293</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>p-value</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Asym.(Intercept)</td>
<td>0.0000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asym.TypeMississippi</td>
<td>0.0000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asym.Treatmentchilled</td>
<td>0.0007</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asym.TypeMississippi:Treatmentchilled</td>
<td>0.0462</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>lrc</td>
<td>0.0000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c0</td>
<td>0.0000</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### As.(I) Asy.TM Asym.T A.TM:T lrc

<table>
<thead>
<tr>
<th>Term</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Asym.TypeMississippi</td>
<td>-0.559</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asym.Treatmentchilled</td>
<td>-0.540</td>
<td>0.471</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asym.TypeMississippi:Treatmentchilled</td>
<td>0.252</td>
<td>-0.640</td>
<td>-0.622</td>
<td></td>
</tr>
<tr>
<td>lrc</td>
<td>-0.540</td>
<td>0.056</td>
<td>-0.016</td>
<td>0.132</td>
</tr>
<tr>
<td>c0</td>
<td>-0.086</td>
<td>-0.001</td>
<td>-0.036</td>
<td>0.063</td>
</tr>
</tbody>
</table>
> CO2.nlmeRE3 <- ranef( CO2.nlme3, aug = T )
> plot( CO2.nlmeRE3, form = ~ Type * Treatment )

**Figure:** Predicted random effect Virus Covariate
After covariates have been introduced in the model to account for intergroup variation, a natural question is which random effects are still needed.
After covariates have been introduced in the model to account for intergroup variation, a natural question is which random effects are still needed.

The ratio between a random-effects standard deviation and the absolute value of the corresponding fixed effect gives an idea of the relative intergroup variability for the coefficient, which is often useful in deciding which random effects should be tested for deletion from the model.
A final assessment of the quality of the fitted model is provided by the plot of the augmented predictions.

```r
> plot( augPred(CO2.nlme3, level = 0:1),layout = c(6,2) )
```

**Figure:** Predicted value of uptake by groups
Thank you for your listening!